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, Giorgio Chiovelli, 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.

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

ABSTRACT

We examine the long-run consequences of ethnic partitioning, a neglected aspect of the Scramble for Africa, 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 many observable characteristics. Second, using georeferenced data on political violence, that include both state-driven conflict and violence against civilians, we find 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 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. Fifth, using individual-level data we document that respondents identifying with split groups have lower access to public goods and worse educational outcomes. 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 development 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 beginning of the colonial rule, another major event took place that according to the African historiography had malicious long-lasting consequences. During the "Scramble for Africa" -that starts with the Berlin Conference of $1884-1885$ and is completed by the turn of the 20th century- 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 limited knowledge of local conditions. Despite their arbitrariness, boundaries outlived the colonial era. As a result in many African countries today a significant fraction of the population belongs to ethnic groups that are partitioned among different states.[1]

Many African scholars (e.g., Asiwaju (1985), Wesseling (1996), Dowden (2008), Thomson (2010)) have maintained that the main channel of Europeans' influence on development was not colonization per se, but the improper border design. Herbst (2000) succinctly summarizes the argument: "*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 and conflict primarily by splitting groups across the newly-minted African states. Horowitz (1985) argues that ethnic partitioning led to irredentism and helped create an ideology of secession and nationalism. Moreover, split groups have often been instrumentally used by governments to destabilize neighboring countries, setting the stage for discrimination of split ethnicities in the political sphere and the eruption of ethnic wars.

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 states (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 neighboring nations, perform economically worse (Alesina, Easterly, and Matuszeski (2011)). Nevertheless, to the best of our knowledge there is no empirical work directly exploring the consequences of ethnic partitioning for African groups (the relevant unit of analysis), where the arbitrary border design and the large number of split groups offer the opportunity to cleanly identify the impact of partitioning. This study is a step in this direction.

[1] Asiwaju (1985) identifies 177 partitioned ethnicities. Englebert, Tarango, and Carter (2002) estimate that partitioned groups constitute on average 40% of the total population; Alesina, Easterly, and Matuszeski (2011) estimate that in several African countries the percentage of the population that belongs to a 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%)).

Results To formally assess the claim that African borders were drawn with little respect to the local political geography, we investigate whether partitioned ethnicities differ systematically from non-split groups across several geographic-ecological traits. With the exceptions of the land mass of the historical ethnic homeland and the presence of lakes, there are no significant differences between split and non-split homelands along a comprehensive set of covariates. And there are no systematic differences across several pre-colonial, ethnic-specific, institutional, cultural, and economic features, such as the size of the settlements, the type of subsistence economy, and proxies of pre-colonial conflict. These results offer support to a long-standing assertion within the African historiography regarding the largely arbitrary nature of African borders, 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 georeferenced information for the $1997 - 2013$ period on incidents of political violence, including battles between government forces, rebels and militias and violence against civilians, we document that civil conflict is higher in the homelands of partitioned ethnicities. This applies to conflict intensity, duration, casualties, and the likelihood of conflict. Our estimates suggest that conflict intensity (likelihood) is approximately 40% (8%) higher in areas where partitioned ethnicities reside, as compared to homelands of ethnicities that have not been separated by national borders. The results are similar when we restrict estimation to ethnic homelands near the national borders.

We then exploit the richness of the data to examine what type of conflict is more likely to afflict partitioned homelands. In line with the thesis put forward by African historians, that split groups are often used by neighboring countries to stage proxy wars and destabilize the government on the other side of the border, we find that military interventions from adjacent countries are more common in the homelands of partitioned groups, rather than in nearby border areas where non-split groups reside. We also examine the impact of ethnic partitioning on the different forms of political violence. Partitioning matters crucially for two-sided conflict between government troops and rebel groups "*whose goal is to counter an established national governing regime by violent acts*" and to a lesser extent with one-sided violence against civilians. These patterns are corroborated with a different georeferenced conflict database (Uppsala Conflict Data Program Georeferenced Event Dataset, UCDP - GED) that records only deadly events associated with civil wars. In contrast, there is no link between ethnic partitioning and riots and protests, which are predominantly a capital-city phenomenon; and there is no association between partitioning and conflict between non-state actors. These results are in accord with African historiography pointing out that partitioned groups face discrimination from the national government and often engage in rebellions (often with the support of their co-ethnics on the other side of the border) to counter repression.

In an attempt to dig deeper on the partitioning - repression - civil war nexus we use the Ethnic Power Relations (EPR) dataset (Wimmer, Cederman, and Min (2009)) that offers an assessment of formal and informal degrees of political participation of ethnic groups in the political arena over the post-independence

period. The within-country analysis shows that partitioned ethnicities are significantly more likely (11% – 14% increased likelihood) to engage in 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 7 percentage points higher compared to non-split groups.

We complement the group-based and the location-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 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). This applies both to respondents residing in their ethnicity's ancestral homeland and to individuals residing outside of it (both in non-split and in partitioned ethnic homelands).

Related Literature Our paper belongs to the genre of studies that investigate the historical origins of comparative development (see Nunn (2014) for a review). The literature has mainly focused on the impact of colonization via institutions (e.g., Acemoglu, Johnson, and Robinson (2005), Acemoglu, Reed, and Robinson (2014)), infrastructure (e.g., Huillery (2009), Jedwab and Moradi (2015)), and human capital (e.g., Easterly and Levine (2015), Wantchekon, Klasnja, and Novta (2015)). We emphasize instead an aspect of the colonial legacy that has been largely neglected by economics research: the drawing of political boundaries in the end of the 19th century that resulted in a large number of partitioned ethnicities after independence. As such our work is related to Alesina, Easterly, and Matuszeski (2011), who show that countries with more straight-line-like borders and nations where a significant part of their population also resides in different countries underperform economically.

A related body of research traces the origins of African countries' weak state capacity to the pre-colonial period. Nunn (2008) and Nunn and Wantchekon (2011) document that the slave trades (1400 – 1900) have shaped development by spurring ethnic conflict and lowering trust. Gennaioli and Rainer (2006, 2007) and Michalopoulos and Papaioannou (2013) show that pre-colonial political centralization at the group level is a significant correlate of contemporary development both across and within countries. Our paper relates to these contributions, as we also study the long-run implications of historical legacies focusing on ethnic traits. Yet, rather than studying pre-colonial features, we examine the impact of ethnic partitioning during colonization. Assessing the impact of ethnic-specific characteristics in Africa is crucial, as Michalopoulos and Papaioannou (2014) show that states' capacity to broadcast power within a country rapidly diminishes for regions further from the capitals (Herbst (2000)).[2]

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), Blattman and Miguel (2010)

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

for reviews, and Collier and Sambanis (2005) for case studies in Africa). Of most relevance are works studying the role of ethnic heterogeneity. Since the influential work of Easterly and Levine (1997), Africa's underdevelopment and conflict intensity has been linked to its widespread ethnolinguistic diversity. While the correlation between ethnic fragmentation and civil war is weak (Fearon and Laitin (2003)), ethnic polarization (Montalvo and Reynal-Querol (2005), Esteban, Mayoral, and Ray (2012)), and inequality across and within ethnic lines (Huber and Mayoral (2014), Esteban and Ray (2011)) correlates significantly with civil conflict. And a growing literature in political science (and recently in economics) shows the prevalence of ethnic politics, ethnic discrimination and repression from the central government, and poor public goods provision across all parts of the continent (Posner (2005), Franck and Rainer (2012), Hodler and Raschky (2014), Luca, Hodler, Raschky, and Valsecchi (2015), Burgess, Jedwab, Miguel, Morjaria, and Padro-i-Miguel (2015)). Moreover, Wimmer, Cederman, and Min (2009) show that the likelihood of ethnic conflict increases when groups are excluded from national power.

We complement this research by uncovering that ethnic minorities partitioned across Africa's borders present a much greater problem for governance than non-split groups. Because split ethnicities are more capable of organizing rebellions through assistance from co-ethnics across the border, armed conflict between partitioned groups and the governments are more likely. We show that the heightened propensity of split groups to participate in conflict is particularly strong for ethnicities and periods when excluded from the central government. This finding is consistent with Fearon and Laitin (2003) who link conflict onset to opportunity cost rather than grievances. Moreover, our finding that foreign interventions from neighboring countries are more common in the homelands of partitioned ethnicities implies that the latter serve as vehicles of instability.

The correlations found in studies linking cross-country variation in border features and ethnic composition to development proxies (income or conflict) are informative (e.g., Alesina, Easterly, and Matuszeski (2011), Englebert, Tarango, and Carter (2002), Bosker and de Ree (2014)), but they cannot be easily 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 determines both economic performance and conflict. As the recent literature on state capacity shows, nation building, development, and conflict are inter-linked and jointly determined by hard-to-account-for factors related to the societal structure, geography, and historical legacies (Besley and Persson (2011b)). Thus, selection, reverse causality, 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 some of the shortcomings of cross-country works. First, by showing that there are no systematic differences in geographic, economic, and cultural characteristics between split and non-split ethnic homelands, our analysis offers large-scale econo-

metric evidence on the accidental nature of most African borders, at least with respect to the ethnic partitioning dimension.[3] Second, using information on the spatial distribution of ethnicities in the end of 19th century, well before the current national boundaries came into effect, alleviates concerns related to the migratory flows ignited by the border design itself. Since borders were drawn by Europeans with limited respect to local conditions and did not change at independence, we focus on cases where country boundaries were not the result of political, economic, and military developments. Third, focusing on ethnic groups is conceptually appealing in the context of Africa, where ethnic identification is strong, ethnic segregation high and political violence 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 group-level data from the Ethnic Power Relations Database on political discrimination and ethnic wars as well as individual-level data from the DHS we shed some light on the potential mechanisms at work. In this regard our empirical study builds on Besley and Persson (2011a), who stress the need to jointly study one-sided violence (repression), two-sided violence (civil war), and public goods.

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 the different aspects of conflict affecting partitioned homelands, so as to shed light on the potential mechanisms at work. In Section 6 we explore the connection between partitioning, ethnic-based discrimination from the national government and ethnic wars. Section 7 presents the individual-level analysis linking education and access to public 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 West Africa, signing bilateral agreements on spheres of influence. During the next 40 years,

[3] Admittedly, 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 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 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 ethnic 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. Asiwaju (1985) notes that "*the Berlin conference, despite its importance for the subsequent history of Africa, was essentially a European affair: there was no African representation, and African concerns were, if they mattered at all, completely marginal to the basic economic, strategic, and political interests of the negotiating European powers*". In many cases, European leaders were in such a rush that they did not wait for the information arriving from explorers, geographers, and missionaries. As the British prime minister at the time Lord Salisbury (Robert Cecil) put it, "*we have been engaged in drawing lines upon maps where no white man's feet have ever tord; we have been giving away mountains and rivers and lakes to each other, only hindered by the small impediment that we never knew exactly where the mountains and rivers and lakes were.*" Asiwaju (1985) summarizes that "*the study of European archives supports the accidental rather than a conspiratorial theory of the marking of African boundaries.*" 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 largely accidental border design. First, at the time Europeans had little knowledge of local geography, 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.[4] 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.[5] Fourth, Europeans were unwilling to change colonial borders despite new information arriving from the ground.[6] Fifth, as locals could freely move across colonial borders, African chiefs did not oppose much the colonial design, as little changed on the

[4] 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.*"

[5] Poor demarcation and imprecise colonial treaties of the exact boundaries have contributed to conflict after independence. Examples include the war between Tanzania and Uganda in 1978 over the Kagera region (a 1800 km^2 strip of land) and the conflict between Burkina Faso and Mali over the Agacher strip in 1985.

[6] 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.*"

ground. Asiwaju (1985) cites the Ketu king, saying that "*we regard the boundary (between Benin-Dahomey and Nigeria) separating the English and the French, not the Yoruba.*" Wesseling (1996) summarizes the situation: "*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 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 departing Europeans did not touch this issue. The leaders of the newly-crafted African states believed that nation-building and industrialization would sideline ethnic divisions. African 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 facilitates causal inference.

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.[7] In line with this argument, Wimmer, Cederman, and Min (2009) estimate that 20% of all civil wars in Africa have a secessionist component.[8] While, compared to the number of civil wars in Africa, there have been few instances of secession (Englebert (2009)), irredentism and the associated ideology have played an important role in some major conflicts, mostly in Somalia, Mali, and Senegal. Somalis, for example, were split during colonization between four different European colonies, while Ethiopia also got a slice, the Ogaden region which is almost exclusively occupied by Somalis. The five-pointed star in the flag of Somalia symbolizes the desire of unifying the five regions inhabited by Somali clans (Italian

[7] Horowitz (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.*"

[8] Civil 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.

Somaliland, Northern Kenya, Southern Ethiopia, French Somaliland - Djibouti, and British Somaliland); three long-lasting wars have been partly driven by the desire of Somalis in Ethiopia to become part of Somalia (Meredith (2005)). UCDP describes the event as follows: *"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."* Similarly, in the initial years after independence Kenya experienced conflict in the Northern Frontier District when Somali insurgents fought for annexation to Somalia (Touval (1967)). In Section 8 of the Supplementary Appendix we discuss in detail the case of the partitioning of the Somalis and perform a counterfactual analysis of its impact on conflict.

Repression 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 within their ancestral homeland on the other side of the border. Members of 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. Thus quite often episodes of repression lead to civil wars, as partitioned groups have a lower opportunity cost of conflict. Moreover, the instrumental use of split ethnicities by neighboring governments provides a pretext for their inferior treatment by home governments.

The recurrent conflict in the Casamance region in Southern Senegal, where the partitioned Diola (Jola) and some smaller groups reside, offers an illustration. As Gambia effectively splits Senegal, Casamance is disconnected from the central government in Dakar. Moreover, Casamance was ruled independently from the rest of Senegal for most of the colonial time. Locals objected to the land reform of 1964 that transferred to the state all non-registered land, effectively transferring property to the capital over local ethnic groups, that had communal property rights. The violent riots in 1980 were soon followed by the formation of the separatist, "Movement of the Democratic Forces of Casamance (MDFC)" in 1982. While initially MDFC used low-level violence, in the 1990*s* conflict intensified as MDFC was supported by Guinea-Bissau and Gambia, where the Diola exert significant influence. The Senegalese government has accused the Gambian President Yahya Jammeh, a Diola himself, and Guinea-Bissau's army for assisting MDFC insurgents, providing them with arms and shelter (Humphreys and Mohamed (2005), Evans (2004)).[9] Moreover, MFDC rebels from Senegal participated in the 1998 civil war in Guinea-Bissau, aiding General Mane in his efforts to dispose President Vieira (Wagane (2006)).

[9]There is a debate whether MFDC is a Jola-based irredentist movement or it reflects the aspirations of other groups in the region. MFDC has consistently asserted that it represents all Casamance groups, denying accusations from the central government that it is a Diola movement.

Spillovers Population displacements across the border are more common within split groups. Such refugee flows, however, may change the ethnic composition in adjacent countries fomenting conflict. A pertinent example is the Alur, a group partitioned 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, many Alur were pushed to Uganda. This in turn generated opposition from the Buganda 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*".[10]

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. Border drawing shaped each and every country-specific geographic and cultural characteristic including a country's ethnic heterogeneity, land size, and access to the coast that affect development. Herbst (2000) argues that civil conflict is more pervasive in large African countries because their size limits their ability to broadcast power across their territories. Collier (2007) discusses how the border design resulted in Africa having the largest proportion of landlocked countries hampering their growth potential. While our analysis focuses on a single aspect of the Scramble for Africa, that of ethnic partitioning, by exploiting within-country variation we are able to account for common–to–all–homelands, country-specific characteristics. Moreover, in the Supplementary Appendix, we examine how these different nationwide by-products of the border design interact with ethnic partitioning in influencing conflict intensity.

3 Ethnic Partitioning and Border Artificiality

3.1 Identifying Partitioned Ethnic Groups

We identify partitioned groups projecting contemporary country 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 the European colonization in the late 19th and early 20th century (Figure $1a$).[11] 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 the conflict databases do not cover the Comoros. This leaves us with 833 groups. We also exclude 8 regions where population according

[10] Fearon 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.).

[11] Murdock's map is based on primary sources covering the period $1860-1940$. Most observations correspond to 1890, 1900, and 1910. After intersecting ethnic boundaries with country borders, we drop ethnicity-country polygons of less than 100 km^2, as such small areas are most likely an outcome of error in the underlying mapping of ethnicities.

to the earliest post-independence census is zero.[12] So our analysis focuses on 825 ethnicities.

Figure 1*a*

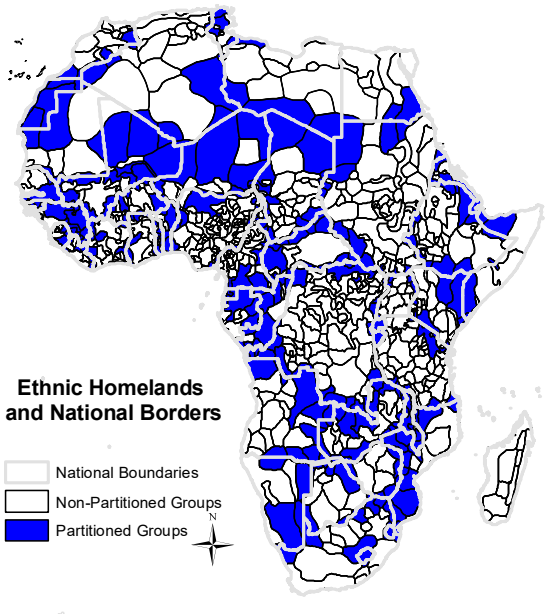

Figure 1*b*

The homeland of 357 groups falls into more than one country. Yet for several of these groups the overwhelming majority of their ancestral land (usually more than 99%) belongs to a single country. For example, 99.5% of the area of the Ahaggaren falls into Algeria and only 0.5% 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 1*b*).[13] Appendix Table *A* lists partitioned ethnicities. When we use a broader threshold of 5% we identify 266 partitioned groups.

Our procedure identifies most major ethnic groups that have been split by the African borders. For example, the Maasai are partitioned between Kenya and Tanzania (62% and 38%, respectively), the Anyi between Ghana and the Ivory Coast (58% and 42%), and the Chewa between Mozambique (50%), Malawi (34%), and Zimbabwe (16%). Other examples include the Hausa (split between Nigeria and Niger) and the Ewe (split by the Togo-Ghana border). We also checked whether our coding is in line with Asiwaju (1985), who provides the only comprehensive (to our knowledge) codification of partitioned African groups. Our strategy identifies almost all ethnic groups that Asiwaju (1985) lists as partitioned.[14]

[12] These 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.

[13] We apply the same threshold, as in our previous 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.

[14] Our approach of identifying split groups is imperfect. Ethnic groups' homelands partially overlap and there is certainly

It is perhaps instructive to assess how much of the cross-country variation in ethnic diversity in Africa can be attributed to ethnic partitioning. In this regard, we estimated simple cross-country regressions linking the widely-used ethnic fragmentation measures (of Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003) and Desmet, Ortuño-Ortín, and Wacziarg (2012)) to the log number of partitioned groups in a country (with and without controls for size); we find that approximately a fourth to a third of the cross-country variation of the measures of ethnic diversity can be accounted for by partitioned ethnicities.

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, and Burundi. And some borders were delineated in the early 20th century, when Europeans conceivably had some knowledge of local conditions.[15] Moreover, some contemporary borders in Western Africa follow the French administrative divisions. And in some cases (Cameroon-Nigeria; Ghana-Togo) there were referenda on the redrawing of these border segments at independence. Yet what is key for establishing causality is not that all borders were randomly drawn (though many were); what is needed for causal inference is that there are no systematic differences between partitioned and non-split ethnic homelands with respect to (un)observable characteristics that may independently affect contemporary conflict.

In this section we examine whether there are significant differences between the two sets of ethnicities across a host of observable traits. We estimate simple (linear probability) models associating the binary ethnic partitioning index ($SPLIT$) with various geographic, ecological, natural resource variables and proxies of pre-colonial conflict and development.[16] 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.[17]

noise in Murdock's map. As such our partitioning index is noisy. For example, our procedure identifies as non-split the Ogaden (it enters as partitioned when we adopt the 5% threshold) and the Sab groups in Ethiopia. Our readings suggest that these groups have been impacted by the Ethiopian-Somali border. Since our classification is solely based on the intersection of the historical tribal map with the contemporary country boundaries, such errors are unlikely to be systematic (correlated with contemporary conflict or the key controls). In presence of classical measurement error our estimates will be attenuated.

[15] Yet our reading suggest that even in cases where European were aware of borders splitting ethnicities (as in the case of the Abyssinia-Ethiopia border), this did not seem to factor in their decisions.

[16] Appendix 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.

[17] Cameron, Gelbach, and Miller (2011) explicitly cite spatial correlation as an application of the multi-way clustering method. Murdock (1959) assigns the 833 ethnicities into 96 ethnolinguistic clusters. We also used the method of Conley (1999) to account for spatial dependence of an unknown form, finding similar standard errors.

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 were more likely to be partitioned. In column (2) we augment the specification with two dummy variables that identify ethnic homelands with a large lake and a main river, respectively. The coefficient on the lake dummy is positive and significant at the 10% level, while the river indicator enters with a small and statistically insignificant coefficient. These results are in accord with the narrative of Europeans attempting to use natural barriers while delineating spheres of influence, apparently with limited success. 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 using a malaria 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 predicts partitioning. Both variables enter with insignificant estimates. In column (5) we include indicators identifying ethnic areas with diamond mines and petroleum. While in the initial phase of colonization Europeans were mostly interested in agricultural goods and minerals, adding these indicators allows investigating whether partitioned and non-split groups differ across these aspects that correlate with contemporary 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 adjacent ethnicities that are of the same ethnolinguistic family, to examine whether Europeans took into account broad cultural differences when delineating the 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 to a lesser extent the presence of lakes) correlates with ethnic partitioning.

Pre-colonial Features and Ethnic Partitioning While at the time of the colonial border design Europeans had limited understanding of local political geography, it is useful to examine the association between ethnic partitioning and pre-colonial conflict, as recent cross-country works (Fearon and Laitin (2012)) and cross-regional studies reveal a legacy of conflict from the pre-colonial times to the present (Besley and Reynal-Querol (2014), Nunn and Wantchekon (2011), Depetris-Chauvin (2014)). Table 1 - Panel *B* examines the association between ethnic partitioning and proxies of pre-colonial conflict.

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 Mali in the end of the 15th century or the war between the Banyoro and Buganda kingdoms around 1600 AD). Specification (1) shows the lack of a systematic association between ethnic partitioning and pre-colonial violence, as reflected in an indicator that takes the value one for ethnic homelands that experienced conflict over the period $1400-1700$. Column (2) shows that ethnic partitioning and proximity to the nearest pre-colonial conflict are not related (the results are similar with log distance). These results suggest 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 conducted by members of different groups or even between members of the same ethnicity*" (Nunn and Puga (2012)). Djankov and Reynal-Querol (2010) present cross-country evidence of a positive 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 proximity of a group to a large pre-colonial kingdom, using data from Besley and Reynal-Querol (2014). There is no systematic association between ethnic partitioning and the group being part of a large kingdom or the distance to the centroid of the closest pre-colonial kingdom. So, pre-colonial political centralization, that has been found to confer long-lasting beneficial effects on regional development (Michalopoulos and Papaioannou (2013)) does not seem to correlate with partitioning.

In column (7) we associate ethnic partitioning to the pre-slave trade level of 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 disproportionately impacted by the border design.

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

These checks corroborate that in the beginning of the colonial era, apart from a group's landmass, there were no differences between split and non-split groups. However, one would like to verify that also ex-post, i.e., after the borders were set, the resulting split groups within a country are no different than non-split ones. In Appendix Table 9 we report "balancedness tests" along various geographic, ecological, and natural resource characteristics both for the full sample of country-ethnic homelands and for the country-ethnic homelands close to the national border. The "similarity regressions" show that within countries with the exception of (log) land area for groups close to the border, there are no systematic differences in numerous observable characteristics between split and non-split groups.

Summary Our results are consistent with the historical account on the largely arbitrary nature of African borders. Yet, they do not imply that *all* African borders were randomly designed, something that is not the case. The 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 political violence events across all African countries from 1997 to 2013. Political violence is defined as the use of force by a group with a political purpose or motivation. ACLED is by far the most complete georeferenced conflict dataset; and while the data are noisy they have several desirable features.[18]

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 rampant, but these incidents are often deadly, economically harmful, and devastating for the victims and the local community.

Second, ACLED categorizes conflict into four groups, allowing for a finer decomposition. The main categories are (percentage of total events): (1) Battles, either without change of control (32%) or where rebels or government troops gain control (4%); (2) Violence against 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, so, we can study the impact of partitioning on conflict intensity. Battles and violence against civilians are by far the most deadly types, as 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 events are classified by the main conflict actors (government, rebels, militias, foreign interventions) allowing us to examine whether partitioning is mostly linked to state-driven violence and interventions from nearby 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 more war-prone nations data come from around twenty sources. This diversity assuages concerns of systematic biases in reporting from government

[18] Parallel 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).

controlled media. The data are mostly based on international sources, such as the BBC (around 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 the *Human Rights Watch* and *Amnesty International*, and the United Nations.[19] 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.

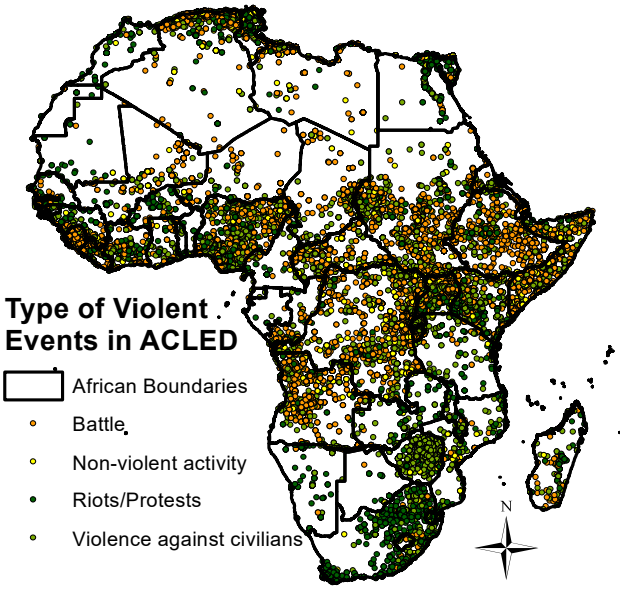

Figure 2*a*

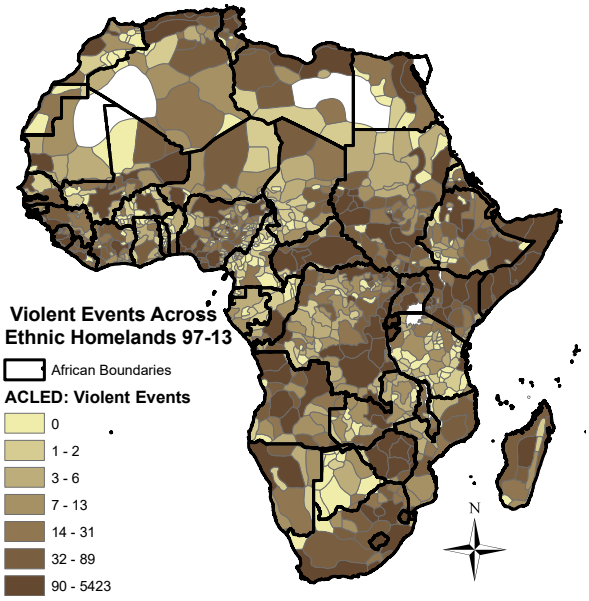

Figure 2*b*

Figure 2*a* illustrates the spatial distribution of conflict events. The map plots 64,650 precisely georeferenced incidents of political violence. In total there are 79,765 recorded events, but given the spatial nature of our study, we drop events where the location of the conflict is not accurately known. 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 Sierra Leone. Violence is also pervasive in Somalia, Ethiopia, and Zimbabwe. In contrast, there are few events in Botswana, Zambia, Tanzania, Namibia, and Gabon. There is also considerable variation within countries. For example, while conflict incidence in Tanzania is low, there are 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. Battles and violence against civilians are correlated, but the correlation is far from perfect (0.64; see Appendix Table 7). For example, in Zimbabwe we observe lots of violence against

[19] Going over the documentation it seems that the data are based on verified information and not simply the reproduction of state-press releases. For example, in Zimbabwe, most events come from the BBC, Reuters, and the Zimbabwe Human Rights NGO Forum, a coalition of 19 NGOs that get data from their representatives on the ground.

civilians (3,701 incidents) and few battles (59). Conversely, in Ethiopia and Angola we predominantly observe conflict between government troops and rebels rather than violence against civilians.

To construct conflict intensity at the country-ethnic homeland level, we project ACLED's mapping (Figure 2*a*) on the intersection of Murdock's ethnolinguistic map with contemporary borders (Figure 1*b*). Figure 2*b* portrays the spatial distribution of all conflicts 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 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, the part of Lobi's homeland in Ivory Coast is assigned to Ivory Coast, while Lobi's land mass in Burkina Faso gets a Burkina Faso indicator. At the country-ethnic homeland level, we have 518 partitioned areas and 694 non-split homelands.[20] Given the lack of systematic association 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 partitioning, we augment the specification with a spillover index ($SPIL$), reflecting the fraction of adjacent groups in the same country 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 groups. [The mean (standard deviation) of $SPIL$ is 0.41 (0.32).]

The conditioning set, $X'_{i,c}$, follows Michalopoulos and Papaioannou (2013, 2014) and other related works (e.g., Fenske (2013, 2014)) and includes log land area, log population according to the first post-independence census, indicators for the presence of rivers and lakes and several geographic, ecological, and natural resource measures. a_c denotes country-specific constants that account for countrywide factors that may affect conflict, related to the type of colonial rule, colonial and contemporary 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)).[21] The negative binomial model ac-

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

[21] Due 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.

counts 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.[22] 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 dropping the top 5% of the dependent variable. To isolate the impact of ethnic partitioning on the likelihood of conflict, we always report linear probability model (LPM) estimates where the dependent variable is an indicator that takes on the value one if a country-ethnic area has been affected by conflict over the sample period. 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 specifications. Panel *A* gives (unconditional) NB-ML estimates with country-specific constants focusing on conflict events, while Panel *B* gives country-fixed effects LS 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 columns (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.[23] 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 an indicator for the presence of a major city in 1400. The coefficient on the ethnic 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) and in column (6) we exclude homelands where capitals fall. The estimates imply that partitioned ethnicities experience an increase of approximately 145 log points in the number of conflict incidents. This translates into an 57% increase in political violence ($\exp(0.45) - 1 = 0.568$). The effect of ethnic partitioning on 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 in some specifications is significant at the 90% level. This implies that the negative

[22] This 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), among others.

[23] Distance to the coast 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 in almost all specifications.

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 (61.3 km). This allows us to compare 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; this assures 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.[24]

Table 2-Panel B reports LPM estimates with country fixed effects. Looking at the "extensive" margin accounts for the non-linear nature of the dependent variable; it also sheds light on the margin at which ethnic partitioning operates. The likelihood of conflict is approximately $7\% - 8\%$ higher for partitioned, as compared to non-split, groups. The magnitude is similar $(0.08 - 0.09)$ when we restrict estimation to groups close to the national border.[25] The LPM reveals sizable spillovers, as $SPIL$ always enters with a highly significant estimate. The specification in (4) implies that compared to ethnic homelands where none of the nearby groups are split $(SPIL = 0)$, in homelands where half of the adjacent groups are partitioned $(SPIL = 0.5)$ the likelihood of conflict increases by 7%.

Observables vs Unobservables A noteworthy result of both the NB ML and the LPM estimates is the stability of the coefficient on the ethnic partitioning index. The NB estimate on $SPLIT$ in the specification that includes country fixed effects and a rich set of controls is similar to the parsimonious specification (in (1)), where we simply condition on log land area, log population and the presence of water bodies. The heuristic test of Altonji, Elder, and Taber (2005) implies that the bias from unobservable features has to be very large, way larger than the impact of the geographic and location traits and country-specific fixed factors. The R^2 jumps from 0.37 in the parsimonious model without country fixed effects to 0.55 when we add the latter, as country-level characteristics matter crucially for conflict. When we further add the geographic and location controls, the R^2 increases to 0.65. As pointed out by Oster (2015), the sizable increase in the model fit, as we include country constants and relevant controls, coupled with the coefficient stability imply that it is unlikely that unobservable omitted variables spuriously drive our estimates. This is because as the model's fit increases, the portion of the variance that is left to be explained

[24] The estimates in columns (10) and (11) are identical because all outliers (observations where conflict exceeds 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.

[25] We obtain similar results when we replace the country-fixed effects with regional constants and estimate the limited dependent variable model with logit or probit ML. The probit marginal effect with the full set of controls is 0.09 and 0.12 in the full and the border sample, respectively.

by unobservables shrinks.

4.4 Ethnic Partitioning and Conflict Intensity

ACLED reports both deadly events and incidents of violence without casualties (that, nevertheless, involve conflict actors). Which type of conflict is more common across split homelands? To answer this question, we constructed measures of conflict reflecting the number of deadly incidents, the likelihood of deadly conflict, fatalities and conflict duration. By employing these different proxies of conflict severity, we also address concerns that the comprehensive nature of ACLED lumps together events of political violence that differ substantially in the underlying intensity of violence and the casualties involved (Eck (2012)).

Table 3 reports the results. Columns (1) and (6) give NB-ML estimates looking on the number of 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 similar to that of the petroleum dummy (coefficient 0.41). Columns (2) and (7) report LPM estimates, where the dependent variable is a binary index identifying homelands that have experienced at least one deadly incident. There is a 6% to 8% increased likelihood of a deadly event in the homelands of split groups. Again the LPM estimates reveal sizable spillovers. Columns (3) and (8) report NB-ML estimates associating fatalities (aggregated across all types of conflict in all years for each country-ethnic area) to ethnic partitioning. Given the extreme skewness of casualties, the estimate is somewhat unstable;[26] yet $SPLIT$ enters with a significantly positive coefficient both in the full and the border sample. In columns (4) and (9) we focus on conflict duration, i.e., the number of years that there has been some conflict in each homeland, while in columns (5) and (10) we focus on the duration of deadly conflict. Since outliers are not an issue when we examine duration (the mean - variance equality holds), we report country-fixed-effects Poisson ML estimates. There is a strong link between partitioning and conflict duration. The estimate in (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 shows that if a homeland is surrounded exclusively by split groups then conflict duration further increases by 60%, as compared to homelands where none of the adjacent groups is split.

Example Senegal offers an illustration of our results. ACLED records 565 events across its 12 ethnic homelands. In the isolated Casamance region in the South, where the Diolas/Jolas (a major group of half a million people) and the Banyun (a smaller group of approximately $10,000$ people) are partitioned by the colonial border between France and Portugal, we observe 154 and 85 events, respectively.[27] This is

[26] The mean (median) of fatalities is 317 (3) with a standard deviation of $3,307$. This is because of few extreme outliers. The threshold for the top 1% percentile is 435 and the maximum value is $107,554$. See Appendix Table 2.

[27] The 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 at the Portuguese side of the border (Tomas (2006)). Moreover, Casamance was ruled directly from French administrators till 1939,

42.5% of all events (63% if we exclude conflicts in the capital) though these two regions jointly correspond to 11.2% of Senegal's area and only 6% of the country's population. Conflict severity is also high. In these two homelands we observe 61.5% of the country's 182 deadly events and 74% of the country's $1,210$ fatalities. The overwhelming majority of these events involve government troops (129 events) and/or rebels (114 in the homeland of the Diola and 63 in the 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).

In the Supplementary Appendix we discuss extensively two more case studies where partitioning has played a prominent role, namely, conflict in Eastern Congo and in Eastern Africa, where the Somalis are split across five countries.

4.5 Ethnic Partitioning and Type of Conflict

In Table 4 we take advantage of ACLED's detailed conflict classification to distinguish between battles, violence against civilians, and riots and protests. Panel A reports NB-ML estimates and Panel B shows linear probability models with country constants.

Battles Examples of battles include the fights of the Lord's Resistance Army, the Sudanese People's Liberation Army, and Uganda's People Defence Force; and the fighting between the Rwandan forces against Hutu rebels in Rwanda and Eastern Congo. Battles result often (on average 47%) in fatalities; for example, ACLED describes that in a single event in September 1999 the Ugandan army killed 42 Pian warriors from the Karamojong group that is split between Uganda, Sudan, and Kenya. The specifications in (1) and (4) show that (compared to non-split ethnicities) partitioned groups experience $55\% - 60\%$ ($exp(0.45) - 1 = 0.57$) more battles between government forces and militias/rebels. The LPM coefficient on $SPLIT$ is also positive and significant implying that battles are 9% more likely to take place in the historical homelands of partitioned ethnicities. $SPIL$ also enters with a positive (though noisy) estimate, suggesting the weak presence of spatial externalities.

Violence Against Civilians 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 against civilians in Darfur and the assaults of the Central Intelligence Organization in Zimbabwe. Violent events include the burning of churches, hostage-taking and child-soldiering raids by rebels in Nigeria and in Sierra Leone. Going over the event narratives reveals that they are often devastating (43% of these events result in at least in one fatality). For example, in a single

when its administration was transferred to Dakar.

event in Eastern Congo in May 1997 "*ADLF rebels moved in and took control of Mbandaka slaughtering* 200 *Rwandan Hutu refugees*". The NB-ML estimate in the full sample (in (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 larger 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 coefficient 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 LPM reveals sizable spillovers. The coefficient on $SPIL$ is 0.18 and significant at the 1% level; a one standard deviation (0.34) increase in the share of adjacent groups within the country increases the likelihood of one-sided violence 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), 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, it is not surprising that there is no association with ethnic partitioning.

4.6 Sensitivity Checks

We 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 the ACLED increases considerably in 2011, 2012 and 2013, we repeat estimation focusing on the period $1997-2010$. (2) We estimate 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 drop 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 and Poisson models have good small-sample properties. (4) We do 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 factors that vary smoothly by border proximity. (7) We include ethnic-family fixed effects (on top of country fixed effects) to account 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 homelands from each of the five main African regions. (9) We estimate formal spatial models that account for spillovers. (10) We account for conflict spillovers from regions in the same country and the same ethnolinguistic family. (11) We control for the historical legacy of violence from the pre-colonial period. (12) We condition on regional income (overall there is a small and usually insignificant effect of partitioning on proxies of regional

income). Across all these permutations the coefficient on the ethnic partitioning index retains its economic and statistical significance. And most specifications reveal sizable spillovers of ethnic partitioning.

4.7 Heterogeneous Effects

We searched for potential heterogeneous effects of ethnic partitioning. In particular, we explored whether the coefficient on partitioning varies by ethnic features related to: (1) the group's population share 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 ethnicities; (6) whether the bilateral border intersecting split groups 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 belongs to. The analysis (reported in Section 4.1 of the Supplementary Appendix) does not reveal much heterogeneity. We also examined whether the impact of partitioning depends on level of country's ethnic, linguistic, religious diversity, country size and geographic position. Besides some weak evidence that partitioning is particularly harmful for ethnicities in landlocked countries, its effect on conflict is quite homogeneous.

5 Further Evidence. Ethnic Partitioning and Conflict by Key Actors

In this Section we utilize ACLED's grouping of events by conflict actors to shed on the parties involved in violence. We then complement the analysis using georeferenced data on civil wars using an alternative conflict database (UCDP GED).

5.1 ACLED

5.1.1 Data

ACLED categorizes events by main conflict actors, namely: (1) government forces; (2) rebel 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*"; (5) riots and (6) protests, defined "*as violent and non-violent spontaneous groupings (respectively)*"; (7) violence against civilians; and (8) outside/external forces.

We merge rebels and militias (since there is some degree of arbitrariness distinguishing between the two)[28]; and we distinguish foreign interventions from international peace-keeping forces (United Nations or African Union) and from government troops of neighboring countries. If neighboring countries intervene

[28] 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 militias).*"

to assist their co-ethnics across the border, we would expect a significant link between ethnic partitioning and military interventions from adjacent countries. In contrast, there is no reason to expect other types of foreign interventions (from the UN, AU, or NATO) to be related to ethnic partitioning.

5.1.2 Results

Table 5 reports NB-ML (in Panel A) and linear probability model (in Panel B) estimates linking conflict by each actor to ethnic partitioning.[29]

Government Forces. The specifications in (1) and (5) reveal a strong link between partitioning and conflict where government forces are involved. The NB estimates in the full sample imply that there are 70% more conflicts involving state troops whereas the LPM suggests that the likelihood of such conflict is $11\% - 12.5\%$ higher in the homelands of partitioned ethnicities. The LPM specifications indicate sizable externalities of ethnic partitioning; a one standard deviation (0.34) increase in the share of adjacent groups that are split increases the likelihood of state-driven violence by $4.5\% - 6.5\%$.

Rebels and Militias. There is a significant association between ethnic partitioning and conflict where rebels and militias participate (columns (2) and (6)). The LPM suggests that the probability of conflict involving rebel groups is approximately $6.5\% - 8.5\%$ higher in the homelands of partitioned ethnicities. Since ACLED classifies as rebel groups those that explicitly challenge national authority via violent means, these results show that the partitioning - conflict link operates (to some extent) via groups challenging the central government. In line with this interpretation when we separately focus on rebels and militias, we find a stronger effect of partitioning for conflict of rebel groups as compared to militias (results not shown).[30]

Interventions from Neighboring Countries. In columns (3) and (7) we examine whether interventions from neighboring countries are related to ethnic partitioning. This is a key conjecture of the African historiography linking the Scramble for Africa with political violence. While we do report NB-ML specifications (where $SPLIT$ enters with a highly significant coefficient), we focus on the LPM estimates, as the dependent variable is highly skewed. Overall 269 country-ethnic homelands (22.2%) experienced an incursion from a neighboring country. Examples include the interventions of Ugandan and Rwandan troops in DRC, the fighting of Military Forces of Kenya against rebels in Southern Somalia, and the interventions of the military forces of Chad in Mali and the Central African Republic. The estimates 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 and 190 non-split ethnic homelands, interventions from neighboring countries have taken place in 113 regions (19%). 94 of these homelands (83%) are partitioned, while overall 69% of

[29] Since we have already reported specifications with riots and protests and violence against civilians (in Table 4), for brevity we do not repeat them in Table 5.

[30] In the full sample the NB-ML (linear probability model) estimate with rebels only is 0.88 (0.087), while for militias only it is 0.23 (0.056). Moreover, events featuring rebels are quite deadly, especially when fighting against government troops.

ethnic homelands are split. Not only the likelihood but also the frequency of interventions from government forces of nearby countries is higher in the homelands of partitioned ethnicities. In the border sample, we observe 708 interventions from nearby countries in the homelands of split groups, as compared to just 100 interventions in the homelands of non-split groups. Perhaps indicative of the highly targeted nature of military aggressions from neighboring states is the absence of spatial externalities associated with it.

Interventions by International Forces. ACLED also reports conflict 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, and the military interventions of NATO in Libya. We examined whether ethnic partitioning correlates with such type of outside interventions –that we use as a "placebo" as a priori these interventions should not be associated with partitioning. We focus again on the LPM estimates as the variable is highly skewed. The coefficient on $SPLIT$ is small and statistically indistinguishable from zero.

5.2 UCDP GED

5.2.1 Data

To shed further light on the link between ethnic partitioning and conflict we used data from the Uppsala Conflict Data Program Georeferenced Events Dataset (UCDP GED) that covers the period $1989-2010$ (Sundberg, Lindgren, and Padskocimaite (2010), Sundberg and Melander (2013)). The UCDP focuses on deadly incidents associated with civil wars, as identified by the UCDP-PRIO Armed Conflict Database. 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 has played a role include the fights between the Ethiopian government and Somali rebels in the Ogaden region and the fighting between the Tuareg rebels and government troops in Mali and Niger. In total there are $7,512$ state-conflict events.

(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 $5,219$ such events.

(3) Non-state 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.*" Examples include conflict between the various factions of the African National Congress (ANC) and between the ANC and other anti-apartheid movements during the democratic transition in South Africa; and conflict between militias in Kenya's Rift

Valley. In total UCDP includes 3,645 events, though 60% of these incidents are from South Africa during the democratic transition. Amodio and Chiovelli (2014) analyze political violence in South Africa during the democratic transition using UCDP non-state conflict data.

Appendix Figures 4 and $5a - 5c$ portray the distribution of conflict events across ethnic homelands according to UCDP-GED, while Appendix Table 6 gives the number of each type of conflict by country.

5.2.2 Results

Table 6 reports the results linking each type of civil conflict to ethnic partitioning. Appendix Table 9 reports tests of means and medians across ethnic homelands.

State-driven conflict. In Panel A we associate state conflict to ethnic partitioning. The NB estimate in (1) implies that state conflict intensity is 65% higher in the homelands of split ethnicities ($exp(0.50) - 1 = 0.65$). So, this estimate is quite similar to the one obtained with ACLED's data 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 column (4). Columns (2) and (5) give LPM estimates. 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) focusing on the duration (in years) of state-driven conflict. $SPLIT$ enters with a significant coefficient in both samples; the duration of state conflict is approximately 75% higher for partitioned ethnic groups. Across all specifications $SPIL$ enters with a positive coefficient that is significant in the full sample. We also estimated NB-ML models linking fatalities from state conflict to ethnic partitioning (results not shown). While the casualty estimates are highly skewed, there is a strong link between partitioning and the number of fatalities. The coefficients (s.e.) on $SPLIT$ and $SPIL$ in the full sample are 0.78 (0.35) and 1.74 (0.70), respectively, implying large effects.

One-sided violence. Panel B focuses on one-sided violence. 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 (in (2) and (5)). Yet the estimates do not pass the standard significance thresholds. The same applies to the coefficient on the spillover measure ($SPIL$). The results are similar with fatalities (results not shown). $SPLIT$ and $SPIL$ enter with positive though weakly insignificant coefficients (t-stats around 1.4). 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 somewhat 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 war. ACLED reports events that are not part of a full-scale civil war and/or incidents where the perpetrators are not clearly identified. This difference in coverage explains the modest correlation (0.43) between the ACLED-based and UCDP GED-based measures.

Non-State Conflict. In Panel C we look at 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 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.[31]

Example. The UCDP GED mapping of conflict in Casamance in Southern 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 in 322 and 427 fatalities, out of a total of 827 deaths from state-driven conflict in Senegal (90%). Looking at one-sided violence yields a similar picture. There have been 79 violent events against civilians resulting into 243 fatalities in Casamance whereas one-sided violence in Senegal outside these two groups is minimal (15 events). The UCDP documentation states that all these events involved either state conflict against the MFDC or one-sided civilian violence conducted by MFDC rebels.

5.2.3 Summary

Both the ACLED and UCDP analysis shows that ethnic partitioning matters crucially for two-sided political violence where government troops fight against rebels and militias. The analysis further shows that there is a weaker link between ethnic partitioning and one-sided violence; and no association with conflict where only non-state actors are involved.[32] This ordering is in line with the theoretical work of Besley and Persson (2011a), who argue that when the opportunity cost of conflict is low and political institutions are non-cohesive, then political violence takes more often the form of two-sided violence (civil war) rather than one-sided violence (repression). Below we examine this issue in greater detail.

6 Partitioning and Ethnic Power Relations

6.1 Motivation and Data

Political violence is multifaceted; so far we have focused on the instances that materialize into conflict. However, a growing empirical literature in African political economy provides compelling evidence that national politics are characterized by ethnic favoritism, patronage, and discrimination (Posner (2005)); Asiwaju (1985) discusses case studies pointing out that split groups not only participate in state-driven conflict, but also become targets of abusive policies. We thus examine the impact of ethnic partitioning on various forms of political violence using data from the Ethnic Power Relations (EPR) dataset (Wimmer, Cederman, and Min (2009)).

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

[32] Arguably conflict incidents are not perfectly measured by either ACLED or UCDP. To account for error-in-variables in the Supplementary Appendix we combine the two datasets to obtain a more precise picture on the presence of conflict. The link between ethnic partitioning and civil conflict is quite strong.

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 around the world. The coverage for Africa spans 40 countries and 196 groups. Using a multitude of sources, we linked the 196 EPR groups to 593 ethnicities on Murdock's (1959) map.[33] Among the 593 ethnicities 234 (39.5%) are partitioned and 359 (60.5%) are non-split. EPR provides ethnic-specific information on political representation (or exclusion) in the national government. Participation in the governing coalition may take the following forms (in ascending order) from junior to senior to dominant partner, to being the monopolist in the national politics. Among groups excluded from the central government there are three mutually exclusive categories: those enjoying some regional autonomy; those that are powerless but are not discriminated against; and ethnicities that face active discrimination from the central government. EPR also identifies civil wars with an explicit ethnic angle.

6.2 Ethnic Wars

We start by exploring the link between partitioning and civil wars that have an ethnic dimension. Using ethnic wars as the outcome of interest has several advantages. First, we look at major breakouts of violence. The coding of civil wars is based on the widely-used UCDP/PRIO Armed Conflicts Data Set (Petter, Wallensteen, Eriksson, Sollenberg, and Strand (2002)). 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.*" Second, 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 georeferenced conflict data. Third, EPR has a long time horizon covering the entire post-independence period. Fourth, by looking at politically relevant groups, we check the robustness of our findings to focusing on ethnicities with a presence in the national political sphere.

A simple tabulation reveals the stark disparities with respect to civil war participation between split and non-split groups. On the one hand, 72 out of the 234 partitioned ethnicities (31%) have taken part in an ethnic war. On the other hand, only 19% (69 out of the 359 non-split groups) have participated

[33] Such 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 the "Bushmen and their kin" cluster in Murdock (1959). In other instances, the matching is less straightforward. For example, in Nigeria EPR blends the "Hausa-Fulani and Muslim Middle Belt" in a single category. In this case we used the *A-MAR* correspondence (Wilkenfeld, Brancati, Fearon, Gurr, Laitin, Pate, and Saideman (2014)). We also used the georeferenced version of EPR so as to identify the corresponding location of groups on Murdock's map. This method is the least satisfactory and, hence, was only used for roughly 10% of cases. Results are unaffected if such matches are excluded.

in an ethnic war. Examples of split groups that have been involved in ethnic wars include the Afar in North-East Ethiopia, that have also faced large-scale discrimination and marginalization policies by the central government for many years. For example, in 1975 the Dergue administration nationalized all land and annulled the de facto autonomy of the Afar leading to a secessionist rebellion (Vaughan (2003)).

In columns (1)-(4) of Table 7 we formally assess the impact of ethnic partitioning on ethnic-war incidence. 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 partitioned groups have roughly an 11% increased likelihood of participating in an ethnic war, as compared to non-split groups. This magnitude is similar to the LPM 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 by roughly 7%, contributing significantly to the eruption of ethnically-tainted large-scale violence.

6.3 Ethnic Discrimination

We now turn our focus on repression using the EPR's information on ethnic discrimination, defined as: "*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, 110 (out of 593) groups have being discriminated by the national government at some point (18.5%).[34] This average masks considerable differences between partitioned and non-split ethnicities. 58 of the 234 split 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 52 of the 359 non-split groups faced discrimination. Examples of discriminated partitioned groups include the Bushmen (San/Basarwa) in Botswana that have faced restrictions on residence, limited access to the civil service and no recognition of their traditional chiefs. Another example is that of 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 (5)-(8) report LPM estimates that explore the association between ethnic dis-

[34] On average these 110 ethnicities have faced discrimination post-independence for 21 years.

crimination and partitioning. Column (5) reports the unconditional specification, while in column (6) we include country-specific constants. The coefficient of $SPLIT$ is 0.078 and highly significant. 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 likelihood of discrimination is approximately 7.2% 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 externalities on political discrimination from partitioning.

6.4 Political Violence (Repression and Civil Wars)

So far, we have analyzed the relationship between partitioning and two forms of political violence, ethnic wars and ethnic discrimination, without attempting to disentangle which form of political violence comes first. The case of the Diolas in Senegal discussed above suggests that discrimination may precede conflict, but other cases such as that of the Somalis in Northern Kenya covered in Section 8 of the Online Appendix indicate that the reverse is not uncommon. Moreover, Besley and Persson (2011a) criticize theoretical and empirical works on civil war and repression for having moved in parallel and stress the importance of studying these aspects of political violence jointly, since they have common roots and are naturally interrelated. Besley and Persson (2011a) develop a model where groups compete for the control of central government whose resources can either be used for group-specific consumption or for public goods investments. Their model that allows both for one-sided (repression) and two-sided violence (civil war) links group-level opportunity cost of fighting and nationwide resource rents to an ordered measure of political violence when institutions are non-cohesive. Following their approach we examined the impact of ethnic partitioning on a trichotomous index of political violence that equals two if the group has ever engaged in an ethnic war, one when the group has been subject to ethnic discrimination from the central government but not in a civil war, and zero when the group has neither been discriminated against nor involved in an ethnic war.

The descriptive statistics are telling. Among partitioned groups 72 out of the 234 (30.8%) have engaged in an ethnic war, while 22 (9.40%) have suffered from ethnic discrimination only. The corresponding likelihoods of civil war and political repression for non-split groups are 19.2% (69 groups out of 359) and 6.7% (24 ethnicities), respectively. Table 7, columns (9)-(12) report ordered logit ML estimates associating the trichotomous index of political violence to ethnic partitioning. To avoid the incidental parameters problem, in columns (9) and (10) we replace the country dummies with regional constants; in columns (11) and (12) we follow Besley and Persson (2011a) and implement the fixed-effects logit estimator of Ferrer-i-Carbonell and Fritjetrs (2004). In all specifications the ethnic partitioning index enters with a highly significant coefficient. As logit coefficients are not easily interpretable, we compared the model predictions for the likelihood of each outcome (0 = peace, 1 = repression, 2 = ethnic war) for split and non-partitioned ethnicities. The estimates in (9)-(10) imply that the likelihood of civil war increases from 19.5% to 31.5%

for partitioned ethnicities, while the likelihood that an ethnicity will experience political discrimination is 7% for non-split groups and 9% for partitioned ones. These estimates, that take into account various ethnic characteristics and regional heterogeneity, are close to the simple tabulations, showing that unobservables are unlikely to matter much.

So, ethnic partitioning is more strongly associated with ethnic wars compared to repression. This result is in line with Besley and Persson (2011a) theoretical argument that in weakly institutionalized countries, a low opportunity cost of fighting translates more often into two-sided violence. In our context this is driven by two empirical regularities. First, political violence in Africa overall takes more often the form of two-sided violence, rather than repression (24% versus 18.5%); second, when a group faces discrimination from the national government, it is also more likely to engage in an ethnic war (out of the 110 groups that have ever been discriminated against, 64 have also participated in an ethnic war), i.e., one-sided violence often escalates into full-fledged conflict.

In the Supplementary Appendix we provide additional sensitivity checks, showing that the ethnic partitioning- discrimination-ethnic war link is present when: (1) we classify a group as split if at least 5% of its homeland falls into more than one country; (2) we drop iteratively each African region; and (3) we estimate alternative linear and non-linear models with the ordered index of political violence.

6.5 Further Evidence

We should stress here that the disproportionate incidence of ethnic wars and political discrimination among split groups does not imply that the latter are never part of the central government. 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 some capacity (either as junior or senior partners) in the various government coalitions over time. This result echoes Francois, Rainer, and Trebbi (2015) finding that groups' participation in the government is widespread. Empirically, this oscillation of split groups between a state of conflict and discrimination at some point and members of governing alliances at other instances translates into an insignificant relationship between partitioning and the probability of a group having ever been a partner in the central government.

Moreover, this nuanced political status of split groups suggests that the vicious cycle of discrimination and conflict in which they are often embroiled, comes into play primarily when excluded from the governing coalition. We checked the empirical validity of this conjecture exploiting the time-series information of group's political status from the EPR. Specifically, we run year-country-ethnicity regressions with country and year fixed effects associating the onset of ethnic wars in period t with ethnic partitioning separately for groups that have been excluded from political power in any of the past three (or five) years and those that have been included in the national power-sharing coalition during the same time. The estimates reported in the Supplementary Appendix imply that the likelihood of ethnic war onset is 0.6% higher for

partitioned as compared to non-split groups, when ethnicities are excluded from national power. This effect is considerable, as the yearly likelihood of war onset across the 593 groups during post independence is 0.7%. In contrast, the difference in the likelihood of civil war between split and non-split ethnicities when they take part in the central government is much smaller (0.15%).[35] These findings show that the link between exclusion from political power and ethnic-based civil wars, documented by Cederman, Wimmer, and Min (2010), disproportionately affects split groups. The rationale is straightforward. Partitioned ethnicities have a lower cost of engaging in armed rebellions against the government, as they can get assistance from their co-ethnics on the other side of the border; hence split groups are more likely to react violently against their marginalization and exclusion from the central government.

7 Ethnic Partitioning and Individual Well-Being

Our evidence suggests that partitioned ethnic groups are more likely to engage in conflict (predominantly against the government), to experience violence against their civilian population and to suffer from repression. To further understand how the ethnic partitioning - political violence link operates, in this Section we employ micro-level data from the Demographic and Health Surveys (DHS) to examine how individuals of partitioned groups fare economically compared to citizens from non-split groups in the same country. Exploiting individual-level variation has some straightforward advantages. 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 impact of ethnic partitioning. Third, since we observe people residing within and outside their group's historical homeland, we can evaluate whether ethnic partitioning has negative repercussions irrespective of respondents' residence or whether the negative effects are concentrated among individuals residing in partitioned homelands.

7.1 Data and Specification

The Demographic and Health Surveys are based on nationally representative samples and include information on households' wealth, education, occupation, and health. We use all georeferenced surveys with information on the respondents' ethnic identity. Our sample comprises 20 countries and covers 88,171 male respondents. We focus on two outcome variables. First, we use household's composite wealth index that ranges from 1 to 5 and corresponds to the quintiles of the distribution of household wealth in each country. This index reflects access to basic public goods (electricity, sewage system, piped water) and ownership of

[35] We also pooled all ethnicities across all years and interacted the ethnic partitioning index with an indicator that equals one in period t when a group has been excluded for at least one year from the central government in any of the previous three (or five) years. The interaction term between exclusion from the governing coalition and ethnic partitioning enters with a positive and highly significant estimate (0.5%), implying that the likelihood of civil war onset increases considerably for partitioned ethnicities when they are politically excluded. A similar pattern obtains in the cross-section. See in the Online Appendix the discussion in Section 5.2 and the corresponding Tables 34*A* and 34*B*.

various assets. Second, we use years of schooling.

Our empirical specification reads:

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

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 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_c, that among others, capture 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 an 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. To account for spatial correlation and the fact that the split indicator takes on the same value for individuals belonging to the same 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 estimates linking the composite 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. Individuals who identify with split ethnicities have on average lower access to public goods and worse living conditions. In column (2) we add a vector of location controls, namely distance terms to the national border, the capital, and the coast. To adequately capture the capital city effect we also include an indicator for enumeration areas, whose distance to the capital is less than the 25th percentile.[36] The coefficient drops, but retains significance at standard confidence levels. In column (3) we account for individual characteristics. The estimate implies that respondents identifying with split ethnicities have roughly 0.20 points of lower wealth as compared to individuals from non-split ethnicities in the same country; this translates into a standardized beta coefficient of 0.07, which is half of the beta coefficient on the capital city indicator. So the impact of identifying with a split group is half of residing in the capital city. In columns (4)-(6) we focus on enumeration areas close to the border, using as a cutoff the median distance (80 kilometers). In all specifications the coefficient on the ethnic partitioning index is negative and significant at 1% level and similar to the analogous estimates in the full sample. In columns (7)-(12) years of schooling serves as the dependent variable. The estimates imply that -conditional

[36] The 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. The dummy on the capital city indicator is also positive and significant.

on location and various individual characteristics- individuals from partitioned ethnicities have on average 0.4 more years of formal education. The beta coefficient on the ethnic partitioning index is around 0.052, implying an economic effect moderately smaller to that of residing in capitals (the beta coefficient on the capital city dummy is 0.076).

7.3 Ethnic Identity and Ethnic Homelands

So far, we have two main findings. First, political violence is more frequent in the homelands of split ethnicities. Second, the scars of ethnic 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 of 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 and the EPR-based analysis suggests that it is the latter that is going on. To shed light on this, we augment the baseline specification with a location-based indicator 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 (identity-based and location-based) coincide.[37]

Table 9 reports the results. In column (1) the identity-based partitioning indicator enters negatively and significantly, whereas the location-based index does not. The insignificance of the latter is driven by the location controls; when we do not account for those both partitioning indicators enter with significantly negative estimates. In column (2) we control for individual characteristics; the pattern remains unchanged. In columns (3)-(4) we restrict estimation to areas close to the border. The coefficient on the identity-based partitioning index remains negative (-0.26); this suggests that even when we focus on the border and control for numerous individual and location features, members of partitioned ethnicities have worse living conditions as compared to those identifying with non-split groups. In columns (5) and (6) we introduce an interaction term between the partitioned ethnic identity indicator and the partitioned ethnic location indicator; this variable identifies individuals that reside in partitioned homelands and identify with split groups. The interaction enters with an insignificant coefficient suggesting that the negative impact of partitioning is not different for members of partitioned groups residing in split homelands. The pattern is similar for schooling outcomes in columns (7)-(12). The identity-based measure of partitioning enters all permutations with a significantly negative estimate, implying that individuals from partitioned ethnicities have on average half a year less of schooling.

[37] Out of $88,171$ individuals, $25,631$ (29%) self-identify with a split group and reside in split homelands; $36,694$ (41.5%) neither identify with a split group nor they reside in partitioned homelands; $13,256$ (15%) reside in split homelands, but identify with a non-split group; and $12,590$ (14.5%) reside in non-split homelands but identify with split groups.

7.4 Sensitivity Checks and Further Evidence

In the Supplementary Appendix we further investigate the impact of ethnic partitioning on public goods. First, we exploit within enumeration-area variation, so as to fully account for differences in location-specific, time-invariant characteristics. While these specifications are quite restrictive (as the sample includes $7,898$ villages/towns/cities), they reveal that members of split groups have systematically worse access to utilities and are less educated than respondents from non-split groups residing in the same place.

Second, to further account for unobserved differences between "movers" and "non-movers", we estimated the link between ethnic partitioning and economic performance separately for "non-movers" (respondents currently residing within their group's ancestral homeland) and "movers" (individuals living outside their ethnicity's historical homeland). And we further distinguished between "movers" residing in non-partitioned ethnic homelands and "movers" in split homelands, other than their own ethnicity. The pattern is fairly uniform. Respondents identifying with a split ethnicity register lower levels of wealth and have fewer years of schooling irrespective of their location.

Third, we examined the persistence of ethnic partitioning's impact distinguishing between "old" and "young" respondents. The negative impact of partitioning on individual outcomes does not differ across generations pointing to its ongoing importance.

8 Conclusion

We study the consequences of a neglected aspect of the colonization in Africa, the drawing of political boundaries among European powers, which led to the partitioning of several ethnicities across African states upon independence.

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 measures and ethnic-specific proxies of pre-colonial conflict and early development. With the exception of the size of the historical homeland, there are no significant differences between partitioned and non-partitioned groups. Hence, our results offer support to the claim of the African historiography on the largely accidental drawing of the 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 to be its major consequence. We exploit a new dataset that reports precisely geocoded information for $64,650$ conflict events of various types over the period $1997-2013$ for all African countries. Exploiting within-country across-ethnic-homeland variation we uncover that political violence is prevalent in partitioned homelands which experience deadly incidents over prolonged periods of violence.

Third, we take advantage of the richness of the data to shed some light on the mechanisms at work. We present evidence suggesting that neighboring countries use the homelands of partitioned groups to stage

military interventions. This suggests that ethnic partitioning is associated with a lower opportunity cost of fighting, as neighboring countries often offer military, political and economic support to their co-ethnics on the other side of the border. We also find that ethnic partitioning is mostly associated with state-driven conflict, where government troops (or state-backed militias) clash with rebels opposing the national government. Ethnic partitioning is also linked to increased violence against the civilian population. In contrast, there is no association between ethnic partitioning and conflict involving non-state actors, rioters and protesters.

Fourth, using data from the Ethnic Power Relations database that reports information on ethnic-based discrimination from the national government and civil wars with an explicit ethnic angle for politically relevant groups in 40 countries post independence, we examine in detail the ethnic partitioning - political violence nexus. Partitioned ethnicities are significantly more likely to experience political discrimination and are more likely to participate in ethnic-based civil wars. The impact of ethnic partitioning on conflict is relatively stronger, as compared to its impact on one-sided political violence/discrimination and is particularly salient in periods when split groups are excluded from the central government.

Fifth, using micro-data from the Demographic and Health Surveys, covering more than $85,000$ respondents in 20 African countries, we document that individuals identifying with partitioned groups have fewer household assets, poorer access to public utilities, and lower education. 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 circumstances of members of split ethnicities irrespective of their actual residence.

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.[38] And, since border artificiality and ethnic partitioning are not an exclusive African phenomenon, subsequent works could study their effect in other world regions, such as the Middle East and the Caucasus.

[38] For example, ethnic partitioning may offer some economic benefit insomuch 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, 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., 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.

Allison, P. D., and R. P. Waterman (2002): "Fixed-Effects Negative Binomial Regression Models," *Sociological Methodology*, 32(1), 247–265.

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.

Amodio, F., and G. Chiovelli (2014): "Ethnicity, Migration and Conflict: Evidence from Contemporary South Africa," mimeo, University of Bologna, Economics Department.

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 (2011a): "The Logic of Political Violence," *Quarterly Journal of Economics*, 126(3), 1411–1445.

Besley, T., and T. Persson (2011b): *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.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Padro-i-Miguel (2015): "The Value of Democracy: Evidence from Road Building in Kenya," *American Economic Review*, 105(6), 1817–1851.

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.

Cederman, L.-E., A. Wimmer, and B. Min (2010): "Why do ethnic groups rebel? New data and analysis.," *World Politics*, 62(1), 87–119.

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.

Collier, P. (2007): *The Bottom Billion*. Oxford Universiy Press, Oxfod, United Kingdom.

Collier, P., and A. Hoeffler (2007): "Civil War," in *Handbook of Defense Economics, Volume 2, Defense in a Globalized World,*, ed. by T. Sandler, and K. Hartley, pp. 711–740. Elsevier, Amsterdam, Netherlands.

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.

Depetris-Chauvin, E. (2014): "State History and Contemporary Conflict: Evidence from Sub-Saharan Africa," mimeo, Brown 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.

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

——— (2015): "The European Origins of Economic Development," revised draft of 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. (2009): *Africa, Unity, Sovereignty and Sorrow*. Lynne Rienner Publishers, Inc., Boulder, Colorado.

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

Esteban, J., L. Mayoral, and D. Ray (2012): "Ethnicity and Conflict: An Empirical Study," *American Economic Review*, 102(4), 1310–1342.

Esteban, J., and D. Ray (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.

Evans, M. (2004): "Senegal: Mouvement des Forces Democratiques de la Casamance (MFDC)," Chatman House Briefing Paper, Africa Programme AFP BP 04/02.

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.

Fenske, J. (2013): "Does Land Abundance Explain African Institutions?," *Economic Journal*, 123(573), 1363–1390.

——— (2014): "Ecology, Trade, and States in Pre-colonial Africa," *Journal of the European Economic Association*, 12(3), 612–640.

Ferrer-i-Carbonell, A., and P. Fritjetrs (2004): "How Important is Methodology for the Estimates of the Determinants of Happiness?," *Economic Journal*, 114(2), 641–659.

Fisman, R., and E. Miguel (2007): "Corruption, Norms, and Legal Enforcement: Evidence from Diplomatic Parking Tickets," *Journal of Political Economy*, 115(6), 1020–1048.

Franck, R., and I. Rainer (2012): "Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education and Health in Sub-Saharan Africa," *American Political Science Review*, 106(2), 294–325.

Francois, P., I. Rainer, and F. Trebbi (2015): "How Is Power Shared In Africa?," *Econometrica*, 83(2), 465–503.

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.

Hodler, R., and P. A. Raschky (2014): "Regional Favoratism," *Quarterly Journal of Economics*, 129(2), 995–1033.

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.

Luca, G. D., R. Hodler, P. A. Raschky, and M. Valsecchi (2015): "Ethnic Favoritism: An Axiom of Politics?," CESifo Working Paper Series No. 5209.

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., 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.

——— (2014): "Historical Development," in *Handbook of Economic Growth, Volume 2*, ed. by P. Aghion, and S. Durlauf, vol. 2, chap. 7, pp. 347–402. 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.

Oster, E. (2015): "Unobservable Selection and Coefficient Stability: Theory and Evidence," mimeo, Brown University, Department of Economics.

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.

Wagane, F. (2006): "The Casamance Separatism : From Independence Claim to Resource Logic," Thesis. Naval Postgraduate School.

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.

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***	0.0534***	0.0528***	0.0640***
	(0.0153)	(0.0159)	(0.0175)	(0.0186)	(0.0150)	(0.0156)	(0.0172)
Lake Indicator		0.0961*	0.0963	0.0965*	0.0910	0.0933*	0.0900
		(0.0565)	(0.0645)	(0.0567)	(0.0554)	(0.0557)	(0.0624)
River Indicator		-0.0053	-0.0164	-0.0092	-0.0064	-0.0057	-0.0174
		(0.0337)	(0.0324)	(0.0325)	(0.0344)	(0.0346)	(0.0347)
Elevation			-0.0411				-0.0673
			(0.0709)				(0.0726)
Suitability for Agriculture			0.1239				0.1591
			(0.0974)				(0.1078)
Malaria Stability Index				0.0195			-0.0415
				(0.0982)			(0.1097)
Distance to the Coast				0.0000			-0.0001
				(0.0001)			(0.0001)
Diamond Mine Indicator					0.0289		0.022
					(0.0647)		(0.0647)
Oil Indicator					-0.0774		-0.1066*
					(0.0545)		(0.0625)
Nearby Groups in the Same Family						-0.0727	-0.0662
						(0.0579)	(0.0622)
Adjusted R-square	0.050	0.055	0.059	0.056	0.058	0.057	0.065
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***
	(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*
	(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
	(0.0337)	(0.0351)	(0.0348)	(0.0347)	(0.0337)	(0.0337)	(0.0331)
Pre-colonial Conflict Indicato	-0.0663						
	(0.0733)						
Distance to Pre-colonial Conflict		-0.0444					
		(0.0839)					
Slave Trades Indicator			0.0045				
			(0.0322)				
Log Number of Slaves (normalized by land area)				0.0063			
				(0.0080)			
Pre-colonial Kingdom Indicator					0.0466		
					(0.0469)		
Distance to Pre-colonial Kingdom						0.0009	
						(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***
	(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	0.3910	0.4619*	0.4920*	0.4834*	0.4256*	0.0879	0.5679	0.4328	0.4514	0.4514	0.3928
	(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
R-square	0.203	0.528	0.645	0.633	0.168	0.182	0.148	0.343	0.546	0.574	0.574	0.544
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*
	(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***
	(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	0.430	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 (NB) 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 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 in the same country. 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 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 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		
	Battles	Civilian Violence	Riots & Protests	Battles	Civilian Violence	Riots & Protests
	(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 (NB) 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		
	All Events	Indicator	Duration	All Events	Indicator	Duration
	NB-ML	LPM	NB-ML	NB-ML	LPM	NB-ML
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: State (Government Forces) Civil Conflict						
SPLIT (Partitioning)	0.4978**	0.0487*	0.3390**	0.8053***	0.0799**	0.5469**
	(0.2411)	(0.0294)	(0.1422)	(0.2335)	(0.0393)	(0.2389)
SPIL (Adjacent Split)	1.1577***	0.0902*	0.6868**	0.4340	0.0424	0.255
	(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
	(0.2416)	(0.0292)	(0.1474)	(0.2615)	(0.0404)	(0.2237)
SPIL (Adjacent Split)	0.4708	0.0829*	0.4935**	0.0901	0.0161	0.1659
	(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
	(0.4062)	(0.0317)	(0.3374)	(0.5178)	(0.0283)	(0.4917)
SPIL (Adjacent Split)	-0.8703	-0.0344	-0.5089	-0.6593	-0.0268	-0.6964
	(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 (NB) maximum likelihood (ML) 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 and Political Violence (Discrimination and Ethnic Wars)

	Ethnic War				Ethnic Discrimination				Ordered Political Violence			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SPLIT (Partitioning)	0.1155**	0.1402***	0.1068***	0.1101***	0.1030***	0.0778***	0.0715***	0.0718**	0.6175**	0.5802**	0.7805***	0.8856***
	(0.0509)	(0.0381)	(0.0388)	(0.0390)	(0.0342)	(0.0276)	(0.0276)	(0.0282)	(0.2543)	(0.2379)	(0.3017)	(0.3138)
SPIL (Adjacent Split)				0.2570**				0.0169		1.1716		2.0160**
				(0.1155)				(0.0764)		(0.6807)		(0.9168)
Adjusted R-square	0.019	0.423	0.474	0.487	0.018	0.475	0.521	0.521	—	—	—	
Log Likelihood	—	—	—	—	—	—	—	—	-422.612	-417.453	-170.503	-166.053
Observations	593	593	593	593	593	593	593	593	593	593	405	405
Country Fixed Effect	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	Yes	Yes
Simple Controls	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes

The table reports linear probability model estimates (in columns (1)-(8)) and ordered logit estimates (in columns (9)-(12)), associating ethnic-based political violence 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 a major or minor civil war with an explicit ethnic dimension over the period 1960-2010. The dependent variable in columns (5)-(8) 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 (9)-(12) is an ordered index of political violence. The trichotomous index of political violence equals two if the ethnic group is engaged in a major civil war (two-sided conflict); the index equals one when the group is subject to political discrimination from the national government but not in civil war (one-sided violence); the index equals zero when the ethnicity is neither discriminated from the national government nor involved in civil war (the construction of the ordered index of political violence follows Besley and Persson (2011). 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 in the same country. The specifications in columns (2)-(4) and (6)-(8) include a vector of country fixed effects (constants not reported). The specifications in columns (3)-(4) and (7)-(12) include log of land area, the log of population in 1960, an indicator for lakes and 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 and an indicator for country-ethnic areas that are by the sea coast (location controls); and 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 (geographic controls). Columns (9) and (10) report ordered logit estimates, replacing the country fixed-effects with regional constants. Columns (11)-(12) report fixed-effects ordered logit estimates employing the approach of Ferrer and Fritjers (2004). For specifications (1)-(10) the table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. For specifications (11) and (12) the table gives standard errors clustered at the country level. ***, **, 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)
Partitioning - Identity	-0.3853***	-0.2369**	-0.2069**	-0.2818***	-0.2557***	-0.2454***	-0.9907**	-0.6594**	-0.4897*	-0.5301*	-0.4701*	-0.4089*
Double-clustered s.e.	(0.1129)	(0.0972)	(0.0905)	(0.1018)	(0.0810)	(0.0789)	(0.4045)	(0.3045)	(0.2520)	(0.3086)	(0.2658)	(0.2444)
Non-Indigenous	0.1936**	0.1873**	0.1890***	0.1359*	0.1792**	0.1828**	0.3116	0.2922	0.3163	0.1273	0.2028	0.2056
Double-clustered s.e.	(0.0887)	(0.0744)	(0.0721)	(0.0783)	(0.0767)	(0.0750)	(0.2695)	(0.2485)	(0.2328)	(0.2043)	(0.1829)	(0.1693)
Adjusted R-squared	0.028	0.145	0.171	0.053	0.134	0.162	0.176	0.227	0.281	0.164	0.198	0.244
Observations	88171	88171	88171	44090	44090	44090	88043	88043	88043	44030	44030	44030
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Individual 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 years of education (in columns (7)-(12)) with ethnic partitioning. The ethnic partitioning index (Partitioning-Identity) takes on the value of one for individuals that identify with a partitioned ethnicity and zero otherwise. 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 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 location controls in columns (2), (3), (5), (6), (8), (9), (11), and (12) includes the distance of each enumeration area to the capital city, the distance to the coast, the distance to the national border and an indicator that takes on the value of one for enumeration areas close to the capital city (distance to the capital less than the 25th percentile). The set of individual controls in columns (3), (6), (9), and (12) includes a vector of year-of-birth fixed effects, a vector of 6 marital-status fixed effects, and a vector of 7 religion fixed effects. 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 to the border; 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.

Table 9: The Long-Run Effects of Ethnic Partitioning on Individual Well-Being and Education. DHS Data Channels; Location and Identity

	Composite Wealth Index						Education					
	All Observations		Border Observations		All Observations		All Observations		Border Observations		All Observations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Partitioning - Identity	-0.2469**	-0.2150**	-0.2733***	-0.2612***	-0.2818**	-0.2519**	-0.7277**	-0.5513**	-0.5504**	-0.4812*	-0.9405**	-0.7710**
Double-clustered s.e.	(0.1012)	(0.0933)	(0.0832)	(0.0809)	(0.1388)	(0.1255)	(0.3190)	(0.2646)	(0.2724)	(0.2510)	(0.4496)	(0.3634)
Partitioning - Location	0.0413	0.033	0.0873	0.0777	0.0072	-0.0031	0.2801	0.252	0.3993*	0.3574*	0.072	0.037
Double-clustered s.e.	(0.0751)	(0.0760)	(0.0838)	(0.0826)	(0.0980)	(0.1008)	(0.1988)	(0.1858)	(0.2185)	(0.1958)	(0.2458)	(0.2379)
Non-Indigenous	0.1885**	0.1900***	0.1845**	0.1874**	0.2058**	0.2083**	0.3006	0.3237	0.227	0.2268	0.4062	0.4326
Double-clustered s.e.	(0.0740)	(0.0717)	(0.0760)	(0.0743)	(0.0888)	(0.0845)	(0.2454)	(0.2300)	(0.1773)	(0.1638)	(0.2970)	(0.2701)
Partitioned Location & Identity					0.0708	0.0750					0.4323	0.4464
Double-clustered s.e.					(0.1400)	(0.1350)					(0.4113)	(0.3625)
Adjusted R-squared	0.145	0.171	0.135	0.162	0.145	0.171	0.227	0.282	0.199	0.245	0.227	0.282
Observations	88171	88171	44090	44090	88171	88171	88043	88043	44030	44030	88043	88043
Country Fixed Effects	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
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

The table reports OLS estimates associating a composite wealth index (in columns (1)-(6)) and years of education (in columns (7)-(12)) with ethnic partitioning. The ethnic partitioning identity index (Partitioning-Identity) takes on the value of one for individuals that identify with a partitioned group. The location-based ethnic partitioning index (Partitioning-Location) takes on the value of one for individuals that currently reside in ethnic homelands that have been partitioned by the national border and zero otherwise. 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). All specifications include a vector of country fixed effects (constants not reported). The set of location characteristics controlled for in all columns include the distance of each enumeration area 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 for enumeration areas close to the capital city (distance to the capital less than the 25th percentile). The set of individual controls in even-numbered columns includes a vector of year-of-birth fixed effects, a vector of 6 marital-status fixed effects, and a vector of 7 religion fixed effects. 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 to the border; 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 ethnic identity 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, CEPR and NBER

Elias Papaioannou
London Business School, CEPR and NBER

October 23, 2015

Abstract

The Supplementary Appendix is structured into eight sections. Section 1 gives descriptive evidence, summary statistics and presents the key patterns of the spatial distribution of all conflict types across Africa. Section 2 reports additional to the main text results showing that ethnic partitioning is not systematically related to various ethnic-specific precolonial 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 geographic, ecological, natural resources and location characteristics. Section 3 reports numerous sensitivity checks on the impact of ethnic partitioning on various aspects of civil conflict (intensity, likelihood, duration, and fatalities). Section 4 explores potential heterogeneous effects of ethnic partitioning on conflict with respect to both ethnic and country characteristics. Section 5 reports robustness checks on the link between partitioning, ethnic-based discrimination from the central government and participation in ethnic wars using data from the Ethnic Power Relations Database. Section 6 reports further evidence on the negative impact of identifying with a partitioned ethnicity using individual-level data from the Demographic and Health Surveys. Section 7 presents a case-study illustrating the effect of ethnic partitioning on conflict in Central-Eastern Africa, the most conflict-prone region in Africa. Section 8 reports an out-of-sample counterfactual analysis that approximates the impact of partitioning the Somalis in Eastern Africa.

*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 partitioned ethnicities (as coded by Murdock (1959) and identified by our method) 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 of the main text). Panel A looks at all homelands (825 observations) whereas in Panel B we focus on the 413 groups whose centroid is close to the national borders (border distance less than the median, 102 km).

Appendix Table 2 reports summary statistics for all variables at the country-ethnic homeland level, which is the unit of analysis in Sections $4-5$; where we 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 for the border sample, using as a cutoff the median 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). We focus on precisely geocoded incidents of political violence. There are 8 conflict categories. (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 or 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 engage in non-violent actions that are, however, within the context of an ongoing civil conflict and dispute (e.g., recruitment drives, incursions or rallies); (6) Riots and protests; (7) Violence again civilians, where armed groups attack civilians; (8) Non-Violent transfer of control. For our analysis we merge all battles (categories (1)-(3)); in many specifications we drop non-violent events by main conflict actors and in some other specifications we also exclude riots and protests.

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 yearly events resulted in fatalities (30%). Conflict events increase

considerably in 2011 (5,261) and especially in 2012 (8,753) and in 2013 (12,565). Below we report estimates dropping 2011, 2012, and 2013 to ensure that our estimates are not driven by events during these 3 years.

Appendix Table 4 reports the number (and share) of all and deadly events for each of the 8 ACLED categories. The sample includes 23,381 battles, mostly involving government troops, 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 casualty). 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.

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 only. 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).[1]

Appendix Table 6 reports conflict events 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 regions in Africa; in both countries we observe pervasive violence against civilians and incessant 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; Figures $2a - f$ portray events associated with each conflict actor separately.

[1] So, violence against civilians and riots/protests can be calculated in two different ways, one simply by looking at the main conflict categories (that also includes battles and non-violent events associated with a conflict actor) and also by looking at the actors involved (that include rebels, government troops, etc.). While there are some minor discrepancies between these two classifications, this does not affect the results (correlation around 0.99). Below we report estimates using both the main-category classification (e.g., Appendix Tables 13 and 16) and the "actor-based" classification (e.g., Appendix Tables 14 and 17).

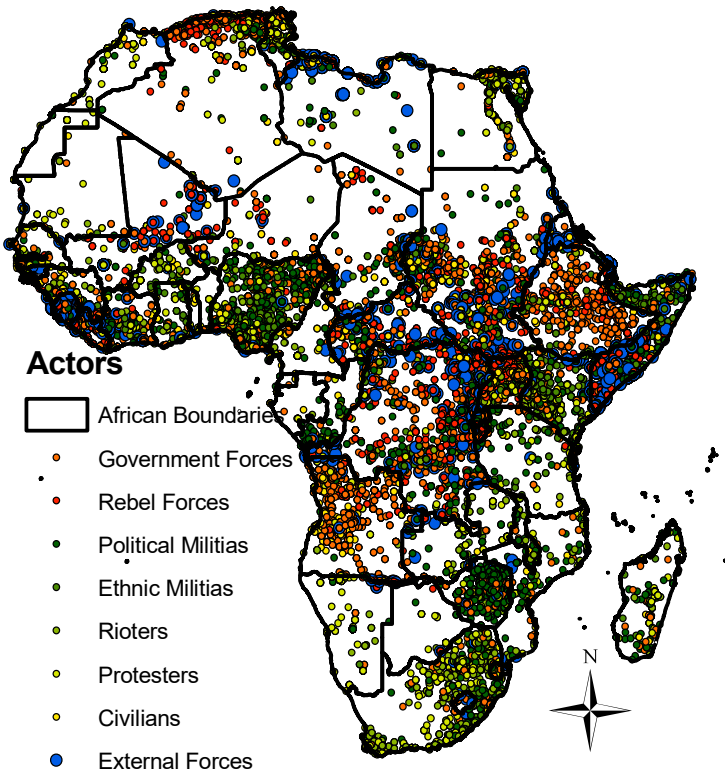

Figure **1**

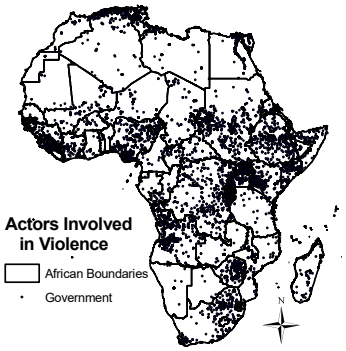

Figure **2a**

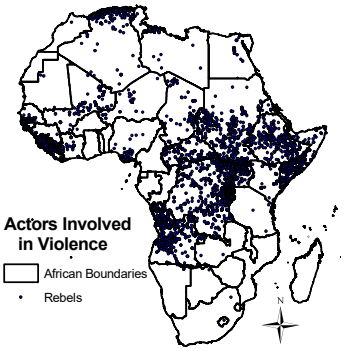

Figure **2b**

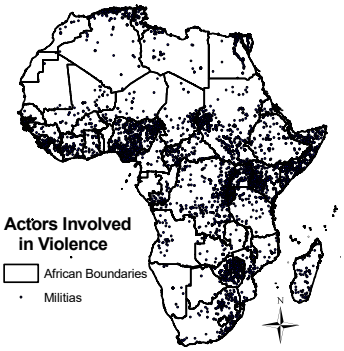

Figure **2c**

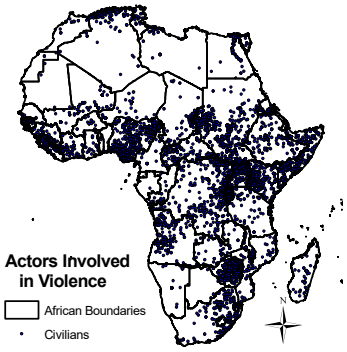

Figure **2d**

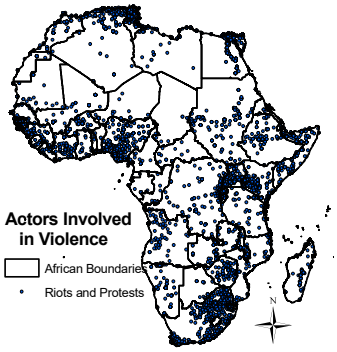

Figure **2e**

Figure **2f**

Figures 3*a* and 3*b* 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.

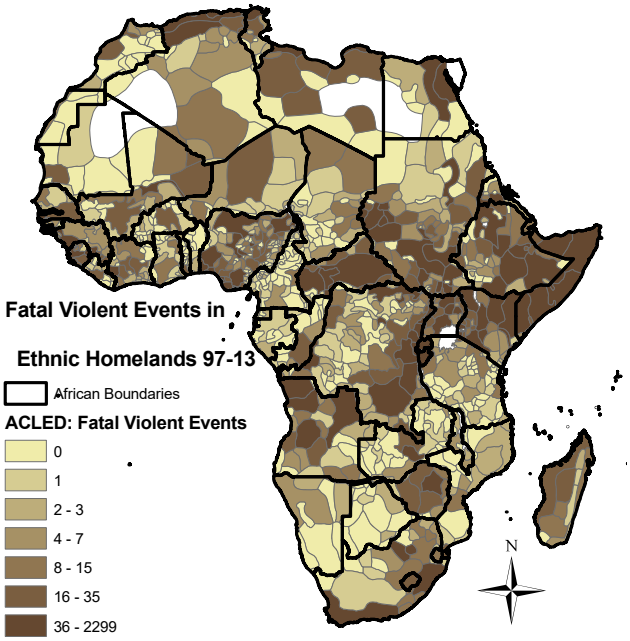

Figure **3a**

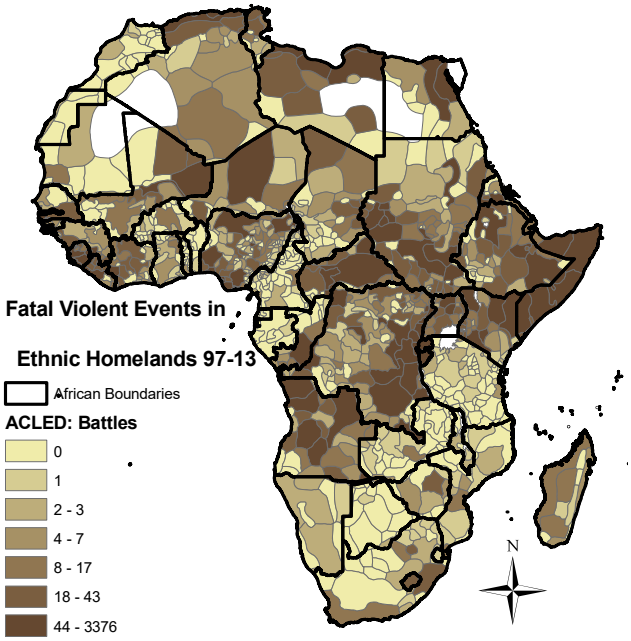

Figure **3b**

1.2.2 UCDP

Figure 4 plots the spatial distribution of all conflict events in the UCDP GED 1.5 database (Sundberg, Lindgren, and Padskocimaite (2010), Sundberg and Melander (2013)). This database focuses on deadly events that are associated with a major or minor civil war, as classified by the widely-used PRIO Armed Conflict Database (ACD); this database distinguishes between a minor and a major civil war when the conflict results in 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 $1989-2010$. Events can be of three types; 88% are daily events, 9% are classified as summary events, and 3% are classified as continuous. We keep all three types of conflicts, though we have experimented dropping summary and continuous events finding similar results. As with ACLED we drop imprecisely georeferenced incidents which amount to 18% of the sample.

UCDP consists of three main types of conflict, each focusing on a different aspect of warfare:

(1) state conflict where government troops and state-backed militias fight either against rebel groups or 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 against each other, without a 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 the 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 violence against civilians takes the maximum values. South Africa is by far the country with the most non-state conflicts (2,125 events) accounting for roughly 60% of all such events in the UCDP. This reflects the intense fighting of ANC and 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 such events.

Appendix Table 6 gives the number of state conflict, one-sided violence and non-state conflict events by country.

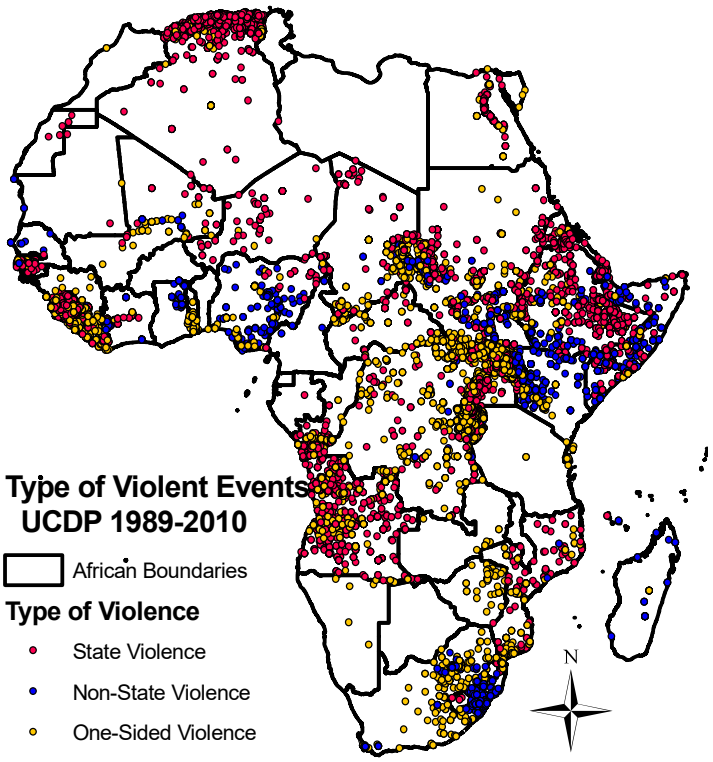

Figure **4**

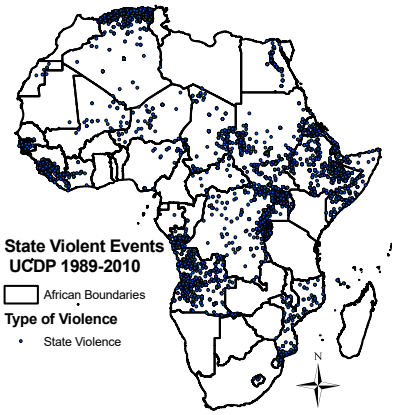

Figure **5a**

Figure **5b**

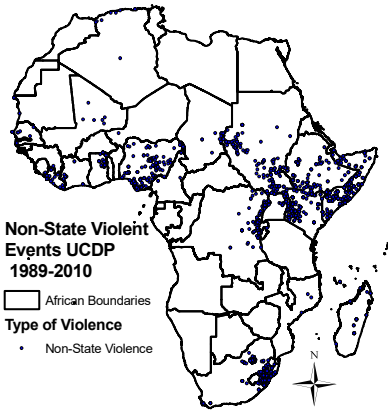

Figure **5c**

1.2.3 ACLED and UCDP

Appendix Table 7 gives the correlation matrix of the various conflict measures. The correlation between battles and violence against civilians 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).

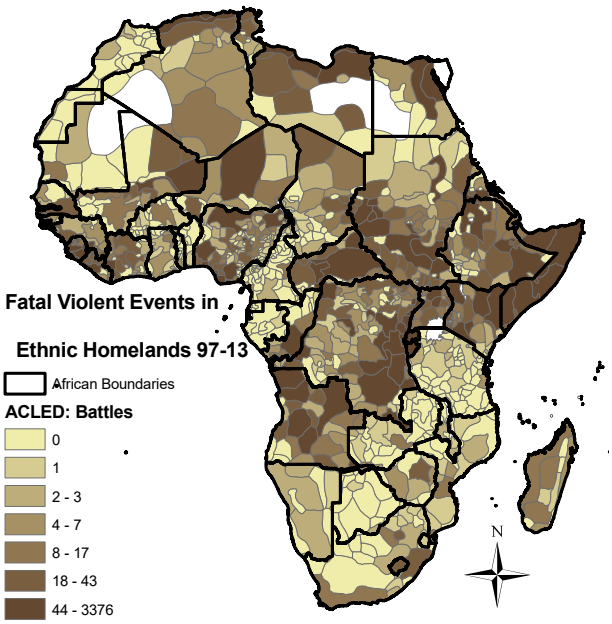

Figure **6a**

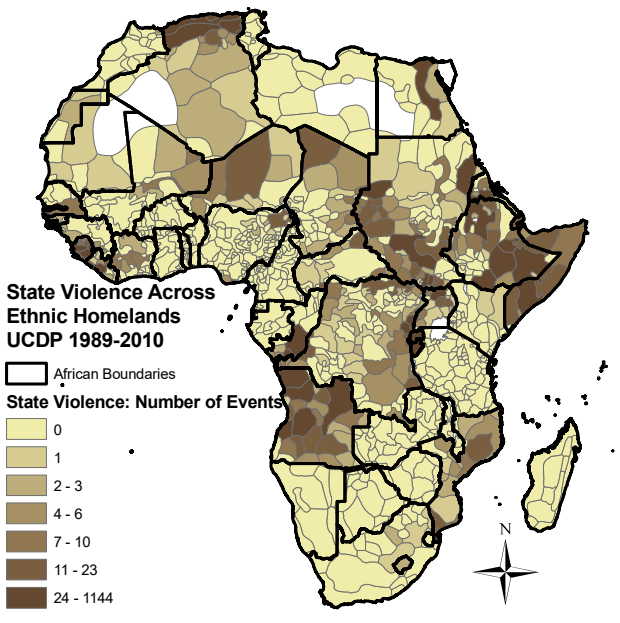

Figure **6b**

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

Similar to Section 3 of the main text, 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 such considerations among Europeans while drawing the colonial borders in the late 19th century.

Appendix Tables 8 reports linear probability model (LPM) estimates across 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 in the main text 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. The results are similar if we omit these variables.

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. As Michalopoulos, Putterman, and Weil (2015) show these measures correlate with contemporary proxies of human capital and individual-level well-being, hence it is interesting to examine whether they correlate with ethnic partitioning. 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 an indicator that equals one for societies with some inheritance rule for property, respectively.[2] 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) classifies 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). Gennaioli and Rainer (2007) and Michalopoulos and Papaioannou (2013) show that pre-colonial political centralization correlates with contemporary regional development. Hence, it is worthwhile to examine the association between political centralization and ethnic partitioning. 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 this proxy of within-group inequality -that may also reflect the degree

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

of pre-colonial economic and institutional development.

In column (8) we examine the role of family institutions, as captured by a polygyny indicator. This is a useful check as groups practising polygyny differ considerably from those that do not (see Fenske (2014a)). The association is 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. The same applies when we use goodness-of-fit measures of non-linear limited dependent variable models (logit and probit). And while Murdock's 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 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 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, as in some instances colonial powers did take into account local politics and geography). Inference requires that partitioned 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 geographic, ecological, natural resource, and location features to ethnic partitioning. The unit of observation is an ethnicity-country region which is our benchmark unit of analysis linking partitioning to conflict in the main body of the paper (Sections $4-5$). 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). Panel A shows that split and non-split country-ethnic homelands are comparable with respect to their land area,

water bodies, elevation, land quality for agriculture, malaria suitability, 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 reflects the explanatory power of ethnic partitioning once we net out the impact of country fixed effects, is zero. Unsurprisingly, the only differences between split and non-split groups within countries are with regards to distance to the border (column (13)) and distance to the capital (column (14)).

Yet, in Panel B we show that in the border sample these differences are neutralized with partitioned and non-split groups scoring comparable values in the respective distance terms. 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 the variable indicating the presence of rivers enters marginally significant the implied 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 negligible; 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 ethnic homelands. For each variable we report summary statistics for partitioned and non-split groups and the corresponding test-of-means and medians equality. Panel A focuses on all 825 homelands, while Panel B reports statistics for ethnic regions close to the national border (using the median distance from the centroid of a group; 102 kilometers). Standard 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).

All Civil Conflict Incidents (ACLED) 83% of all ethnic areas have experienced at least one conflict event (including riots and protests) over the period $1997 - 2013$. 87.8% of partitioned ethnicities experienced conflict, while the likelihood of a civil conflict incidence for non-partitioned ethnicities is 6.3 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 points 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 column (3) is highly significant; the median number of events for partitioned groups is 13 while for non-split groups is just 3. To account for outliers in columns (4) and (5) we exclude ethnic homelands where capitals fall and in columns (6) and (7) homelands where the number of conflict events exceeds the $99th$ percentile. The mean differences for split and non-split ethnicities (17 and 22, respectively) become statistically significant. There are also large differences between partitioned and non-split ethnicities with respect to the median value of conflict events (13 compared to 3). The differences in conflict intensity between partitioned and non-split ethnic groups are also sizable when we focus on areas close to the border (Panel B). While the average (median) number of all civil conflict incidents for partitioned ethnicities is 66 (12), for non-split ones the average (median) is 32.3 (1). These simple mean and median tabulations across ethnic homelands (825 ethnicities) clearly illustrate the negative repercussions of ethnic partitioning on contemporary African development.

Battles between government troops, rebels, and militias (ACLED) On average 72.5% of split-by-the-national-border ethnic homelands have experienced at least one battle between government forces, rebels, and 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 (of 7.5 and 17.5 percentage points) are statistically significant at standard confidence levels. The same pattern emerges when we examine the association between ethnic partitioning and the intensity of battles between government troops, militias and rebels. 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 the civilian population is 0.78, while the corresponding probability for non-split ethnic homelands is 0.68. The difference is larger (0.15) when we focus on ethnic homelands close to the national border. The intensity of civilian violence is also linked to ethnic partitioning. 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 for 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. This is hardly surprising as riots and protests mostly take place in capitals and do not usually reflect ethnic-based political violence.

State Conflict (UCDP) The tests of means reveal sizable differences on both the likelihood and intensity of state-based conflict (as recorded by the UCDP-GED database). The likelihood of state conflict (over $1989 - 2010$) in split homelands is 44.6%, while in non-split homelands the likelihood is 15 percentage points lower, 29.9%. Again the difference is larger (close to 24 percentage points) when we focus on border areas, illustrating that ethnic partitioning matters crucially for two-sided conflict between national government forces and rebels. 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 likelihood for non-split homelands is 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 in partitioned as compared to non-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 4.7 events in the sample of non-split groups (in both the full and the border

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 with respect to conflict where government is neither directly nor indirectly involved. The differences in the likelihood of non-state conflict between split and non-split ethnicities are very small. And conflict intensity, as captured by the number of events, is somewhat higher in non-split ethnic areas; yet even this pattern is driven by outliers. When we drop the top 1% of the distribution, 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-specific constants (in Panel A) and country-fixed-effects linear probability model (LPM) estimates (in Panel B). [The table "mirrors" Table 2 in the main body of the paper.] The ethnic partitioning index is always statistically significant (usually at the 1% 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).[3] The LPM specifications suggest that there is an 8% -10% increased

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

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 entire sample period ($1997 - 2010$). Please also note that the coefficient on the ethnic partitioning index is quite stable across the various permutations (especially in the full sample). Moreover, the coefficient in the full sample with the rich set of controls and country fixed effects (0.51, in column (4)) is quite similar to the estimate of the more parsimonious specification (0.48, in column (1)), where we simply control on log land area, log population around independence and the presence of water bodies. Since the model fit increases considerably when we add country constants and the location controls, this suggests that the significant link between civil conflict and ethnic partitioning is unlikely to be driven by hard-to-observe omitted factors (Altonji, Elder, and Taber (2005), Oster (2015)).

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 fatalities and conflict duration. The linear probability estimates suggest that the likelihood of deadly conflict is 10 percentage points higher for partitioned ethnic homelands. Moreover, on average conflict lasts 25% longer in the homelands of partitioned groups.

Appendix Table 13 associates battles between government troops and rebels/militias, one-sided violence against civilians and riots and protests with ethnic partitioning. There is a 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)). The estimates in columns (3) and (4) show that violence against civilians 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). Similarly to our results in the main text, ethnic partitioning is not systematically linked to riots and protests.

Appendix Table 14 repeats our analysis linking conflicts by actor to ethnic partitioning focusing on the period $1997 - 2010$ and it corroborates our key findings. [The table is similar to Table 5 in the main part of the paper.] Both the negative binomial ML and the LPM specifications reveal large effects of ethnic partitioning on conflict between government forces and rebels/militias that oppose the state. Ethnic partitioning is also linked to a higher incidence and intensity of violence against the civilian population, though quantitatively its impact is smaller than the one on battles involving government forces. And, ethnic partitioning is unrelated to riots and protests. The specifications in columns (5) and (11) further show that

military interventions from neighboring countries are way more common in the homelands of split ethnicities; this finding, that is in line with the anecdotal evidence and the narrative of the African historiography, reveals a key mechanism linking ethnic partitioning to conflict. Split groups are often used by the governments of adjacent countries to destabilize their neighbors, at the same time partitioned ethnicities may get assistance from their co-ethnics on the other side of the border when they stage a rebellion against the national government or when they face repression. In contrast, as specifications (6) and (12) show 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 adjoin partitioned ethnicities. 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 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. The Poisson estimates (that exclude outliers) imply that partitioned ethnic groups experience an increase in civil conflict of approximately 45% ($\exp(0.38) - 1 = 0.46$) as compared to non-split ethnicities. $SPIL$

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 violence against civilians 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 stage 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

Fourth, while the analysis reveals considerable spillovers of ethnic partitioning to nearby ethnic regions (especially when focusing on the likelihood of conflict and on conflict duration), we also estimated specifications without accounting for the share of nearby ethnic groups (in the same country) that have been partitioned ($SPLIT$). Appendix Table 18 reports the results. $SPLIT$ enters with a positive and significant coefficient in all permutations; this applies both to 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 ethnicities.

3.2.5 Alternative Index of Ethnic Partitioning

Fifth, we repeated estimation with an alternative measure of ethnic partitioning using the 5% land threshold to identify split groups. When we use this lower cutoff we identify 266 partitioned ethnicities, while the remaining 559 groups are classified as non-partitioned. 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 (excl. riots and protests) over the sample period (in Panel B). The ethnic partitioning indicator enters all NB-ML specifications with a positive and highly significant coefficient. The linear probability model estimates are stable ($0.065 - 0.07$) both in the full sample and in

the border sample, implying that compared to non-split homelands, the likelihood of a main conflict is approximately 6.5% 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-specific 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 nearest 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 column (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 percentage points when half of the neighboring ethnic homelands are partitioned (as compared to ethnic homelands where none of the nearby groups is split). Consistent 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 estimated specifications adding on top of the country constants ethnic-family fixed effects. These models are very restrictive, as we have 96 ethnic families; yet 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 columns (1) and (4)), linear probability model (LPM) estimates with the binary index of conflict as the dependent variable (in columns (2) and (5)), and Poisson ML

specifications with the duration of conflict (in years) as the dependent variable. The NB-ML estimate on 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 prolonged warfare. Moreover, the coefficient on $SPIL$ is positive (and in the LPM significant at the 1% level), implying considerable spillovers.

Perhaps more importantly, the coefficient on ethnic partitioning 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 increase, adding the ethnic family fixed effects yields a comparable estimate, 0.36. The same applies to the LPM specifications. The coefficient estimates on $SPLIT$ in the specifications with only country fixed effects range from $0.066 - 0.089$ similar to the specification that adds the 96 ethnic family constants, 0.066. 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. The regional classification follows Nunn (2008). 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, as in this region Europeans established sizable colonial communities. In columns (5)-(6) we drop Western African countries, as some of 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 that tends to have long-lasting effects. 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.[4] 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 matrices (linear in distance and quadratic in distance) and three different models: (*i*) simple spatial autoregressive model that accounts for conflict in nearby regions; (*ii*) Durbin's model that accounts for the impact of the various controls of neighboring homelands; and (iii) the generalized spatial model that allows both for a spatial autoregressive term and spillovers of all controls. 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 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 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 of each homeland) and conditioning on total conflict within each ethnolinguistic family (netting out conflict in each homeland). The cross-sectional estimates (with regional constants only) reveal that conflict is higher when an ethnic homeland falls in a conflict-prone country and when there is a lot of conflict across groups in the same family. 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 and Political Centralization

Eleventh, we estimated specifications linking contemporary conflict to ethnic partitioning controlling for the historical legacy of violence. Appendix Table 26 reports the results. Following

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

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), 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)). Conditioning on pre-colonial kingdoms and empires is also useful as we account for any long-lasting impact of political centralization (Gennaioli and Rainer (2006, 2007) and Michalopoulos and Papaioannou (2013)). The coefficient on the ethnic partitioning index remains positive and highly significant.

3.2.12 Accounting for Regional Income Differences

Twelfth, given the significant cross-country 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 events, battles between government forces and militias/rebels, 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.[5] 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).[6] This set of results, therefore, suggests that the effect of ethnic

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

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

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 involving government troops and rebels. Moreover, foreign interventions from neighboring countries' troops is much more likely 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 in the 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 area, a value of one when country-ethnic areas experience conflict based on only one of the two databases, and a value of zero when both datasets indicate the absence of conflict. We also defined a binary index which equals zero when either of the two (or both) databases indicate that no conflict has occurred and the value of one when an ethnic area has experienced some conflict according to both the ACLED and the UCDP. 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-stats $1.5-1.6$). The linear probability model estimates imply that the likelihood of conflict reported in both civil conflict datasets is $6.5\%-9\%$ higher for partitioned ethnic homelands. This estimate is quite similar when we solely rely on ACLED (Table 2) or UCDP (Table 6).

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 the ACLED's and UCDP's 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. The LPM coefficient on $SPLIT$ in columns (2) and (4) suggest that state-driven conflict between government troops and rebels is 11% more likely in the historic homeland of partitioned ethnicity. There is also a somewhat

weaker (though still significant) association between violence against the civilian population and ethnic partitioning. Yet, as in the main part of the paper, this effect is quantitatively smaller as compared to ethnic partitioning's impact on two-sided violence.

4 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. Based on previous works on 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 heterogeneity is also helpful as it sheds light on how and under which circumstances ethnic partitioning spurs conflict.

4.1 Heterogeneity w.r.t. Group Characteristics

In Appendix Table 31 we examine whether specific ethnicity-country characteristics mitigate or accentuate the adverse effect of ethnic partitioning on conflict. Panel A reports negative binomial maximum likelihood estimates (NB) with country-specific 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 underlying characteristic.

In column (1) we explore whether the relationship between partitioning and conflict depends on whether a split region is surrounded by a high or a low share of split groups. These models can be thought as examining possible interactive effects between the ethnic partitioning index ($SPLIT$) and the spillovers measure ($SPIL$). The coefficients are similar in the two sets of split groups. This applies both to the NB and the LPM specifications. So, the adverse impact of partitioning on local conflict is not systematically influenced by the frequency of adjacent split groups. We further searched for potential interactive effects between ethnic partitioning and partitioning in nearby ethnic regions augmenting the baseline specification (equation (1) in the main paper) with an interaction between $SPLIT$ and $SPIL$. The results (not shown) reveal that the interaction term of the ethnic partitioning dummy with the share of adjacent split homelands ($SPLIT*SPIL$) enters with a small and quite unstable coefficient; this implies that the main spillover index (SPIL) adequately captures the spatial externalities of partitioning.

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. In case of more than two partitions we use the population share of the largest neighboring partition. This test is related to 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 cross-country evidence). The results show that the partitioning - 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 index to differ for ethnicities belonging to an ethnic family that consists of many (few) groups in the same country. Ethnic partitioning is systematically linked to conflict for both sets of groups.

In column (5) we address a similar question. We investigate the role of ethnic similarity in the immediate neighborhood of each partitioned group, allowing $SPLIT$ to differ for ethnicities where most of adjacent groups are from the same ethnolinguistic family and those split groups where most neighbors come from different ethnic families. The NB estimates reveal that split groups surrounded by a large share of ethnically similar entities experience on average more conflict. This 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 (2015) 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). 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.

In column (6) we examine whether being split by a squiggly or a straight-line-like border affects the relationship between partitioning and conflict. Following Alesina, Easterly, and Matuszeski (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 weakly indicate that ethnic partitioning is more detrimental for groups split by low-fractal boundaries. 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 $SPLIT$ for the two sets of partitioned groups.

In column (7) we allow the coefficient on ethnic partitioning 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 is Africa are intersected by borders that were separating colonial powers. The remaining $30\% - 35\%$ of partitioned groups are split by political boundaries that were internal administrative borders (such as those of French West Africa). The NB estimates in (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 more countries affects the ethnic-partitioning-conflict nexus. In our sample, 69% of partitioned groups belong to 2 contemporary states with the remaining 31% being split between three, four, five and six countries. Ethnic partitioning is associated with more conflict incidents and a higher likelihood of conflict in both sets of split groups.

Overall the results in 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.

4.2 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 country-level aspects related to diversity, 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, linguistic or religious fractionalization; we do so by allowing the coefficient on ethnic partitioning to differ for countries scoring above and below the median of the widely-used measures of ethnic, linguistic, and religious fractionalization of Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003). While the NB estimates in Panel *A* show that the coefficient on $SPLIT$ is somewhat larger for more heterogeneous countries, the LPM specifications (that are less sensitive to outliers) do not reveal any differences. This applies to all three measures of fractionalization. So, overall the estimates suggest that the impact of ethnic partitioning is quite homogeneous and does not depend on the degree of fractionalization.[7]

In column (4) we examine whether ethnic partitioning has a differential impact on conflict for landlocked countries (e.g., Chad, Burkina Faso, Rwanda, Mali, the Central African Republic). As Collier (2007) notices, the peculiar design of African colonial borders and the division of the continent among colonizers, resulted in African having the highest share of landlocked countries across continents. The NB and the LPM results suggest that ethnic partitioning is particularly deleterious for ethnicities in landlocked countries. This result is interesting, as it shows that two key aspects of the Scramble for Africa, ethnic partitioning and creating many landlocked countries, interact crucially.

In column (5) we allow the impact of the ethnic partitioning ($SPLIT$) to differ between large and small countries (using as a cutoff the cross-country median of land area). 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.

5 EPR-based Analysis. Further Evidence and Sensitivity Checks

5.1 Descriptive Evidence

Appendix Table 33-Panel *A* tabulates ethnic partitioning with EPR's [Ethnic Power Relations database, Wimmer, Cederman, and Min (2009)] classification on whether a group has engaged in a civil war with an explicit ethnic angle or suffered from political discrimination from the national government over the period $1960 - 2010$. EPR reports data on an ethnicity's degree of participation in the central government (the group is classified either as junior, or senior, or

[7] We also experiment with the various ethnolinguistic fragmentation measures of Desmet, Ortuño-Ortín, and Wacziarg (2012) finding similar results (of not much heterogeneity). The results are similar when we group countries based on the degree of cultural fractionalization (using data from Fearon (2003)), ethnic and religious polarization (using the index of Montalvo and Reynal-Querol (2005)), or ethnic inequality (using data from Alesina, Michalopoulos, and Papaioannou (2015)).

dominant partner, or as having monopoly power) and its exclusion from national power (that is further decomposed into being either discriminated against, or powerless, or autonomous). The data cover 196 politically relevant ethnicities from 40 African countries. Using a multitude sources, we linked EPR groups to Murdock's map ethnicities. Our concordance assigned EPR's 196 groups to 593 ethnicities in Murdock (1959) 234 groups (39.5%) are partitioned, while 359 (60.5%) are classified as non-split. We expect this concordance to be of great usefulness beyond the current paper.

Appendix Table 33-Panel *B* tabulates the trichotomous index of political violence for partitioned and non-split ethnic groups. The construction of the ordered index of political violence follows Besley and Persson (2011) who propose studying jointly one-sided violence (repression) and two-sided conflict (civil war). The trichotomous index of political violence equals two if the ethnic group has been involved in an ethnic civil war; equals one when the group is subject to political discrimination from the government but not engaged in civil war; and zero when the ethnicity is neither discriminated nor involved in a civil war.

The descriptives reveal some interesting patterns. First, the likelihood of civil wars is significantly higher for partitioned ethnicities. Second, the likelihood of facing discrimination from the national government is also higher for split groups. Third, in Africa during the post-colonial period, political violence results in most instances in major wars, rather than repression. Fourth, the likelihood of ethnicities experiencing political discrimination from the central government without leading to (or being an outcome of) a civil war, is 9.4% for split groups and 6.7% for non-split groups. This implies that relatively speaking ethnic partitioning matters more for two-sided rather than one-sided violence.

5.2 Exclusion from the Central Government and Ethnic Wars

Split groups are more often discriminated by the national government and more likely to engage in ethnic wars. Nevertheless, as discussed in the main body of the paper these groups do also participate sometimes in the various government coalitions resulting empirically in an insignificant relationship between partitioning and the probability of a group having ever been a junior, senior, or dominant partner in the central government. This nuanced political status of split groups suggests that the vicious cycle of discrimination and conflict in which they are often embroiled, comes into play primarily when excluded from the governing coalition. In supplementary Appendix Table 34*A* we show this to be the case both in the cross-section and in the time series of the EPR.

Starting from the cross-section we split the sample into two groupings [see columns (1) and (2)]. The first cluster includes the 331 ethnicities that have been excluded from the

central government at some point during the $1960-2010$ period (56% of the sample) whereas the remaining 262 are groups that have always participated in some capacity in the central government (44% of the sample). Note that exclusion from the central government does not necessarily translate into being discriminated against by the latter. In fact, among the 331 groups that at some point were excluded only 110 were also discriminated with the vast majority being classified as powerless by the EPR (and a small fraction of those excluded, 14%, having some regional autonomy). Among the 262 groups that have always been included in a governing alliance partitioned groups are no more likely to have participated in an ethnic war compared to non-partitioned ones (the probability of ethnic war in this sample is 7%). The estimated coefficient in column (2) is very small (roughly 1%) and highly statistically insignificant. On the contrary, among the 331 groups that have been excluded at some point split ethnicities have a 17% higher probability of having participated in an ethnic war (the conflict incidence is 37% in this sample).

In columns $3-8$ we further examined the impact of ethnic partitioning on the vicious cycle of discrimination and conflict that seems to characterize African politics exploiting the yearly variation of a group's political status. Specifically, we estimated year-country-ethnicity regressions with country fixed effects and yearly dummies corresponding to the period $1960-2010$ associating the onset of ethnic wars in year t with ethnic partitioning separately for ethnicities that have been excluded from the central government in any of the past three (or five) years and those that have been included in the national power-sharing coalition during the same period. [Cederman, Wimmer, and Min (2010) show that exclusion from political power is a strong predictor of ethnic-based civil wars.] The estimates imply that in periods during which groups are excluded the likelihood of ethnic war onset is roughly 0.55% higher for partitioned as compared to non-split groups. The magnitudes are large as the likelihood of yearly war onset across the 593 groups during the post-independence period is 0.73%. In contrast, the difference in the likelihood of civil war between split and non-split ethnicities when they are both part of the central government is much smaller (0.18%). We verified this pattern by pooling all year-country-group observations between $1960-2010$ in the 40 African countries and interacted the ethnic partitioning index with an indicator that takes the value of one when a group is excluded from the central government in any of the past three years in column (5) (we use a 5-year window in column 8). The interaction terms between exclusion from the national government and ethnic partitioning enter with positive and highly significant estimates (roughly 0.50%), implying that the likelihood of civil war onset increases considerably for partitioned ethnicities when they are left out of the governing coalitions. In these specifications, the direct impact of political exclusion is also significantly positive (around

0.90%), while the direct impact of ethnic partitioning is positive but small and insignificant (around 0.04%). Appendix Table 34*B* replicates the specifications of Appendix Table 34*A* using the 5% land threshold for classifying split groups finding similar results.

These findings reveal that the link between partitioning and political violence established in Table 7 of the main text is magnified in presence of non-cohesive political institutions manifested in the exclusion of groups from central power. The rationale for why split groups react disproportionately is straightforward. Partitioned ethnicities have a lower opportunity cost of conflict afforded by the presence of co-ethnics on the other side of the border, hence are more likely to react violently against their exclusion from the governing coalitions. In equilibrium the government is also discriminating against partitioned groups but both instances of political violence have the same common origin that of non-inclusive political institutions.

5.3 Alternative Estimation Techniques of the Ordered Index of Political Violence

In the main body of the text when we link the ordered $(0, 1, 2)$ index of political violence to ethnic partitioning we use the fixed-effects logit estimation of Ferrer-i-Carbonell and Fritjetrs (2004). This is a three-piece estimation method which involves the following steps. First, we estimated the mean value of the trichotomous index for each country. Second, we defined an indicator that takes on the value of one for ethnicities of the ordered index above the country mean (estimated in step 1) and zero otherwise. Third, we ran a conditional logit model with maximum likelihood with this binary index (defined in step 2) as the dependent variable. We also run multinomial logit specifications that allow estimating separately the impact of ethnic partitioning (and the other controls) on repression and ethnic wars. In line with the results in Table 7, columns (9)-(12)), the coefficient on $SPLIT$ in the civil war model is large (1.05) and highly significant, while in the repression model the coefficient is smaller (0.36) and not statistically significant.

In Appendix Table 35 we report alternative non-linear and linear estimates discussed below.

LS Estimation Columns (1)-(4) report OLS specifications. While OLS is not well-suited when the dependent variable is trichotomous and non-linearities are present, it allows accommodating for country-fixed effects that account for national-level institutions and politics, the identity of colonial power, ethnolinguistic fragmentation, etc. (see Nunn and Wantchekon (2011) for a similar approach in estimating linear models via OLS in presence of an ordering dependent variable). The coefficient on the ethnic partitioning index is positive and highly significant in all permutations; in line with the baseline estimates (in Table 7) ethnic partitioning

increases the likelihood and intensity of political violence.

Ordered Probit ML Estimation Yet non-linear models are more appropriate when the dependent variable is ordered taking just three values. In columns (5)-(8) we thus report ordered probit specifications, estimated with maximum likelihood. To circumvent the "incidental parameter" problem arising from the non-linearity of the model and the inclusion of country constants, we replace the country fixed effects with regional constants. $SPLIT$ enters all specifications with a highly significant coefficient further stressing the strong impact of ethnic partitioning on political violence. As probit coefficients are not easily interpretable we obtained the predictions of the maximum likelihood estimation for split and non-split ethnicities for each of the three outcomes: peace (neither political discrimination nor ethnic wars), repression (group being only subject to political discrimination from the central government), and civil war (ethnicity participating in an ethnic-based war). Figure 7 plots the estimated probabilities of the specification in column (7). Ethnic partitioning matters crucially for ethnic-based civil war; the estimated probabilities for civil war for split and non-split ethnicities are 31.5% and 19.4%. The non-linear model also yields a higher likelihood of repression for split, as compared to non-split groups (9.05% versus 7.07%, respectively). Yet the difference is small (1.95%).

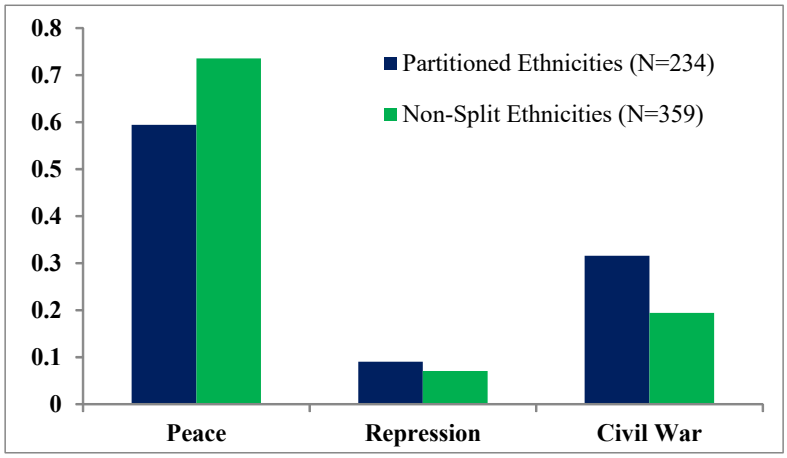

Figure **7**

So, in line with the ordered logit ML estimates in Table 7 (columns (9)-(12)) and the simple linear probability estimates (in columns (1)-(8) of Table 7 and Appendix Tables $36-37$ below) linking ethnic partitioning separately to ethnic wars and political discrimination from the national government, these specifications further show that partitioning is mostly associated with major ethnic-based civil wars, as compared to one-sided violence (where the ethnicity is subject to discrimination from the national government, without being involved in ethnic

conflict).

5.4 Alternative Index of Ethnic Partitioning

In Appendix Table 36 we report within-country linear probability model estimates associating participation in ethnic civil wars (in (1)-(4)) and political discrimination from the national government (in (5)-(8)) 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 columns (1)-(8) of Table 7) the binary ethnic partitioning index enters with a positive and highly significant coefficient. The LPM coefficient in column (4) that besides country fixed effects includes a rich set of controls suggests that the probability to engage in an ethnic war is 12 percentage points higher for split groups. Interestingly, this estimate is quite similar to the baseline estimates linking ethnic partitioning to state-based conflict (in UCDP) and conflict where the government troops participate (in ACLED), see Tables 5 and 6 in the main text. The analogous estimate for discrimination in column (8) suggests that the likelihood of being discriminated from the national government is approximately 8.5 percentage points higher for partitioned ethnicities.

5.5 Regional Effects

Similar to our analysis linking partitioning to conflict, we examined whether a particular region drives the association between ethnic partitioning, discrimination, and ethnic wars dropping iteratively homelands from each African region. In Appendix Tables 37 and 38 we use the baseline ethnic partitioning index and the alternative partitioning index based on the 5% cutoff, respectively. The results show that ethnic partitioning is associated with an increased likelihood of political discrimination from the national government of approximately $6\% - 7\%$. The estimates further show that ethnic partitioning is strongly related to a group's engagement in ethnic-based civil wars across all African regions. The coefficients suggest that the likelihood of ethnic wars is $10\% - 12\%$ higher for split ethnicities. Across all 5 samples (with both the benchmark and the alternative ethnic partitioning index) the coefficient in the civil-war specifications is larger than the analogous of political discrimination. This accords with the ordered logit estimates in Table 7 (and Appendix Table 35) revealing that the impact of partitioning on two-sided violence is quantitatively larger than its impact on one-sided violence.

6 DHS-Based Analysis: Further Evidence and Sensitivity Checks

6.1 Descriptives

Appendix Table 39 - 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. 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. 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 39 - Panel *B* gives summary statistics for the main variables employed in the empirical analysis. Appendix Figure 8 gives the spatial distribution of respondents.

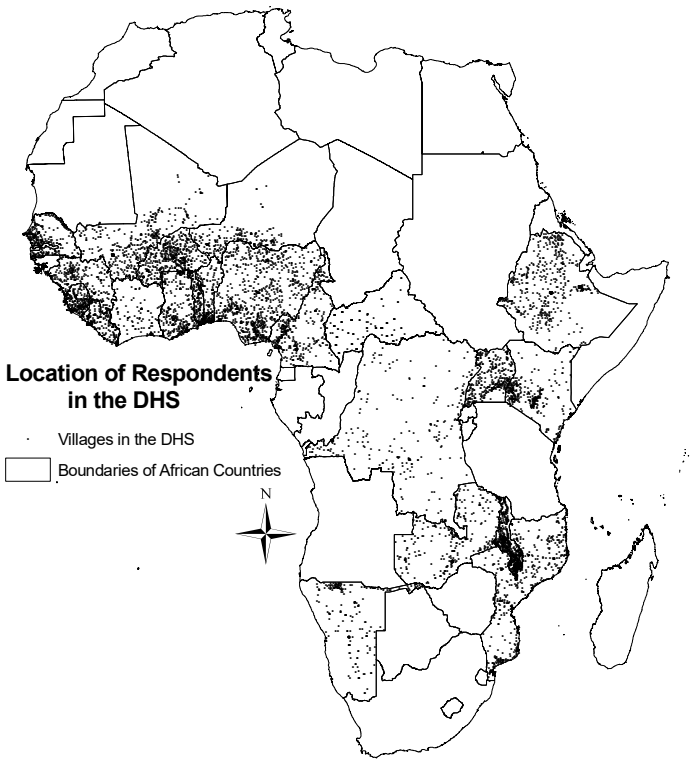

Figure **8**

6.2 Enumeration-Area Fixed Effects Estimates

We repeated estimation exploiting within DHS enumeration-area (village/town/city) variation. Estimating enumeration-area-fixed-effects specifications allows us to control for (observed and unobserved) local factors at a very fine level; as such the ethnic partitioning indicator isolates the impact of the respondent's ethnic origin. Appendix Table 40 reports the within-enumeration-area estimates. These extremely restrictive specifications -that include roughly $8,000$ 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 except when we focus on years of schooling among respondents close to the border. 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 weakly less educated.

6.3 Looking at Movers and Non-Movers

The specifications in Table 9 of the paper leverage the presence of "movers" (i.e., members of split groups residing outside their ancestral homelands) for identification. While in all specifications in Table 9 we control for whether an individual currently resides outside his ancestral homeland, there may be unobserved or hard-to-account for differences between "movers" and "non-movers". Moreover, one may wonder whether by changing one's location the effects of ethnic partitioning can be mitigated. For example, is it the case that individuals from partitioned ethnicities fare better when they move away from their group's ancestral homeland? And does it matter if individuals move to a non-split ethnic area or to a different from their group's partitioned homeland?

To address these questions we repeated estimation in three distinct subsets of the DHS data. The first comprises of respondents that reside inside the boundaries of their ethnicity's ancestral homeland (coined as "non-movers" or "indigenous"). The second and third subsets focus on "movers", i.e., individuals currently residing outside their ethnicity's ancestral homeland. The second subset consists of "movers" residing in non-split ethnic homelands. The third focuses on "movers" residing in homelands of partitioned groups.

Appendix Table 41 reports the results of the within-country analysis for the three different cuts of the DHS data.

Non-Movers First, in columns (1)-(2) and (7)-(8) we focus on "indigenous" ("non-movers"), individuals that currently reside inside their group's ancestral homeland. "Non-movers" consist of 39% of the full sample, in total 34,179 respondents. Focusing on this subset, we are no longer able to distinguish whether the impact of ethnic partitioning is a location-specific attribute (as the georeferenced conflict data suggest) or whether partitioning is associated with the ethnic identity of the respondent (as the analysis with the Ethnic Power Relations Database point to). However, by looking at this sample we mitigate concerns that our baseline estimates (in Table 9) are driven by some sort of selection into migration; this is because we now focus squarely on individuals that live within the confines of their ethnic historical enclaves. The estimates reveal that ethnic partitioning is associated with lower wealth levels and fewer years of schooling. These estimates directly complement the findings linking ethnic partitioning to civil conflict (with ACLED and UCDP), by showing that not only civil wars and one-sided violence against civilians, but also public goods provision is lower in the homelands of partitioned, as compared to non-split, ethnicities.

Movers in Non-Partitioned Ethnic Homelands Second, in columns (3)-(4) and (9)-(10) we examine the link between ethnic partitioning and well-being (public goods provision and education) focusing only on "movers" (i.e., individuals residing outside their ethnic group's ancestral homeland) residing in enumeration areas (villages/town/cities) falling within the ethnic boundaries of non-split groups. These specifications are run on a sample of 30,606 individuals (34.7% of the total number of respondents in the DHS). These estimates do not confound the impacts of residing in a split homeland and identifying with a split ethnicity (something that is inadvertently the case when we focus on "non-movers" only). Moreover, by restricting estimation to enumeration areas within the boundaries of non-split homelands we shed light on whether individuals from partitioned ethnicities can mitigate the adverse consequences of partitioning by relocating towards non-split ethnic territories (that as we have shown are less prone to conflict). The coefficients are all negative, indicating that this is not the case. Respondents identifying with partitioned groups have worse living conditions and are less educated, as compared to respondents, who self-identity with non-split groups.

Movers in Partitioned Ethnic Homelands Third, in columns (5)-(6) and (11)-(12) we focus on "movers", residing in partitioned ethnic homelands. While this is the smallest (of the three) subset, we still have 23,386 respondents (26.5% of DHS sample). The strong negative association between ethnic partitioning (at the identity level) and public goods provision (as reflected in the DHS composite wealth index) and schooling is also present when we focus on individuals residing outside their group's ancestral homeland ("movers"), but residing in the

region of some other split-by-the-national border group. These results that account for selection into location further show that ethnic partitioning is systematically linked to under-provision of public goods and lower levels of schooling. A similar pattern obtains if we perform the analysis among "movers" and split the enumeration areas into rural and urban ones.

Summary Interestingly the point estimates on the binary ethnic partitioning index (at the ethnic identification level) are quite stable across all three subsets ((*i*) "non-movers", (*ii*) "movers" residing in non-split ethnic homelands, and (*iii*) "movers" residing in split ethnic homelands). The coefficient stability across all three data subsets reassures that our baseline estimates do not capture heterogeneity (at least with respect to the relocation dimension) and are less likely to reflect the impact of internal migration. The stability of the implied economic magnitudes further suggests that the adverse legacy of ethnic partitioning cannot be readily alleviated by moving out of one's ancestral split homeland into other homelands (partitioned or not-split) indicating the generalized negative impact of the Scramble for Africa for individuals of partitioned groups.

6.4 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/late 1990*s* African countries, nevertheless, have made considerable steps towards democratic institutions, while over the past decade growth has picked up 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 respondents report their year of birth and repeat estimation separately for "old" and "young" cohorts. Appendix Table 42 reports the results. In odd-numbered columns we restrict estimation to "old" respondents, while in even-numbered columns we restrict estimation to "young" respondents; as the cutoff year we use 1977 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 "old" and "young" cohorts. Moreover, the estimate on the ethnic partitioning index is similar across the two subsets, revealing that the adverse effects of partitioning are quite persistent.

6.5 Summary

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.

7 Conflict in East-Central Africa: A Case Study

In this section we discuss in some detail 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 9).

Tanzania Let us start from Tanzania, a country with little overall conflict; in the 69 ethnic regions of Tanzania ACLED records 250 main conflict incidents (excl. riots and protests). 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) and it mostly 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%). How about conflict in Tanzanian regions, close to the Democratic Republic of Congo (DRC), the country with the highest conflict intensity in Africa? There are zero and two conflict events in the non-split homelands of the Bende and the Fipa, respectively, although both groups reside along the border with DRC. How come? Both groups are non-split, as lake Tanganyika serves as the natural border between the two countries.

Democratic Republic of Congo (DRC) Focusing now on 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; this is more than 20% of all conflict in DRC, a disproportionate fraction considering that the share of the country's population residing in the homelands of the Bashi, the Konjo, and the Ruanda is just 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 (around 2.2%), we observe close to 10% of all conflict in DRC. The UCDP data reveal a similar picture. In the three Rwandan 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 Rwandan 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.[8] 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.

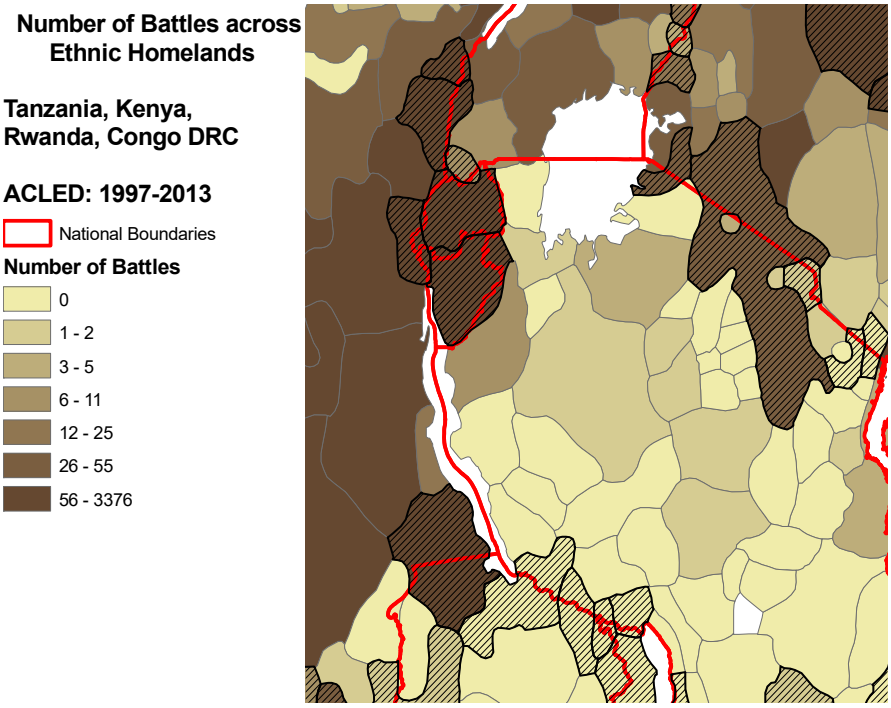

Figure 9

[8] The FDLR and other Hutu-based militias, which fled Rwanda after the 1994 genocide and sought shelter in their homeland in Eastern Congo, had played a major role in the recent civil conflict in Congo. 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.

A Note It is important to keep in mind that the map we use for the historical location of ethnicities was published by Murdock 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, 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; mostly Rwandese who came to the Congo to work in the mines in Katanga as well as those who fled to DRC 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 by many Congolese intellectuals*" (Lemarchand (2008)).

8 A Counterfactual: Somalia-Ethiopia Border

A comprehensive assessment of the consequences of the Scramble for Africa would be feasible conjuring an alternative set of borders and comparing the consequences of the actual arrangement with those of the counterfactual design. Coming up a credible counterfactual is not straightforward. Nevertheless, in this Section we attempt such an analysis for the well-studied case of the Somali which are partitioned across Somalia, Ethiopia, Eritrea, Djibouti, and Kenya. We start by going over the historical account of the colonial boundaries in Eastern Africa. Then we discuss our counterfactual approach and present the estimates. Finally, exploiting an ephemeral rearrangement of the colonial border between Ethiopia and the Italian Somaliland before World War II we conclude by zooming in on the Ethiopia-Somalia border and quantify the impact of the partitioning of the Somali clans, in the Ogaden region in Eastern Ethiopia.

8.1 Historical Background

The case of the Somali is indicative of the devastating consequences of ethnic partitioning in Africa. During the late phase of colonization, the Somali tribes were split among four colonial powers: British East Africa-Kenya, Italian Somaliland, British Somaliland, and French Somaliland, while Ethiopia also got a slice in the Eastern part of the country of traditional Somali homelands in the Ogaden region. The five-edged star in the Somali flag represents this five-way split along with the desire of Somalis to unite into a single state. While the British and

Italian Somalilands merged in the 1960s to form current-day Somalia, Somali clans in Djibouti (French Somaliland), in the Ogaden region in Ethiopia, and the North-Eastern province of Kenya (Woqooyi Bari or Northern Frontier District), were not allowed to join despite their expressed desire in doing so.

As we will show qualitatively and quantitatively below a significant fraction of contemporary conflict in Eastern Africa may be traced to the dismemberment of the Somali.

Figure 10: Somalia Flag

Before presenting the results of our counterfactual exercise, it is useful to briefly discuss the historical account of each of the current partitions of Somalis outside Somalia.

First, regarding the Somali territories in Northern Kenya, in spite of British promises (and an informal referendum where the locals voiced their strong support for joining Somalia), these regions were ceded to Kenyan nationalists in the late 1950s. The Northern Frontier District eventually became part of Kenya at its independence in 1963, as the Kenyatta government was unwilling to let the province join Somalia. A devastating secessionist conflict ensued (the so-called *shifta* war), where ethnic Somali militias -backed by the Somali government in Mogadishu- fought against Kenyan forces. The resulting death toll of more than $5,000$ people (UCDP estimates) further alienated Somalis in Kenya, who have faced discrimination from the central government since then. While the 1967 Memorandum of Understanding between Somalia and Kenya established a cease-fire, the territorial dispute remained unresolved leading to subsequent conflict and repression against the Somalis.[9] In line with this the EPR database classifies the Somalis as discriminated by the national government in Nairobi for almost all of Kenya's post-independence period.

Second, the Somalis residing in Djibouti (French Somaliland) voted in a referendum in

[9] The peak of the conflict was the so-called Wagalla massacre ($10-2-1984$), when according to some estimates $5,000$ Somalis were executed by Kenyan soldiers.

1958 to join Somalia; yet, they were prevented from doing so, both because of voting irregularities and because the Afar and the European population voted for the colony (at the time) to remain with France. Since Somali clans (mostly Esa/Issa) constitute approximately half of the country's population, there has been only minor ethnic conflict, mostly with the Afar, themselves a split group between Ethiopia, Eritrea and Djibouti.[10] According to UCDP's classification the country was under minor civil war in the early 1990s, which an estimate of 500 fatalities.

Third, by far the most contentious issue for the Somalis became the fate of the Ogaden region in Ethiopia. In spite of the area's overwhelming Somali ethnic composition, Ogaden was ceded by the British government to the Ethiopian king Menelik II in the turn of the 20th century (see the 1925 border agreement in Figure 11). However, the Somalis never accepted this and attempted repeatedly to secede from Ethiopia and join the Italian Somaliland. During the short-lived Italian Eastern African ($1936-1941$) and Mussolini's victory over Ethiopia, Ogaden's administration came under the Somalia governorate (provincial government).[11] [Mussolini self-proclaimed that he had liberated the Somalis]. This brief arrangement is illustrated in the 1936 colonial boundaries of Figure 11. Yet after the Italian defeat, the region returned to Ethiopia. Since locals faced discrimination from the Ethiopian government, once British and Italian Somaliland merged to form the independent Somali Republic in 1960, irredentist demands in the Ogaden region peaked; with rebel groups initially pushing for greater autonomy and subsequently for secession. Conflict quickly erupted. Yet in spite of Somali troops' direct interference (in early 1964), the Ethiopian government defeated Somali rebels and troops. The $1963-1964$ war was followed by continuing discrimination against the Somali clans and low-intensity conflict. The EPR classification indicates that Somali tribes in Ogaden have experienced political discrimination from the central government for the entire $1960-2010$ period. The discovery of oil and gas in 1972 and the fall of the Imperial Regime of Haile Selassie in 1974 led to a spike in clashes. The Somali government tried to gain from the regime change in Addis

[10] The partitioning of the Afar has spurred conflict in the North-Eastern regions of Ethiopia. Following land nationalization by the pro-Soviet communist Ethiopian Derg regime in 1975, Afar rebels (Afar Liberation Movement (ALF)) declared an armed struggle against the new Ethiopian government attacking military bases. UCDP event classification reports that the Ethiopian troops offensive that attacked both rebels and the civilian Afar population led to more than $1,000$ deaths in $1975-1976$. Low-intensity conflict and civilian violence continued in the subsequent decades and peaked again in 1993, with Eritrean independence that further split the Afar.

[11] Africa Orientale Italiana was divided into 6 governorates, that were ruled with great autonomy. The Governments of Amhara, Galla and Sidamo, Harar and Shoa constituted the Italian Empire of Ethiopia; these areas correspond to contemporary Ethiopia with the key exception of the Ogaden. The Government of Somalia included the Italian Somaliland to the Ogaden. And Italian Eritrea -which was also ceded to Ethiopia after WWII- corresponds to contemporary Eritea (that seceded from Ethiopia to form an independent state in 1993). Yet Italian Eritrea got also some Ethiopian provinces, as Mussolini wanted to weaken Ethiopia that has fought against him during the Second Italo-Abyssinian War of $1935-36$. Actually the clashes started in the Ogaden, with Ethiopians and Italians quarreling over the exact position of the boundaries.

Ababa and Somali troops invaded Ethiopia in Ogaden in the summer of 1977. Yet Ethiopia secured military assistance from the Soviet Union and Cuba and in a counterattack ousted Somali troops from Ogaden in 1978. According to UCDP estimates, the $1977-1978$ Ogaden War resulted into more than $4,000$ casualties. Throughout the $1980s$ there was sporadic violence, killings, and rebellions.

Although an agreement in 1988 attenuated conflict, the peace agreement left open the issue of the Ogaden region. UCDP description of the agreement is telling: "*the April 1988 agreement is not coded as a peace agreement in UCDP data, since it does not deal with the root issue - or incompatibility - of the conflict: the status of the Ogaden. Over the years Ethiopian President Mengistu had staunchly refused to discuss what he called secondary questions unless Somalia abandoned its claim to the Ogaden. However, the military pressure put on the army in Northern Ethiopia in 1988 forced him to retreat on this, and he agreed to sign an accord that did not solve the basic problem. Mengistu subsequently expressed the hope that "the border issue, which is the main cause of contradiction and conflict between the two countries will be settled legally so as to create, once and for all, a situation in which it will be possible to create all-round cooperation.*" Strife continued, fueled both by local rebels (mostly the Ogaden National Liberation Front) and by the rise of the Islamic Movement in Somalia AIAI (al-Itihad al-Islami: Islamic Union) that was "*aiming to spread political Islam and establish a pan-Somali, Salafist emirate* (UCDP)." According to the UCDP classification the Somali regions have been under low-intensity civil war (defined as yearly fatalities exceeding 25 but being less than 1000) for almost all years since 1993.

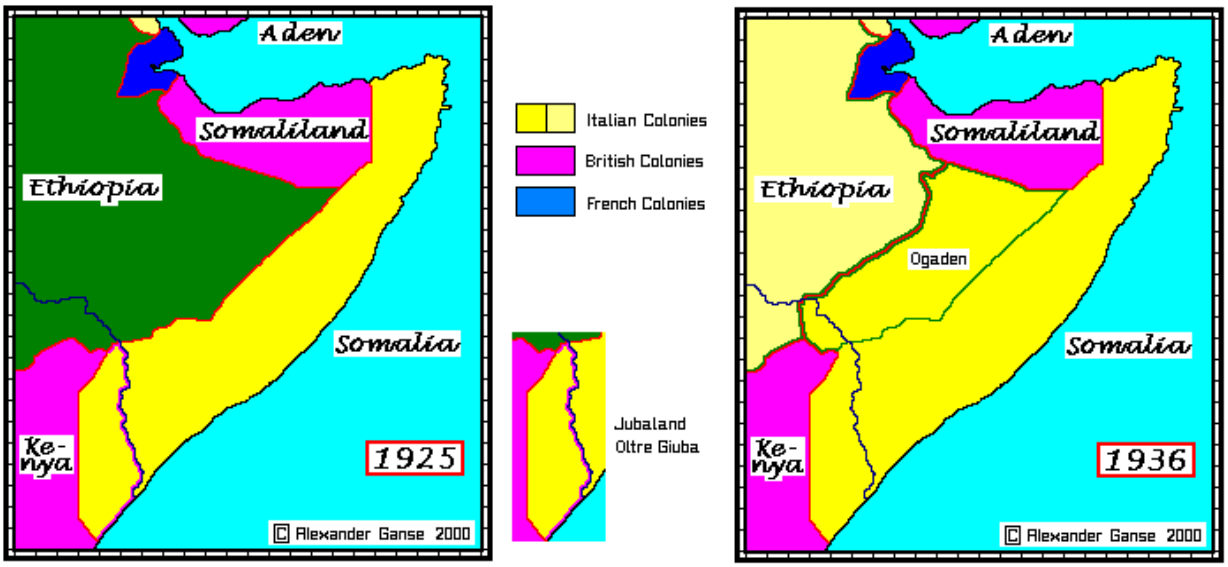

Figure 11: Italian Somaliland - Ethiopia Colonial Border

8.2 Counterfactual

We performed a simple simulation and counterfactual to approximate the impact of the split of Somali clans on conflict. Our out-of-sample counterfactual exercise proceeds as follows.

First, working in the sample of 620 non-split country-ethnicity observations outside the five Eastern African countries with sizable Somali communities (Ethiopia, Somalia, Kenya, Eritrea, and Djibouti) we estimated a negative binomial (NB) ML model associating the number of (deadly) events with the simple, location, geographic, and size controls. To minimize the impact of outliers we estimated the NB-ML model excluding observations with conflict exceeding the top 1% corresponding to more than 706 conflict incidents (though this has no major impact on the analysis).

Second, we obtained out-of-sample predictions of conflict intensity for all ethnic homelands, split and non-split, in the five Eastern African countries (103 country-ethnicity observations). In total there are 107 country-ethnic groups in the five Eastern African countries, but to be consistent with the forecast model, we exclude those with more than 706 events (top 1% of conflict incidents). These forecasts reflect only the impact of location, geography, and size and do not capture the direct impact of ethnic partitioning.[12]

Third, we compared the out-of-sample predictions with the actual number of (deadly) incidents in these five Eastern African countries with large Somali communities.

The forecast model predicts a total of $3,984$ incidents (on average 38.67 per homeland). In practice we observe $5,986$ incidents (on average 55.2 for each country-ethnic region). If we exclude Eritrea that became independent in 1993 and where the partitioning of Somalis does not seem to have played a major direct role, the forecast model predicts a total of $3,807$ events for 94 country-ethnic homelands (average 40.5), while in practice we observe $5,430$ (average 57.8). The simple forecast model predicts conflict well in non-split homelands; the prediction yields an average of 48.6 events per homeland (51.8 if we exclude Eritrea), while we observe 46.9 events (50 if we exclude Eritrea) in the 39 (35) non-split regions in the five (four) Eastern African countries. In contrast, the forecast model underpredicts conflict in partitioned ethnic areas; while we observe $2,685$ events in split homelands in Eastern Africa (on average 68.85 per homeland and 70.8 if we drop Eritrea), the forecast NB-ML model predicts 875 events (on average 22.5 per country-ethnic region and 21.4 if we exclude Eritrea). These tabulations imply that ethnic partitioning accounts for approximately $1,800 - 2,000$ events in the Eastern African countries; this is around a third of total conflict, an estimate quite close to the in-sample estimates (reported in the paper).

[12] Please note that since there are spillovers these forecasts do reflect to some extent spatial externalities. To account for spillovers, we also estimated specifications excluding regions with adjacent split groups. The estimates are similar though more noisy, as the sample drops considerably.

A similar picture emerges when we look at deadly events. The NB-ML model predicts approximately $1,685$ deadly events (on average 16.3 per homeland), while we observe $2,243$ (on average 21.8), and $2,179$ (average 23.2) if we exclude Eritrea. So, ethnic partitioning may account for 558 deadly events, approximately 25% of all deadly incidents in Eastern Africa.

We also implemented a similar test looking at the likelihood of conflict (extensive margin). To do so, we estimated a probit model (with maximum likelihood) associating the conflict indicator with the rich set of geographic and location controls and regional constants for non-split groups, excluding altogether observations from Ethiopia, Kenya, Eritrea, Djibouti, and Somalia; we estimated probit rather than linear probability models to avoid having predicted values outside the $0-1$ interval. Moreover, since we omit the country constants the probit model is preferable as it is not prone to the "incidental parameters" problem. We then obtained out-of-sample predictions of the likelihood of conflict for all ethnic homelands, split and non-split, and compared the predictions with the actual likelihood of conflict. The forecast model predicts that the likelihood of conflict in the 103 homelands of Eastern Africa is around 63% (and 58% when we look at deadly conflict). In reality we observe (deadly) conflict in 72% of the 103 ethnic homelands. This simple comparison implies that ethnic partitioning increases the likelihood of (deadly) conflict by approximately 9% (14%), quite similar to the in-sample linear probability model estimates in the paper.

8.3 A Closer Look in the Ogaden Region

Overview We now take a closer look in Ethiopia in an attempt to quantify the impact of the border settlement between Great Britain, Italy and Ethiopia in the turn of the 20th century that resulted in the partitioning of the Somali tribes residing in the Eastern (Ogaden) province. The sizable share of Somali-tribe lands in Ethiopia was the outcome of many agreements. First, with the 1903 border agreement between British Somaliland (in the North of contemporary Somalia) and Ethiopia, the British rewarded the Ethiopian emperor, Menelik II, with a sizable portion of Somali tribal lands for his support in the Anglo-Somali war. And while the initial treaty included provisions for the region's autonomy, the Ethiopians quickly claimed sovereignty over the area. As the fighting between Somali tribes and the British continued and colonizers wanted Ethiopian help, the colonial powers did not raise the issue. Second, the 1908 Convention between Italy and Ethiopia (that in turn clarified an unofficial agreement of 1897) in its article 8 states that "*all of the Ogaden, and all of the tribes towards the Ogaden, shall remain dependent on Abyssinia (Ethiopia).*" Third, subsequent treaties between Ethiopia, Italy, and Great Britain and Ethiopia's accession to the League of Nations ratified the initial border design, which was demarcated in the thirties ($1932-1934$). Fourth, while during the

Italian rule (1936 − 1941) the border moved westwards and Ogaden became part of the (Italian) Somaliland, when WWII ended the territory was returned to Ethiopia.

UCDP 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.*"

Descriptive Comparisons Murdock's map delineates four Somali tribes that are split by the Ethiopia-Somali border (which mostly follows a longitudinal line), the Ishaak, the Esa, the Mijertein, and the Ogaden.[13] According to the 1960 and the 2000 population estimates, around $3.5\% - 5\%$ of Ethiopia's population resides in these 4 low-population-density ethnic areas. Yet while less than 5% of Ethiopia's population resides in Ogaden, in those four ethnic regions we observe a total of 262 events and 185 deadly events in ACLED over the period $1997 - 2013$.[14] These numbers correspond to 22% of all conflict events, 26% of main events (excl. riots and protests) and 27.5% of deadly events in Ethiopia; in total there are $1,186$ conflict incidents ($1,004$ if we exclude riots and protests that almost exclusively take place in Addis Ababa) and 671 fatal incidents. The considerable impact of the partitioning of the Somali clans in Ethiopia can also be seen by comparing duration and fatalities in the four Somali ethnic regions with the rest of the country. Average (median) conflict duration in the Somali ethnic regions is 8.75 (10) years, while the mean (median) duration in the other 44 ethnic regions in Ethiopia is 5 (3.5) years. Average (median) duration of deadly events in the Somali ethnic regions is 7.5 (7), with deadly conflict in Ogaden lasting for 13 years, while in the other ethnic regions in Ethiopia, mean (median) deadly conflict duration is about half, 4 (2.5). A similar patterns emerges when we use the UCDP datasets that focuses on conflict incidents associated with civil wars. In the four Somali partitioned ethnic areas in Eastern Ethiopia we observe 274 deadly state-driven events, most of them involving the Ogaden clans (245 events). This represents close to half (43% if solely focus on Ogaden) of all state-conflict

[13] The Mijertein and the Ogaden are partitioned at the 5% threshold.

[14] In line with our results there are no riots and protests in these partitioned ethnic areas.

events in Ethiopia that has experienced a total of 566 deadly incidents.

Counterfactual The impact of the partitioning of Somali clans in the Ogaden can be further seen by comparing realized conflict with the out-of-sample forecast (that associates conflict to the rich set of covariates in non-split ethnic regions outside Eastern Africa). The forecast models predict in total 106 events and just 20 deadly events; so the simple comparison between realized conflict in the four ethnic homelands of the Somalis in Ethiopia and predicted conflict (based on population, land area, location, and geography) imply that partitioning has had a major impact, accounting for more than half, approximately two-thirds of conflict in the region. Thus the gap of realized with forecasted conflict (156 events) just in the Ogaden region represents around $10\% - 15\%$ of the overall conflict in Ethiopia (that in total has experienced 1186 conflict events).

8.4 Summary

While the out-of-sample counterfactual is based on many assumptions (including that of no spillovers), the results are quite similar to the in-sample estimates. Moreover, the simple simulation allows approximating the impact of the design of specific borders that have contributed to civil conflict by partitioning specific groups. In the case of the Ethiopia-Somali border, guided by history, we can reasonably assume a counterfactual border design, where the predominantly Somali-inhabited Ogaden was part of Somalia rather than Ethiopia (see Figure 11).

9 Data Sources and Variable Definitions for the Supplementary Appendix Analysis

9.1 ACLED

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 and alternative partitioning index that equals 1 if at least 5% of the historical homeland of an ethnic group is partitioned into different countries. *Source: Calculated intersecting Murdock's (1959) ethnic map of Africa with the Digital Chart of the World (DCW) shapefile. The latter contains the polygons delineating the international boundaries in 2000. Appendix Table A reports the 230 partitioned ethnicities.*

All Civil Conflict Incidents: Sum of all civil conflict incidents according to ACLED. There are 8 event types. (1) Battles without change of control; (2) Battles where rebel groups gain control of the location; (3) Battles where the government regains control of a location; (4) Headquarter of base establishments, where rebel groups establish (via violent or non-violent means) their base; (5) Non-violent conflict events where rebel groups, militias or government forces proceed in non-violent actions (without active fighting) that are however within the context of an ongoing civil conflict and dispute (e.g., recruitment drives, incursions or rallies); (6) Riots and protests; (7) Violence again civilians, where armed groups (rebels, militias or government forces) attack unarmed civilians; (8) Non-Violent transfer of control. 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. We focus on conflict events classified as "high geo-precision quality." See Section 2 for details. We focus in two samples. (i) $1997-2013$ (as in the paper). (ii) $1997-2000$ since reporting increased considerably after 2010. *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 and also when we drop riots and protests). *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 (of any type and also when we drop riots and protests). *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 and also when we drop riots and protests). *Source: ACLED 4.*

Duration Civil Conflict: Duration in years that a country-ethnic homeland has been in conflict (using conflict events of any type and also when we drop riots and protests). *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 and also when we drop riots and protests). *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 and also dropping riots and protests), summed over all years. *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.*

9.2 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.*

9.3 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: Number of years of education. Range from 0 to 24 with a median of 6 years. *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.

Mover (Non-Indigenous) 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/)

Year-of-Birth Fixed Effects: A vector of date-of-birth fixed effects. The 61 dummies correspond to yearly cohorts born between 1935 and 1996. *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 Christians, Other, None. *Source: Demographic and Health Surveys. http://www.measuredhs.com/)*

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

Distance to the Coast: The geodesic distance from each enumeration area (gps coordinates) 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 each enumeration area (gps coordinates) to the nearest national border. *Source: Calculated using ArcGis.*

Capital City Indicator: Dummy variable that takes on the value one for enumeration areas close to the capital city (distance to the capital less than the 25th percentile which corresponds to 130 kilometers).

9.4 EPR Data

Political Discrimination: Binary index that takes on the value of one when a politically relevant ethnic group has been subject to political discrimination from the national government for at least one year during the post-independence period and zero otherwise. An ethnic group is classified as "politically relevant" if at least one political organization claims to represent it in national politics or if its members are subjected to state-led political discrimination. Discrimination is defined as "political exclusion directly targeted at an ethnic community—thus disregarding indirect discrimination based, for example, on educational disadvantage or discrimination in the labour or credit market." "Group members are subjected 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." *Source. Ethnic Power Relations (EPR) vintage 3.01 database. Wimmer, Cederman, and Min (2009).*

Ethnic Wars: Binary index that takes on the value of one when a politically relevant ethnic group has been engaged into two-sided civil war with explicit ethnonationalistic or secessionist aims, for at least one year during the post-independence period. The coding of conflict is based upon the UCDP/PRIO Armed Conflicts Dataset, which defines internal conflict

"as any armed and organized confrontation between government troops and rebel organizations or between army factions that reaches an annual battle death threshold of twenty-five. This definition excludes one-sided conflicts, such as massacres and genocides, as well as communal riots, pogroms, and other non-state conflicts." *Source. Ethnic Power Relations (EPR) vintage 3.01 database. Wimmer, Cederman, and Min (2009)*

Political Violence: Ordered trichotomous $(0, 1, 2)$ index of political violence for politically relevant ethnic groups during the post-independence period. The index takes on the value of two (civil war) when a politically relevant group has been engaged into two-sided civil war with explicit ethnonationalistic or secessionist aims, for at least one year during the post-independence period. The index takes on the value of one (repression) when a politically relevant ethnic group has been subject to political discrimination from the national government for at least one year during the post-independence period, but has not been engaged into an ethnic war. The index takes on the value of zero when a politically relevant group has neither been subject to political discrimination from the national government nor has it been engaged into a civil war with explicit ethnonationalistic or secessionist aims during the post-independence period. " *Source. Ethnic Power Relations (EPR) vintage 3.01 database. Wimmer, Cederman, and Min (2009). The construction of the index follows Besley and Persson (2011).*

10 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 1000*s* 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 1000*s* 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 1000*s* 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 1000*s* 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).*

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

10.1 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 repute 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.*

10.2 Country-level Variables Used in Heterogeneous Effects Section

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

Ethnic, Linguistic and Religious Fractionalization: Index of ethnic, linguistic and religious heterogeneity. Each index reflects the probability that two randomly selected individuals belong to different ethnic, linguistic or religious groups. *Source: Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003).*

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.

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 T. Persson (2011): "The Logic of Political Violence," *Quarterly Journal of Economics*, 126(3), 1411–1445.

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.

Cederman, L.-E., A. Wimmer, and B. Min (2010): "Why do ethnic groups rebel? New data and analysis.," *World Politics*, 62(1), 87–119.

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

Chen, X., and W. D. Nordhaus (2011): "Using Luminosity Data as a Proxy for Economic Statistics," *Proceeedings of the National Academy of Sciences*, 108(21), 8589–8594.

Collier, P. (2007): *The Bottom Billion*. Oxford Universiy Press, Oxfod, United Kingdom.

Collier, P., and A. Hoeffler (2007): "Civil War," in *Handbook of Defense Economics, Volume 2, Defense in a Globalized World,*, ed. by T. Sandler, and K. Hartley, pp. 711–740. Elsevier, Amsterdam, Netherlands.

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 Salience 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.

Ferrer-i-Carbonell, A., and P. Fritjetrs (2004): "How Important is Methodology for the Estimates of the Determinants of Happiness?," *Economic Journal*, 114(2), 641–659.

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.

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.

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.

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. (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.

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 (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.

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.

Oster, E. (2015): "Unobservable Selection and Coefficient Stability: Theory and Evidence," mimeo, Brown University, Department of Economics.

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 (2015): "War and Relatedness," Review of Economics and Statistics, fortcoming.

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
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
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 Confict 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
Fatal Conflict Events	606	8.37	66.23	0	0	127	1558
Fatal Confict 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 (61.3 km, 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	1,460	7.51	41.6%
2011	5,261	8.14	1,460	7.5	27.8%
2012	8,741	13.54	2,347	12.08	26.9%
2013	12,563	19.43	3,511	18.06	27.9%
Total	64,650	100	19,448	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

Conflcit Category	All Events		Fatal Events		Percentage
	Number	Percentage	Number	Percentage	Fatal
Battle-Government regains territory	1,205	1.86	224	1.15	18.59%
Battle-No change of territory	20,892	32.31	9,115	46.86	43.63%
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	1,039	5.35	6.43%
Violence against civilians	20,396	31.56	8,773	45.11	43.01%
Total	64,650	100	19448	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
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 (vintage 4) and UCDP GED (vintage 1.5) 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***	0.0777***	0.0725***	0.0725***	0.0679***	0.0790***	0.0818***	0.0739***
	(0.0171)	(0.0207)	(0.0205)	(0.0178)	(0.0195)	(0.0164)	(0.0222)	(0.0176)
Lake Indicator	0.0866	0.1015	0.1057	0.0808	0.0835	0.0888	0.1176	0.0997
	(0.0678)	(0.0679)	(0.0665)	(0.0799)	(0.0771)	(0.0685)	(0.0785)	(0.0657)
River Indicator	-0.0169	-0.0248	-0.0172	-0.0053	-0.0259	-0.0238	-0.0003	-0.0224
	(0.0430)	(0.0432)	(0.0429)	(0.0419)	(0.0472)	(0.0428)	(0.0426)	(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 variables reflecting ethnic-specific pre-colonial economic, social and political traits (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 ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 9: "Balancedeness Tests". Ethnic Partitioning and Geographic Characteristics within Countries

Panel A: All Country-Ethnic Homelands

	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	-0.1702	-0.0027	0.0023	0.0165	0.0009	-0.0092	-0.0132	-0.0212	-0.0005	-0.0173	-0.024	0.0219	100.8410**	0.1059***
	(0.1045)	(0.0225)	(0.0333)	(0.0365)	(0.0193)	(0.0218)	(0.0189)	(0.0220)	(0.0094)	(0.0274)	(0.0159)	(0.0316)	(13.3237)	(0.0276)
marginal R2	0.003	0.000	0.000	0.001	0.000	0.000	0.001	0.001	0.000	0.001	0.003	0.000	0.154	0.017
Mean DV	2.071	0.108	0.504	0.617	0.407	0.717	0.093	0.084	0.026	0.129	0.039	0.600	107.412	0.505
Observations	1212	1212	1212	1212	1212	1212	1212	1212	1212	1212	1212	1212	1212	1212
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Appendix Table 9: "Balancedeness 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***	-0.0191	0.0864*	0.0010	0.0243	-0.0090	0.0123	0.0095	0.0057	0.0475	0.0118	-0.0268	-1.9752	-0.0128
	(0.1762)	(0.0294)	(0.0399)	(0.0326)	(0.0154)	(0.0214)	(0.0250)	(0.0212)	(0.0045)	(0.0311)	(0.0094)	(0.0303)	(2.8585)	(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 Dep. V	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

Appendix Table 10: Test of Means and Medians for Main Civil Conflict Measures

Panel A: All Ethnic Homelands

	Indicator Likelihood	Number of Incidents - All		Excluding Capitals		Excluding Outliers (top 1%)	
	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=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)	***(0.07)***	***(0.42)***	***(0.00)***	***(0.05)***	***(0.09)***	***(0.02)***	***(0.01)***
<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)	***(0.33)***	***(0.20)***	***(0.03)***	***(0.10)***	***(0.19)***	***(0.03)***	***(0.12)***
<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)	***(0.05)***	***(0.53)***	***(0.03)***	***(0.05)***	***(0.10)***	***(0.05)***	***(0.03)***
<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)	***(0.00)***	***(0.56)***	***(0.00)***	***(0.47)***	***(0.00)***	***(0.21)***	***(0.00)***
<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)	***(0.00)***	***(0.87)***		***(0.31)***		***(0.01)***	
<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	0.000	7.64	0.000	6.44	0.000
difference	**0.147**	**5.696**	**0.000**	**4.278**	**0.000**	**2.998**	**0.000**
difference (p-value)	***(0.00)***	***(0.14)***		***(0.12)***		***(0.07)***	
<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)	***(0.93)***	***(0.37)***		***(0.47)***		***(0.50)***	

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

	Indicator Likelihood	Number of Incidents		Excluding Capitals		Excluding Outliers	
	mean	mean	median	mean	median	mean	median
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<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)	***(0.00)***	***(0.10)***	***(0.01)***	***(0.09)***	***(0.02)***	***(0.00)***	***(0.00)***
<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)	***(0.00)***	***(0.16)***	***(0.15)***	***(0.18)***	***(0.10)***	***(0.01)***	***(0.01)***
<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)	***(0.00)***	***(0.13)***	***(0.01)***	***(0.08)***	***(0.00)***	***(0.00)***	***(0.00)***
<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)	***(0.00)***	***(0.01)***	***(0.01)***	***(0.23)***	***(0.01)***	***(0.02)***	***(0.01)***
<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)	***(0.00)***	***(0.05)***		***(0.06)***		***(0.00)***	
<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)	***(0.00)***	***(0.07)***		***(0.13)***		***(0.00)***	
<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)	***(0.35)***	***(0.39)***		***(0.40)***		***(0.95)***	

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**
	(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
	(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
R-square	0.135	0.429	0.476	0.421	0.093	0.091	0.151	0.342	0.467	0.516	0.516	0.484
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**
	(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
	(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 (NB) 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 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 in the same country. 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 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***
	(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
	(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 in the same country. 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 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	
	NB-ML	LPM	NB-ML	LPM	NB-ML	LPM	NB-ML	LPM	NB-ML	LPM	NB-ML	LPM
	(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
	(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 (NB) 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 in the same country. 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 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 14: Ethnic Partitioning and Conflict Actors over 1997-2010.

	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)
Panel A. Negative Binomial ML Estimates												
SPLIT (Partitioning)	0.5520***	0.5675***	0.1237	0.4065***	1.2291***	0.0602	0.8461***	0.7028***	0.1282	0.5658***	1.1803***	1.0811
	(0.1571)	(0.1476)	(0.0865)	(0.1451)	(0.3299)	(0.2391)	(0.2960)	(0.2921)	(0.2065)	(0.2299)	(0.4290)	(0.9560)
SPIL (Adjacent Split)	0.2273	0.1189	0.0051	0.0349	-0.0842	-1.1965	-0.1332	-0.0912	0.2292	-0.3655	-0.4716	-2.0554**
	(0.3778)	(0.3140)	(0.0920)	(0.3675)	(0.5777)	(0.7729)	(0.4800)	(0.3989)	(0.2856)	(0.4345)	(0.5595)	(1.0131)
Log Likelihood	-2640.84	-3057.78	-669.28	-2426.80	-940.63	-373.20	-953.45	-1120.16	-234.12	-870.70	-383.79	-134.27
Panel B. Linear Probability Model (LPM) Estimates												
SPLIT (Partitioning)	0.1197***	0.0710**	-0.0027	0.0678**	0.0561*	0.017	0.1148**	0.0829**	0.004	0.0749*	0.0638**	0.0455*
	(0.0291)	(0.0317)	(0.0296)	(0.0319)	(0.0321)	(0.0174)	(0.0453)	(0.0417)	(0.0382)	(0.0449)	(0.0314)	(0.0273)
SPIL (Adjacent Split)	0.0633	0.0285	0.0228	0.0459	0.0189	-0.0219	0.0723	0.0523	0.026	0.0711	-0.053	-0.053
	(0.0577)	(0.0642)	(0.0404)	(0.0574)	(0.0469)	(0.0223)	(0.0733)	(0.0657)	(0.0595)	(0.0682)	(0.0474)	(0.0340)
Adjusted R-square	0.453	0.472	0.437	0.424	0.345	0.378	0.467	0.485	0.418	0.436	0.384	0.425
Observations	1212	1212	1212	1212	1212	1212	579	579	579	579	579	579
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 Negative Binomial (NB) 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 (United Nations, 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 in the same country. 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 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**	0.1304**	0.2473***	0.2537***	0.2705***	0.2470***	0.2618**	0.2618**	0.4177***	0.4139***	0.4139***	0.3577***
	(0.0650)	(0.0650)	(0.0699)	(0.0711)	(0.0713)	(0.0742)	(0.1171)	(0.1171)	(0.1241)	(0.1253)	(0.1253)	(0.1228)
SPIL (Adjacent Split)	0.3482***	0.3482***	0.4839***	0.5055***	0.5440***	0.6342***	0.3047	0.3047	0.3959**	0.3425*	0.3425*	0.4491**
	(0.1172)	(0.1172)	(0.1153)	(0.1162)	(0.1189)	(0.1188)	(0.1871)	(0.1871)	(0.1806)	(0.1832)	(0.1832)	(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
Country Fixed Effects	No	No	No	No	No	No	No	No	No	No	No	No
Panel B. Fixed-Effects Poisson ML Estimates (excl. Outliers)												
SPLIT (Partitioning)	0.2546**	0.2372***	0.3712***	0.3892***		0.3435***	0.5054**	0.4868***	0.6954***	0.7213***		0.5835***
	(0.1022)	(0.0794)	(0.1117)	(0.1026)		(0.1016)	(0.2005)	(0.1409)	(0.1754)	(0.1574)		(0.1589)
SPIL (Adjacent Split)	0.1119	0.1443	0.2800	0.3069		0.1975	0.4569	0.4853	0.6001*	0.4922		0.5426
	(0.2264)	(0.2288)	(0.2413)	(0.2346)		(0.2447)	(0.3658)	(0.3998)	(0.3193)	(0.3000)		(0.3986)
Log Likelihood	-19100	-14200	-12000	-11700		-11000	-7200.19	-4746.57	-4061.54	-3934.52		-3697.48
R-square	0.264	0.434	0.527	0.543		0.498	0.210	0.454	0.553	0.596		0.542
Observations	1151	1151	1151	1151		1125	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, Hahn, 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 in the same country.
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 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		
	Battles	Civilian Violence	Riots & Protests	Battles	Civilian Violence	Riots & Protests
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Conditional Negative Binomial ML Estimates						
SPLIT (Partitioning)	0.2589***	0.2229***	0.2841***	0.5204***	0.4289***	0.3690*
	(0.0891)	(0.0831)	(0.0949)	(0.1584)	(0.1461)	(0.1954)
SPIL (Adjacent Split)	0.2976**	0.5722***	0.4158***	0.4559**	0.2999	0.0743
	(0.1462)	(0.1350)	(0.1536)	(0.2256)	(0.2216)	(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***	0.4543***	0.2827**	0.7334***	0.8135***	0.1207
	(0.1089)	(0.1299)	(0.1115)	(0.2577)	(0.2294)	(0.1658)
SPIL (Adjacent Split)	0.5559***	0.3573	0.245	0.5343*	0.1605	0.5298
	(0.1776)	(0.2913)	(0.1945)	(0.2962)	(0.3755)	(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, Hahn, 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-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 (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 in the same 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). 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 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
	(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
	(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 (United Nations, 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 in the same 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). 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 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 Accountring for Spillovers

	All Ethnicity-Country Homelands			Ethnicity-Country Homelands Close to the National Border		
	All	Excl. Outliers	Excl. Capitals	All	Excl. Outliers	Excl. Capitals
	(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**
	(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*
	(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 (NB) 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 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		
	All	Excl. Outliers	Excl. Capitals	All	Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Negative Binomial ML Estimates					
SPLIT-5PC (Partitioning)	0.4977***	0.5385***	0.5024***	0.7601***	0.7597***	0.7390***
	(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
	(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 (NB) 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 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			
	All Events	Battles	Violence	Riots	All Events	Battles	Violence	Riots
	(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
	(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*
	(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
	(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
	(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 (NB) 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 in the same country. 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 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			
	All Events	Battles	Violence	Riots	All Events	Battles	Violence	Riots
	(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
	(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*
	(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 (NB) 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 in the same country. 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 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		
	All Incidents NB-ML	Indicator Conflict LPM	Duration Conflict Poisson-ML	All Incidents NB-ML	Indicator Conflict LPM	Duration Conflict Poisson-ML
	(1)	(2)	(3)	(4)	(5)	(6)
SPLIT (Partitioning)	0.3605**	0.0657**	0.1804***	0.5296**	0.0879	0.3629***
	(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*
	(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 all main civil conflict incidents (of all types) 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 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 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 in the same 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). 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 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	
	All	Border	All	Border	All	Border	All	Border	All	Border
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(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
	(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
	(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*
	(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***
	(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 (NB) 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 in the same country. 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 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 Eucledian Distance			Quadratic in Eucledian 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***	0.0702***	0.0880***	0.0863***	0.0698***	0.0866***
	(0.0245)	(0.0247)	(0.0246)	(0.0245)	(0.0256)	(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**	0.1259*	0.1542**	0.1512**	0.1101	0.1313*
	(0.0731)	(0.0716)	(0.0732)	(0.0725)	(0.0734)	(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 a main conflict (excl. riots and protests) 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 Eucledian distance to the centroid of each country-ethnic region weighting matrix. In columns (4)-(6) we use a quadratic in Eucledian 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 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 Events	Main Events Indicator	Log (1+Main Events)	Deadly Events	Duration Main Events	Duration Deadly	Main Events	Main Events Indicator	Log (1+Main Events)	Deadly Events	Duration Main Events	Duration Deadly
	NB-ML	LPM	OLS	NB-ML	PO-ML	PO-ML	NB-ML	LPM	OLS	NB-ML	PO-ML	PO-ML
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SPLIT (Partitioning)	0.8728***	0.1088***	0.3397***	0.6029***	0.2336***	0.1779**	1.2177***	0.1228***	0.4152***	0.9065***	0.4152***	0.4671***
	(0.1460)	(0.0259)	(0.1039)	(0.1480)	(0.0607)	(0.0710)	(0.2389)	(0.0437)	(0.1004)	(0.2094)	(0.1114)	(0.1267)
Log Conflict Family	0.1887***	0.0239***	0.1402***	0.1588**	0.0935***	0.1075***	0.1333**	0.0203*	0.0917**	0.0967**	0.0969***	0.1110***
	(0.0416)	(0.0074)	(0.0319)	(0.0329)	(0.0197)	(0.0204)	(0.0584)	(0.0107)	(0.0400)	(0.0443)	(0.0363)	(0.0331)
Log Conflict Country	0.1634***	0.0205***	0.1112***	0.1362***	0.0601***	0.0594***	0.1464***	0.0230**	0.1061***	0.1375***	0.0682***	0.0632***
	(0.0307)	(0.0069)	(0.0225)	(0.0287)	(0.0157)	(0.0168)	(0.0347)	(0.0107)	(0.0289)	(0.0370)	(0.0232)	(0.0243)
Log Likelihood	-4031.87	__	__	-3101.18	-2999.45	-2529.91	-1637.02	__	__	-1204.98	-1357.03	-1042.35
Adjusted R-square	__	0.354	0.479	__	__	__	__	0.349	0.406	__	__	__
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-ethnicity 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 estimates 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 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

	All Ethnicity-Country Homelands			Ethnicity-Country Homelands Close to the National Border		
Historical (Pre-colonial)	Conflict	Slave Trades	Kingdom	Conflict	Slave Trades	Kingdom
	(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			
	All Events	Battles	Violence	Riots	All Events	Battles	Violence	Riots
	(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
	(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*
	(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
	(0.1089)	(0.1484)	(0.1689)	(0.2826)	(0.2495)	(0.3287)	(0.2869)	(0.3418)
Log Likelihood	-4107.81	-2916.114	-2874.75	-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
	(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
	(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
	(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 (NB) 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 in the same country. 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 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, Nordhaus et al. (2006)). 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 28: Ethnic Partitioning and Conflict Actors. Linear Probability Model (LPM) Estimates Accounting for Regional Development

	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)	0.1087***	0.0659*	0.0134	0.0581*	0.0652**	0.0063	0.1255***	0.0862*	-0.0020	0.0745*	0.0696**	0.0349
	(0.0280)	(0.0344)	(0.0304)	(0.0309)	(0.0332)	(0.0225)	(0.0435)	(0.0501)	(0.0540)	(0.0448)	(0.0348)	(0.0303)
SPIL (Adjacent Split)	0.1270**	0.1025**	0.0719	0.1690***	0.0677	-0.0111	0.1863***	0.1582**	0.0741	0.1727**	-0.0036	-0.0647
	(0.0532)	(0.0489)	(0.0542)	(0.0574)	(0.0484)	(0.0291)	(0.0622)	(0.0640)	(0.0783)	(0.0822)	(0.0470)	(0.0425)
Log GDP p.c.	-0.0448	-0.0445	-0.0109	-0.0803***	-0.0863*	-0.0314	-0.0607	-0.0752	0.022	-0.1342**	-0.0963*	-0.0189
	(0.0386)	(0.0337)	(0.0311)	(0.0305)	(0.0447)	(0.0351)	(0.0536)	(0.0552)	(0.0704)	(0.0620)	(0.0492)	(0.0294)
Adjusted R-square	0.453	0.472	0.436	0.426	0.349	0.379	0.467	0.486	0.417	0.439	0.388	0.425
Observations	1211	1211	1211	1211	1211	1211	578	578	578	578	578	578
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 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 in the same 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). 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 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			
	Trichotomous All	Trichotomous Main	Binary All	Binary Main	Trichotomous All	Trichotomous Main	Binary All	Binary Main
	(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**
	(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*
	(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

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	
	Trichotomous	Binary	Trichotomous	Binary	Trichotomous	Binary	Trichotomous	Binary
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT (Partitioning)	0.1684***	0.1117***	0.1947***	0.1157**	0.0947**	0.0683**	0.1080	0.0377
	(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
	(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

	Adjacent Split	Population Share	Adjacent Largest	Share Groups Same Family	Share Adjacent Same Family	High - Low Fractal	Across & Within Colony	2-Way Splits vs. More-than-2 Splits
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT - High	0.5394***					0.4216**		
	(0.1667)					(0.1817)		
SPLIT - Low	0.4361***					0.5635***		
	(0.1435)					(0.1325)		
SPLIT - Small Share		0.4927***	0.5012***	0.5897***	0.2148			
		(0.1565)	(0.1707)	(0.1632)	(0.1462)			
SPLIT - Large Share		0.4693***	0.4609***	0.3479***	0.7711***			
		(0.1382)	(0.1250)	(0.1338)	(0.1643)			
SPLIT - Between Colonial Powers							0.6674***	
							(0.1427)	
SPLIT - Within Colonial Power							0.1155	
							(0.1519)	
SPLIT - Two-Way Splits								0.5559***
								(0.1502)
SPLIT - Multiple-Way Splits								0.3572***
								(0.1325)
SPIL (Adjacent Split)	0.4883*	0.4184*	0.4155*	0.4149*	0.4234*	0.4532*	0.4462*	0.4567*
	(0.2727)	(0.2330)	(0.2337)	(0.2375)	(0.2294)	(0.2543)	(0.2291)	(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

	Adjacent Split	Population Share	Adjacent Largest	Share Groups Same Family	Share Adjacent Same Family	High - Low Fractal	Across & Within Colony	2-Way Splits vs. More-than-2 Splits
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT - High	0.0869**					0.0405		
	(0.0396)					(0.0413)		
SPLIT - Low	0.0801**					0.1125***		
	(0.0334)					(0.0320)		
SPLIT - Small Share		0.0781**	0.0377	0.0600	0.0719*			
		(0.0395)	(0.0462)	(0.0383)	(0.0385)			
SPLIT - Large Share		0.0890***	0.1273***	0.1117***	0.0945**			
		(0.0328)	(0.0274)	(0.0382)	(0.0434)			
SPLIT - Between Colonial Powers							0.0885**	
							(0.0378)	
SPLIT - Within Colonial Power							0.0732**	
							(0.0361)	
SPLIT - Two-Way Splits								0.0623
								(0.0403)
SPLIT - Multiple-Way Splits								0.1163***
								(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 (NB) 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 in the same country. All specifications include a vector of country fixed effects (constants not reported). All specifications 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, 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

	Ethnic Fractionalization	Linguistic Fractionalization	Religious Fractionalization	Landlocked - Coastal	Large - Small Land Area
	(1)	(2)	(3)	(4)	(5)
SPLIT - High	0.5556***	0.6128***	0.4963***		
	(0.1713)	(0.1482)	(0.1561)		
SPLIT - Low	0.3379**	0.1032	0.3636**		
	(0.1555)	(0.1757)	(0.1742)		
SPLIT - Landlocked				0.8424***	
				(0.2017)	
SPLIT - Coastal				0.2250*	
				(0.1247)	
SPLIT - Big Countries					0.6110***
					(0.1362)
SPLIT - Small Countries					0.1022
					(0.1983)
SPIL (Adjacent Split)	0.4676**	0.4546**	0.4937**	0.4654**	0.5019**
	(0.2186)	(0.2144)	(0.2192)	(0.2231)	(0.2157)
Log Likelihood	-4108.04	-4105.48	-4108.53	-4103.57	-4105.90
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

	Ethnic Fractionalization	Linguistic Fractionalization	Religious Fractionalization	Landlocked - Coastal	Large - Small Land Area
	(1)	(2)	(3)	(4)	(5)
SPLIT - High	0.0605*	0.0774**	0.0755**		
	(0.0326)	(0.0318)	(0.0330)		
SPLIT - Low	0.1086**	0.0916*	0.0986***		
	(0.0422)	(0.0485)	(0.0377)		
SPLIT - Landlocked				0.1512***	
				(0.0345)	
SPLIT - Coastal				0.0479	
				(0.0336)	
SPLIT - Big Countries					0.0794**
					(0.032))
SPLIT - Small Countries					0.0870*
					(0.046))
SPIL (Adjacent Split)	0.1507***	0.1457***	0.1458***	0.1398***	0.1442***
	(0.0405)	(0.0412)	(0.0407)	(0.0416)	(0.0409)
adjusted R-square	0.446	0.445	0.445	0.447	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 (NB) 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 ethnic (in column (1)), linguistic (in column (2)), and religious (in column (3)) fractionalization index of Alesina et al. (2003). 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 in the same country. All specifications include a vector of country fixed effects (constants not reported). All specifications 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, 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. Panel A reports in parentheses clustered at the ethno-linguistic family dimension standard errors. 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 33. EPR Descriptive Patterns
Ethnic Partitioning and Political Violence

Panel A: Ethnic Wars, Political Discrimination and Ethnic Partitioning

	Political Violence				
	Ethnic Civil Wars		Political Discrimination		Total
Ethnic Groups	No	Yes	No	Yes	
Non-Partitioned	290 (80.78%)	69 (19.22%)	307 (85.52%)	52 (14.48%)	359
Partitioned	162 (69.23%)	72 (30.77%)	176 (75.21%)	58 (24.79%)	234
Total	452	141	483	110	593

Panel B: Ordered Political Violence and Ethnic Partitioning

Political Violence Status	Partitioned		Non-Partitioned		
	Number	Share	Number	Share	Total
Peace	140	59.83%	266	74.09%	406
Repression Only	22	9.40%	24	6.69%	46
Ethnic Civil Wars	72	30.77%	69	19.22%	141
Total	234	359	359		593

Panel A tabulates EPR [Ethnic Power Relations] database classification of ethnic-based civil wars and political discrimination for partitioned and non-split ethnicities. Panel B tabulates the ordered index of political violence for partitioned and non-split ethnic groups. The construction of the ordered index of political violence follows Besley and Persson (2011). The political violence takes three values. Peace (index value 0) when the ethnic group is neither discriminated from the national government nor it is involved in ethnic civil war. Repression (index value 1) when the ethnic group is subject to political discrimination from the government, but the ethnicity is not engaged in an ethnic civil war. Civil War (index value 2) when the ethnic group is engaged in a major civil war (two-sided conflict). For details see Section 6 in the main body for the paper. The EPR database (Wimmer, Cederman, and Min (2009)) covers 40 African countries during the post-independence period.

Appendix Table 34A:
Ethnic Partitioning, Exclusion from Central Government and Ethnic War Incidence and Onset

	Ethnic War Incidence		Ethnic War Onset					
	Cross-Sectional Variation		Time Series Variation					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT (Partitioning)	0.1702***	0.0115	0.0056***	0.0018**	0.0003	0.0054***	0.0019**	0.0005
	(0.0558)	(0.0289)	(0.0020)	(0.0008)	(0.0011)	(0.0019)	(0.0009)	(0.0012)
Excluded from National Power					0.0093***			0.0085***
					(0.0031)			(0.0030)
Excluded **x** SPLIT					0.0053**			0.0049**
					(0.0023)			(0.0024)
Excluded from National Power	Yes	No	Yes	No	.	Yes	No	.
			Any Time in the Past 3 Years			Any Time in the Past 5 Years		
Adjusted R-square	0.49	0.75	0.042	0.041	0.026	0.04	0.043	0.025
Observations	331	262	10186	16641	26827	10492	16335	26827
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	.	.	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports OLS estimates. Columns (1) and (2) exploit cross-sectional variation and the dependent variable is an indicator that takes the value of 1 if the ethnic group has engaged in an ethnic war between 1960-2010. Data on ethnic wars and exclusion from the central 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. The specifications in columns (3)-(8) explore time series variation and the dependent varriable takes tha value of 1 in years that an ethnic conflict erupts. In columns (3)-(5) we classify as excluded those groups that in the last 3 years before the current period have been excluded from the central government for at least one year. We use the 5-year threshold for columns (6)-(8). All columns include a vector of country fixed effects (constants not reported) and a rich set of controls including log of land area, the log of population in 1960, an indicator for lakes and 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 and an indicator for country-ethnic areas that are by the sea coast (location controls); and 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 (geographic controls). Columns (3)-(8) also include year fixed effects. The table reports in parentheses standard errors clustered at the group level for columns (3)-(8) and double clustered standard errors at the country and the ethno-linguistic family dimensions in columns (1) and (2). ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 34B :
Ethnic Partitioning, Exclusion from Central Government and Ethnic War Incidence and Onset

	Ethnic War Incidence		Ethnic War Onset					
	Cross-Sectional Variation		Time Series Variation					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Cross-Sectional		Time Series Variation					
ALT - SPLIT (Partitioning)	0.1865***	0.0167	0.0062***	0.0014**	0.001	0.0060***	0.0015**	0.0011
	(0.0569)	(0.0264)	(0.0020)	(0.0006)	(0.0010)	(0.0020)	(0.0007)	(0.0010)
Excluded from National Power					0.0092***			0.0084***
					(0.0033)			(0.0032)
Excluded **x** ALT - SPLIT					0.0050**			0.0046*
					(0.0024)			(0.0024)
Excluded from National Power	Yes	No	Yes	No	.	Yes	No	.
			Any Time in the Past 3 Years			Any Time in the Past 5 Years		
Adjusted R-square	0.49	0.75	0.042	0.041	0.026	0.04	0.043	0.025
Observations	331	262	10186	16641	26827	10492	16335	26827
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	.	.	Yes	Yes	Yes	Yes	Yes	Yes
All Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports OLS estimates. Columns (1) and (2) exploit cross-sectional variation and the dependent variable is an indicator that takes the value of 1 if the ethnic group has engaged in an ethnic war between 1960-2010. Data on ethnic wars and exclusion from the central government come from the Ethnic Power Relations (EPR) database (Wimmer, Cederman and Min (2009)).
SPLIT-ALT is an indicator variable that identifies partitioned ethnicities as those with at least 5% of the historical homeland falling into more than one contemporary country. The specifications in columns (3)-(8) explore time series variation and the dependent varriable takes tha value of 1 in years that an ethnic conflict erupts. In columns (3)-(5) we classify as excluded those groups that in the last 3 years before the current period have been excluded from the central government for at least one year. We use the 5-year threshold for columns (6)-(8). All columns include a vector of country fixed effects (constants not reported) and a rich set of controls including log of land area, the log of population in 1960, an indicator for lakes and 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 and an indicator for country-ethnic areas that are by the sea coast (location controls); and 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 (geographic controls). Columns (3)-(8) also include year fixed effects. The table reports in parentheses standard errors clustered at the group level for columns (3)-(8) and double clustered standard errors at the country and the ethno-linguistic family dimensions in columns (1) and (2). ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 35: Ethnic Partitioning and Political Violence
Sensitivity Analysis. Alternative Estimation Techniques

	OLS				Ordered Probit ML			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT (Partitioning)	0.2581**	0.2620***	0.1914**	0.1976**	0.3863***	0.5180***	0.3809**	0.3633**
	(0.1005)	(0.0716)	(0.0793)	(0.0796)	(0.1502)	(0.1347)	(0.1410)	(0.1309)
SPIL (Adjacent Split)				0.4877**				0.6805
				(0.2312)				(0.3735)
Adjusted R-square	0.024	0.421	0.469	0.480	-467.367	-432.564	-411.108	-407.905
Observations	593	593	593	593	593	593	593	593
Countries	40	40	40	40	40	40	40	40
Country Fixed Effects	No	Yes	Yes	Yes	No	No	No	No
Regional Fixed Effects	No	No	No	No	No	Yes	Yes	Yes
Simple Controls	No	No	Yes	Yes	No	No	Yes	Yes
Location Controls	No	No	Yes	Yes	No	No	Yes	Yes
Geographic Controls	No	No	Yes	Yes	No	No	Yes	Yes

The table reports OLS estimates (in columns (1)-(4)) and ordered probit ML estimates (in columns (5)-(8)), associating an ethnic-based ordered index of political violence with ethnic partitioning. The dependent variable index of political violence equals two if the ethnic group is engaged in a major civil war (two-sided conflict); the index equals one when the group is subject to political discrimination from the national government but not in civil war (one-sided violence); the index equals zero when the ethnicity is neither discriminated from the national government nor involved in civil war (the construction of the ordered index of political violence follows Besley and Persson (2011). 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 in the same country. The specifications in columns (2)-(4) include a vector of country fixed effects (constants not reported). The specifications in columns (6)-(8) include a vector of region fixed effects (constants not reported). The specifications in columns (3)-(4) and (7)-(8) include log of land area, the log of population in 1960, an indicator for lakes and 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 and an indicator for country-ethnic areas that are by the sea coast (location controls); and 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 (geographic controls). 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: Ethnic Partitioning, Ethnic-based Political Discrimination, and Major Ethnic Civil Wars. Sensitivity Analysis. Alternative Index of Ethnic Partitioning. Linear Probability Model Estimates

	Ethnic War				Ethnic Discrimination			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT-5PC (Partitioning)	0.1057**	0.1451***	0.1176***	0.1180***	0.1244***	0.0933***	0.0864***	0.0864***
	(0.0472)	(0.0392)	(0.0402)	(0.0388)	(0.0311)	(0.0266)	(0.0279)	(0.0282)
SPIL (Adjacent Split)				0.2518**				0.0203
				(0.1140)				(0.0759)
Adjusted R-square	0.017	0.425	0.476	0.488	0.027	0.479	0.498	0.523
Observations	593	593	593	593	593	593	593	593
Countries	40	40	40	40	40	40	40	40
Country Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Simple Controls	No	No	Yes	Yes	No	No	Yes	Yes
Location Controls	No	No	Yes	Yes	No	No	Yes	Yes
Geographic Controls	No	No	No	Yes	No	No	No	Yes

The table reports linear probability model estimates, associating ethnic-based civil wars and political discrimination from the national government 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 a major or minor civil war with an explicit ethnic dimension over the period 1960-2010. The dependent variable in columns (5)-(8) 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)).
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 in the same country. The specifications in columns (2)-(4) and (6)-(8) include a vector of country fixed effects (constants not reported). The specifications in columns (3)-(4) and (7)-(12) include log of land area, the log of population in 1960, an indicator for lakes and 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 and an indicator for country-ethnic areas that are by the sea coast (location controls); and 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 (geographic controls). 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 37: 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	
	Ethnic War	Discrim.	Ethnic War	Discrim.	Ethnic War	Discrim.	Ethnic War	Discrim.	Ethnic War	Discrim.
	(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***
	(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
	(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 38: 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 and Employing an Alternative Index of Ethnic Partitioning.

	Excluding									
	North		East		West		Central		South	
	Ethnic War	Discrim.	Ethnic War	Discrim.	Ethnic War	Discrim.	Ethnic War	Discrim.	Ethnic War	Discrim.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SPLIT-5PC (Partitioning	0.1355***	0.0958***	0.1147***	0.0765**	0.1182**	0.0779**	0.0752**	0.0771**	0.1329***	0.0925***
	(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
	(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 39: DHS Descriptive Patterns and Summary Statistics
Ethnic Partitioning at the Identity Level and at the Location Level

Panel A: Descriptives

Ethnic Partitioning		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	88043	5.460	4.699	6	0	24
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 40: The Long-Run Effects of Ethnic Partitioning.
Enumeration-area (Village) Fixed Effects Estimates

	All Observations (Individuals)				Observations close to the Border			
	Composite Wealth Index		Education		Composite Wealth Index		Education	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partitioning - Identity	-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 Indicator	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-squared	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
Enumeration-Area FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Location Controls	No	Yes	No	Yes	No	Yes	No	Yes

The table reports OLS estimates associating a composite wealth index (in columns (1)-(2) and (5)-(6)) and years of education (in columns (3)-(4) and (7)-(8)) with ethnic partitioning. The ethnic partitioning index (Partitioning-Identity) takes on the value of one for individuals that identify with a partitioned ethnicity and zero otherwise. 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).

All specifications in include a vector of enumeration area (village/town/city) fixed effects (constants not reported). The set of individual controls includes a vector of year-of-birth fixed effects, a vector of 6 marital-status fixed effects, and a vector of 7 religion fixed effects. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses standard errors clustered at the respondent's ethnicity. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 41: The Long-Run Effects of Ethnic Partitioning on Individual Well-Being and Education. DHS Data Channels; Location and Identity. Looking at "Movers" & "Non-Movers"

	Composite Wealth Index						Education					
	Non-Movers		Movers in Non-Partitioned Homelands		Movers in Partitioned Homelands		Non-Movers		Movers in Non-Partitioned Homelands		Movers in Partitioned Homelands	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Partitioning - Identity	-0.2070**	-0.1843**	-0.2932*	-0.2290	-0.2109**	-0.1984**	-0.5686*	-0.4687*	-0.9172**	-0.5913*	-0.6031**	-0.5113**
Double-clustered s.e.	(0.0936)	(0.0909)	(0.1615)	(0.1479)	(0.0894)	(0.0903)	(0.3309)	(0.2798)	(0.4493)	(0.3438)	(0.2650)	(0.2536)
Adjusted R-squared	0.117	0.140	0.210	0.246	0.147	0.177	0.246	0.293	0.256	0.328	0.208	0.255
Observations	34179	34179	30606	30606	23386	23386	34179	34179	30606	30606	23386	23386
Country Fixed Effects	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
Individual 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 years of education (in columns (7)-(12)) with ethnic partitioning. The ethnic partitioning identity index (Partitioned-Identity) takes on the value of one for individuals that identify with a partitioned group. Columns (1), (2) and (7) and (8) focuse on individuals that reside *inside* their ethnicity's ancestral homeland ("non-movers"). In the rest of the columns we focus on those 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 specifications in columns (3)-(4) and (9)-(10) restrict estimation to individuals (movers) currently residing in non-partitioned ethnic homelands. The specifications in columns (5)-(6) and (11)-(12) restrict estimation to individuals (movers) currently residing in partitioned ethnic homelands.

All specifications include a vector of country fixed effects (constants not reported) and a set of location controls. This includes the distance of each enumeration area to the capital city, the distance to the coast, the distance to the national border and an indicator that takes on the value of one for enumeration areas close to the capital city (distance to the capital less than the 25th percentile). The individual controls is added in even-numbered columns and it includes a vector of year-of-birth fixed effects, a vector of 6 marital status fixed effects, and a vector of 7 religion fixed effects. 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 42: The Long-Run Effects of Ethnic Partitioning
Examining Persistence

	DHS Composite Wealth Index				Education			
	Old	Young	Old	Young	Old	Young	Old	Young
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partitioning - Identity	-0.2290**	-0.1882**	-0.2468**	-0.1856**	-0.5861**	-0.4047*	-0.6729**	-0.4354*
Double-clustered s.e.	(0.0968)	(0.0865)	(0.0986)	(0.0902)	(0.2782)	(0.2336)	(0.2900)	(0.2437)
Partitioning - Location			0.0683	-0.0112			0.3331	0.1343
Double-clustered s.e.			(0.0735)	(0.0789)			(0.2058)	(0.1803)
Non-Indigenous Indicator	0.1910**	0.1785**	0.1940***	0.1783**	0.3626	0.255	0.3774	0.257
Double-clustered s.e.	(0.0751)	(0.0700)	(0.0741)	(0.0702)	(0.2825)	(0.1950)	(0.2792)	(0.1934)
Adjusted R-squared	0.172	0.174	0.172	0.174	0.283	0.266	0.283	0.266
Observations	43283	44888	43283	44888	43231	44812	43231	44812
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports OLS estimates associating a composite wealth index in columns (1) - (4) and years of education in columns (5) - (8) with ethnic partitioning. The ethnic partitioning identity index (Partitioning-Identity) takes on the value of one for individuals that identify with a partitioned group. The location-based ethnic partitioning index (Partitioning-Location) takes on the value of one for individuals that currently reside in ethnic homelands that have been partitioned by the national border and zero otherwise. Even-numbered columns report estimates restricting estimation to individuals born after 1977 (young). Odd-numbered columns report estimates restricting estimation to individuals born before (or on) 1977 (old). The composite wealth index (dependent variable) 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).
All specifications include a vector of country ethnic homeland fixed effects (constants not reported). All specifications include a vector of year-of-birth fixed effects, a vector of 6 marital-status fixed effects, and a vector of 7 religion fixed effects (individual controls). We also include as location controls the distance of each enumeration area to the capital city, the distance to the coast, the distance to the national border and an indicator that takes on the value for enumeration areas close to the capital city (distance to the capital less than the 25th percentile). 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

No	Ethnicity Name	% of Homeland	Country Code	# Partitions	No	Ethnicity Name	% of Homeland	Country Code	# Partitions
1	ABABDA	0.72	EGY	2	120	LAKA (ADAMAWA)	0.69	TCD	3
	ABABDA	0.28	SDN	2		LAKA (ADAMAWA)	0.20	CMR	3
2	ADELE	0.48	GHA	2		LAKA (ADAMAWA)	0.11	CAF	3
	ADELE	0.52	TGO	2	121	LAMBA	0.39	ZAR	2
3	AFAR	0.17	DJI	3		LAMBA	0.61	ZMB	2
	AFAR	0.22	ERI	3	122	LAMBYA	0.17	MWI	3
	AFAR	0.61	ETH	3		LAMBYA	0.33	TZA	3
4	ALUR	0.16	ZAR	2		LAMBYA	0.50	ZMB	3
	ALUR	0.84	UGA	2	123	LIGBI, DEGHA (SE)	0.72	GHA	2
5	AMBA	0.87	ZAR	2		LIGBI, DEGHA (SE)	0.28	CIV	2
	AMBA	0.13	UGA	2	124	LOBI	0.42	CIV	2
6	AMBO	0.41	AGO	2		LOBI	0.58	BFA	2
	AMBO	0.59	NAM	2	125	LUGBARA	0.45	ZAR	3
7	AMER	0.56	ERI	2		LUGBARA	0.04	SDN	3
	AMER	0.44	SDN	2		LUGBARA	0.51	UGA	3
8	ANA	0.33	BEN	2	126	LUNGU	0.31	TZA	2
	ANA	0.67	TGO	2		LUNGU	0.69	ZMB	2
9	ANUAK	0.75	ETH	2	127	LUVALE	0.81	AGO	3
	ANUAK	0.25	SDN	2		LUVALE	0.01	ZAR	3
10	ANYI	0.42	GHA	2		LUVALE	0.17	ZMB	3
	ANYI	0.58	CIV	2	128	MADI	0.42	SDN	2
11	ASBEN	0.89	NER	2		MADI	0.58	UGA	2
	ASBEN	0.11	DZA	2	129	MAKONDE	0.56	MOZ	2
12	ASSINI	0.51	GHA	2		MAKONDE	0.44	TZA	2
	ASSINI	0.49	CIV	2	130	MALINKE	0.03	GMB	6
13	ATTA	0.51	MAR	2		MALINKE	0.13	CIV	6
	ATTA	0.49	DZA	2		MALINKE	0.27	MLI	6
14	ATYUTI	0.13	GHA	2		MALINKE	0.04	GNB	6
	ATYUTI	0.87	TGO	2		MALINKE	0.25	GIN	6
15	AULLIMINDEN	0.55	MLI	3		MALINKE	0.29	SEN	6
	AULLIMINDEN	0.40	NER	3	131	MAMBILA	0.57	CMR	2
	AULLIMINDEN	0.05	DZA	3		MAMBILA	0.43	NGA	2
16	AUSHI	0.27	ZAR	2	132	MANDARA	0.35	CMR	2
	AUSHI	0.73	ZMB	2		MANDARA	0.65	NGA	2
17	AVATIME	0.51	GHA	2	133	MANGA	0.60	NER	2
	AVATIME	0.49	TGO	2		MANGA	0.40	NGA	2
18	AZANDE	0.62	ZAR	3	134	MANYIKA	0.39	MOZ	2
	AZANDE	0.15	CAF	3		MANYIKA	0.61	ZWE	2
	AZANDE	0.23	SDN	3	135	MASAI	0.38	KEN	2
19	AZJER	0.24	LBY	3		MASAI	0.62	TZA	2
	AZJER	0.00	NER	3	136	MASALIT	0.13	TCD	2
	AZJER	0.75	DZA	3		MASALIT	0.87	SDN	2

20	BABUKUR	0.82	ZAR	2
	BABUKUR	0.18	SDN	2
21	BAJUN	0.37	KEN	2
	BAJUN	0.63	SOM	2
22	BALANTE	0.73	GNB	2
	BALANTE	0.27	SEN	2
23	BANYUN	0.48	GNB	2
	BANYUN	0.52	SEN	2
24	BANZIRI	0.14	ZAR	2
	BANZIRI	0.86	CAF	2
25	BARABRA	0.31	EGY	2
	BARABRA	0.69	SDN	2
26	BARARETTA	0.18	ETH	3
	BARARETTA	0.44	KEN	3
	BARARETTA	0.38	SOM	3
27	BARGU	0.77	BEN	4
	BARGU	0.03	NER	4
	BARGU	0.19	NGA	4
	BARGU	0.02	BFA	4
28	BASHI	0.09	BDI	3
	BASHI	0.83	ZAR	3
	BASHI	0.08	RWA	3
29	BATA	0.29	CMR	2
	BATA	0.71	NGA	2
30	BAYA	0.20	CMR	2
	BAYA	0.80	CAF	2
31	BERABISH	0.80	MLI	2
	BERABISH	0.20	MRT	2
32	BERTA	0.75	ETH	2
	BERTA	0.25	SDN	2
33	BIDEYAT	0.21	LBY	4
	BIDEYAT	0.40	TCD	4
	BIDEYAT	0.03	EGY	4
	BIDEYAT	0.36	SDN	4
34	BIRIFON	0.52	GHA	3
	BIRIFON	0.47	BFA	3
35	BOBO	0.20	MLI	2
	BOBO	0.80	BFA	2
36	BOKI	0.22	CMR	2
	BOKI	0.78	NGA	2
37	BONDJO	0.14	ZAR	2
	BONDJO	0.86	COG	2
38	BONI	0.67	KEN	2
	BONI	0.33	SOM	2
39	BORAN	0.46	ETH	2
	BORAN	0.54	KEN	2
40	BRONG	0.84	GHA	2
	BRONG	0.16	CIV	2
41	BUEM	0.40	GHA	2
137	MASHI	0.12	AGO	2
	MASHI	0.88	ZMB	2
138	MASINA	0.82	MLI	3
	MASINA	0.09	BFA	3
	MASINA	0.09	MRT	3
139	MATAKAM	0.70	CMR	2
	MATAKAM	0.30	NGA	2
140	MBERE	0.02	TCD	3
	MBERE	0.24	CMR	3
	MBERE	0.74	CAF	3
141	MBUKUSHU	0.74	AGO	3
	MBUKUSHU	0.15	BWA	3
	MBUKUSHU	0.12	NAM	3
142	MBUNDA	0.89	AGO	2
	MBUNDA	0.11	ZMB	2
143	MENDE	0.18	LBR	3
	MENDE	0.82	SLE	3
144	MINIANKA	0.01	CIV	3
	MINIANKA	0.72	MLI	3
	MINIANKA	0.27	BFA	3
145	MOMBERA	0.72	MWI	2
	MOMBERA	0.28	ZMB	2
146	MPEZENI	0.11	MWI	2
	MPEZENI	0.89	ZMB	2
147	MUNDANG	0.80	TCD	2
	MUNDANG	0.20	CMR	2
148	MUNDU	0.30	ZAR	2
	MUNDU	0.70	SDN	2
149	MUSGU	0.76	TCD	2
	MUSGU	0.24	CMR	2
150	NAFANA	0.74	GHA	2
	NAFANA	0.26	CIV	2
151	NALU	0.41	GNB	2
	NALU	0.59	GIN	2
152	NAMA	0.18	ZAF	2
	NAMA	0.82	NAM	2
153	NAUDEBA	0.87	BEN	2
	NAUDEBA	0.13	TGO	2
154	NDAU	0.86	MOZ	2
	NDAU	0.14	ZWE	2
155	NDEMBU	0.26	AGO	3
	NDEMBU	0.39	ZAR	3
	NDEMBU	0.35	ZMB	3
156	NDOGO	0.01	ZAR	3
	NDOGO	0.18	CAF	3
	NDOGO	0.81	SDN	3
157	NDUKA	0.23	TCD	2
	NDUKA	0.77	CAF	2
158	NGAMA	0.30	TCD	2

	BUEM	0.60	TGO	2
42	BULOM	0.85	SLE	2
	BULOM	0.15	GIN	2
43	BUSA	0.14	BEN	2
	BUSA	0.86	NGA	2
44	BWAKA	0.81	ZAR	3
	BWAKA	0.15	CAF	3
	BWAKA	0.04	COG	3
45	CHAGA	0.24	KEN	2
	CHAGA	0.76	TZA	2
46	CHAKOSSI	0.27	GHA	2
	CHAKOSSI	0.73	TGO	2
47	CHEWA	0.34	MWI	3
	CHEWA	0.50	MOZ	3
	CHEWA	0.16	ZMB	3
48	CHIGA	0.12	RWA	3
	CHIGA	0.87	UGA	3
49	CHOKWE	0.81	AGO	2
	CHOKWE	0.19	ZAR	2
50	COMORIANS	0.82	COM	2
	COMORIANS	0.18	MYT	2
51	DAGARI	0.67	GHA	2
	DAGARI	0.33	BFA	2
52	DARI	0.78	TCD	2
	DARI	0.22	CMR	2
53	DAZA	0.27	TCD	2
	DAZA	0.73	NER	2
54	DELIM	0.55	ESH	2
	DELIM	0.45	MRT	2
55	DENDI	0.60	BEN	3
	DENDI	0.39	NER	3
56	DIALONKE	0.36	MLI	3
	DIALONKE	0.58	GIN	3
	DIALONKE	0.06	SEN	3
57	DIDINGA	0.04	KEN	3
	DIDINGA	0.89	SDN	3
	DIDINGA	0.07	UGA	3
58	DIGO	0.62	KEN	2
	DIGO	0.38	TZA	2
59	DIOLA	0.14	GMB	3
	DIOLA	0.07	GNB	3
	DIOLA	0.78	SEN	3
60	DUMA	0.63	GAB	2
	DUMA	0.37	COG	2
61	DZEM	0.74	CMR	3
	DZEM	0.03	GAB	3
	DZEM	0.24	COG	3
62	EGBA	0.41	BEN	3
	EGBA	0.52	NGA	3
	NGAMA	0.70	CAF	2
159	NGERE	0.65	CIV	3
	NGERE	0.29	LBR	3
	NGERE	0.06	GIN	3
160	NGUMBA	0.65	CMR	2
	NGUMBA	0.35	GNQ	2
161	NGWAKETSE	0.86	BWA	2
	NGWAKETSE	0.14	ZAF	2
162	NSENGA	0.15	MOZ	3
	NSENGA	0.78	ZMB	3
	NSENGA	0.06	ZWE	3
163	NSUNGLI	0.78	CMR	2
	NSUNGLI	0.22	NGA	2
164	NUKWE	0.44	AGO	4
	NUKWE	0.24	BWA	4
	NUKWE	0.05	ZMB	4
	NUKWE	0.26	NAM	4
165	NUSAN	0.30	BWA	3
	NUSAN	0.37	ZAF	3
	NUSAN	0.33	NAM	3
166	NYAKYUSA	0.12	MWI	2
	NYAKYUSA	0.88	TZA	2
167	NYANGIYA	0.17	SDN	2
	NYANGIYA	0.83	UGA	2
168	NYANJA	0.64	MWI	2
	NYANJA	0.36	MOZ	2
169	NYASA	0.05	MWI	3
	NYASA	0.68	MOZ	3
	NYASA	0.27	TZA	3
170	NZANKARA	0.14	ZAR	2
	NZANKARA	0.86	CAF	2
171	PANDE	0.38	CAF	2
	PANDE	0.62	COG	2
172	POPO	0.72	BEN	2
	POPO	0.28	TGO	2
173	PUKU	0.31	CMR	3
	PUKU	0.49	GNQ	3
	PUKU	0.19	GAB	3
174	REGEIBAT	0.34	ESH	2
	REGEIBAT	0.66	MRT	2
175	RESHIAT	0.83	ETH	3
	RESHIAT	0.06	KEN	3
	RESHIAT	0.11	SDN	3
176	RONGA	0.60	MOZ	3
	RONGA	0.35	ZAF	3
	RONGA	0.05	SWZ	3
177	RUANDA	0.02	BDI	5
	RUANDA	0.06	ZAR	5
	RUANDA	0.89	RWA	5

	EGBA	0.07	TGO	3
63	EKOI	0.38	CMR	2
	EKOI	0.62	NGA	2
64	ESA	0.03	DJI	3
	ESA	0.52	ETH	3
	ESA	0.44	SOM	3
65	EWE	0.44	GHA	2
	EWE	0.56	TGO	2
66	FANG	0.37	CMR	4
	FANG	0.07	GNQ	4
	FANG	0.54	GAB	4
	FANG	0.02	COG	4
67	FON	0.86	BEN	3
	FON	0.14	TGO	3
68	FOUTADJALON	0.01	MLI	4
	FOUTADJALON	0.11	GNB	4
	FOUTADJALON	0.88	GIN	4
	FOUTADJALON	0.01	SEN	4
69	FUNGON	0.81	CMR	2
	FUNGON	0.19	NGA	2
70	GADAMES	0.25	LBY	3
	GADAMES	0.27	TUN	3
	GADAMES	0.48	DZA	3
71	GIL	0.80	MAR	2
	GIL	0.20	DZA	2
72	GOMANI	0.86	MWI	2
	GOMANI	0.14	MOZ	2
73	GREBO	0.33	CIV	2
	GREBO	0.67	LBR	2
74	GRUNSHI	0.68	GHA	2
	GRUNSHI	0.32	BFA	2
75	GUDE	0.83	CMR	2
	GUDE	0.17	NGA	2
76	GULA	0.61	TCD	2
	GULA	0.39	CAF	2
77	GUN	0.48	BEN	2
	GUN	0.52	NGA	2
78	GURENSI	0.74	GHA	3
	GURENSI	0.13	TGO	3
	GURENSI	0.13	BFA	3
79	GURMA	0.15	BEN	4
	GURMA	0.12	NER	4
	GURMA	0.01	TGO	4
	GURMA	0.72	BFA	4
80	GUSII	0.53	KEN	2
	GUSII	0.47	TZA	2
81	HAMAMA	0.80	TUN	2
	HAMAMA	0.20	DZA	2
82	HAUSA	0.14	NER	2
	RUANDA	0.02	TZA	5
	RUANDA	0.02	UGA	5
178	RUNDI	0.76	BDI	4
	RUNDI	0.04	RWA	4
	RUNDI	0.20	TZA	4
179	RUNGA	0.74	TCD	3
	RUNGA	0.26	CAF	3
180	SABEI	0.56	KEN	2
	SABEI	0.44	UGA	2
181	SAHO	0.43	ERI	2
	SAHO	0.57	ETH	2
182	SAMO	0.12	MLI	2
	SAMO	0.88	BFA	2
183	SANGA	0.26	CMR	3
	SANGA	0.19	CAF	3
	SANGA	0.55	COG	3
184	SEKE	0.34	GNQ	2
	SEKE	0.66	GAB	2
185	SHAMBALA	0.10	KEN	2
	SHAMBALA	0.90	TZA	2
186	SHEBELLE	0.58	ETH	2
	SHEBELLE	0.42	SOM	2
187	SHUWA	0.62	TCD	3
	SHUWA	0.17	CMR	3
	SHUWA	0.21	NGA	3
188	SONGHAI	0.57	MLI	3
	SONGHAI	0.36	NER	3
	SONGHAI	0.07	BFA	3
189	SONINKE	0.68	MLI	3
	SONINKE	0.03	SEN	3
	SONINKE	0.29	MRT	3
190	SOTHO	0.24	LSO	2
	SOTHO	0.76	ZAF	2
191	SUBIA	0.11	BWA	4
	SUBIA	0.53	ZMB	4
	SUBIA	0.06	ZWE	4
	SUBIA	0.30	NAM	4
192	SUNDI	0.37	ZAR	2
	SUNDI	0.63	COG	2
193	SURI	0.71	ETH	2
	SURI	0.29	SDN	2
194	SWAZI	0.45	ZAF	2
	SWAZI	0.55	SWZ	2
195	TABWA	0.57	ZAR	2
	TABWA	0.43	ZMB	2
196	TAJAKANT	0.15	MAR	4
	TAJAKANT	0.14	ESH	4
	TAJAKANT	0.66	DZA	4
	TAJAKANT	0.05	MRT	4

	HAUSA	0.86	NGA	2
83	HIECHWARE	0.81	BWA	2
	HIECHWARE	0.19	ZWE	2
84	HLENGWE	0.82	MOZ	3
	HLENGWE	0.00	ZAF	3
	HLENGWE	0.18	ZWE	3
85	HOLO	0.84	AGO	2
	HOLO	0.16	ZAR	2
86	IBIBIO	0.11	CMR	2
	IBIBIO	0.89	NGA	2
87	IFORA	0.30	MLI	2
	IFORA	0.70	DZA	2
88	IMRAGEN	0.10	MAR	3
	IMRAGEN	0.74	ESH	3
	IMRAGEN	0.16	MRT	3
89	ISHAAK	0.20	ETH	2
	ISHAAK	0.80	SOM	2
90	IWA	0.33	TZA	2
	IWA	0.67	ZMB	2
91	JERID	0.90	TUN	2
	JERID	0.10	DZA	2
92	JIE	0.24	KEN	2
	JIE	0.76	UGA	2
93	KABRE	0.39	BEN	2
	KABRE	0.61	TGO	2
94	KANEMBU	0.73	TCD	3
	KANEMBU	0.25	NER	3
	KANEMBU	0.02	NGA	3
95	KAONDE	0.21	ZAR	2
	KAONDE	0.79	ZMB	2
96	KAPSIKI	0.65	CMR	2
	KAPSIKI	0.35	NGA	2
97	KARA	0.85	CAF	2
	KARA	0.15	SDN	2
98	KARAMOJONG	0.27	KEN	2
	KARAMOJONG	0.73	UGA	2
99	KARE	0.75	ZAR	2
	KARE	0.25	CAF	2
100	KGATLA	0.13	BWA	2
	KGATLA	0.87	ZAF	2
101	KISSI	0.12	LBR	3
	KISSI	0.02	SLE	3
	KISSI	0.86	GIN	3
102	KOBA	0.89	BWA	2
	KOBA	0.11	NAM	2
103	KOMA	0.57	ETH	2
	KOMA	0.43	SDN	2
104	KOMONO	0.49	CIV	2
	KOMONO	0.51	BFA	2
197	TAMA	0.30	TCD	2
	TAMA	0.70	SDN	2
198	TAWARA	0.57	MOZ	2
	TAWARA	0.43	ZWE	2
199	TEDA	0.34	LBY	3
	TEDA	0.35	TCD	3
	TEDA	0.31	NER	3
200	TEKE	0.31	ZAR	3
	TEKE	0.03	GAB	3
	TEKE	0.66	COG	3
201	TEKNA	0.53	MAR	2
	TEKNA	0.47	ESH	2
202	TEM	0.17	BEN	2
	TEM	0.83	TGO	2
203	TENDA	0.57	GIN	2
	TENDA	0.43	SEN	2
204	THONGA	0.58	MOZ	3
	THONGA	0.42	ZAF	3
205	TIENGA	0.22	NER	3
	TIENGA	0.78	NGA	3
206	TIGON	0.32	CMR	2
	TIGON	0.68	NGA	2
207	TIGRINYA	0.51	ERI	3
	TIGRINYA	0.44	ETH	3
	TIGRINYA	0.05	SDN	3
207	TLOKWA	0.14	BWA	3
	TLOKWA	0.77	ZAF	3
	TLOKWA	0.09	ZWE	3
208	TOMA	0.29	LBR	2
	TOMA	0.71	GIN	2
209	TONGA	0.84	ZMB	2
	TONGA	0.16	ZWE	2
210	TRIBU	0.25	GHA	2
	TRIBU	0.75	TGO	2
211	TRIPOLITANIANS	0.74	LBY	2
	TRIPOLITANIANS	0.26	TUN	2
212	TUBURI	0.25	TCD	2
	TUBURI	0.75	CMR	2
213	TUKULOR	0.39	SEN	2
	TUKULOR	0.61	MRT	2
214	TUMBUKA	0.74	MWI	2
	TUMBUKA	0.26	ZMB	2
215	TUNISIANS	0.87	TUN	2
	TUNISIANS	0.13	DZA	2
216	UDALAN	0.82	MLI	3
	UDALAN	0.05	NER	3
	UDALAN	0.13	BFA	3
217	VAI	0.76	LBR	2
	VAI	0.24	SLE	2

105	KONGO	0.77	AGO	3	218	VENDA	0.70	ZAF	2
	KONGO	0.23	ZAR	3		VENDA	0.30	ZWE	2
106	KONJO	0.81	ZAR	2	219	VILI	0.20	AGO	4
	KONJO	0.19	UGA	2		VILI	0.22	ZAR	4
107	KONKOMBA	0.24	GHA	2		VILI	0.11	GAB	4
	KONKOMBA	0.76	TGO	2		VILI	0.47	COG	4
108	KONO	0.74	SLE	2	220	WAKURA	0.28	CMR	2
	KONO	0.26	GIN	2		WAKURA	0.72	NGA	2
109	KONYANKE	0.30	CIV	2	221	WANGA	0.79	KEN	2
	KONYANKE	0.70	GIN	2		WANGA	0.21	UGA	2
110	KORANKO	0.39	SLE	2	222	WUM	0.88	CMR	2
	KORANKO	0.61	GIN	2		WUM	0.12	NGA	2
111	KOTA	0.41	GAB	2	223	YAKA	0.16	AGO	2
	KOTA	0.59	COG	2		YAKA	0.84	ZAR	2
112	KOTOKO	0.67	TCD	2	224	YAKOMA	0.40	ZAR	2
	KOTOKO	0.33	CMR	2		YAKOMA	0.60	CAF	2
113	KPELLE	0.48	LBR	3	225	YALUNKA	0.25	SLE	2
	KPELLE	0.52	GIN	3		YALUNKA	0.75	GIN	2
114	KRAN	0.16	CIV	2	226	YAO	0.13	MWI	3
	KRAN	0.84	LBR	2		YAO	0.65	MOZ	3
115	KREISH	0.10	CAF	2		YAO	0.22	TZA	3
	KREISH	0.90	SDN	2	227	YOMBE	0.13	AGO	3
116	KUNDA	0.84	MOZ	3		YOMBE	0.48	ZAR	3
	KUNDA	0.15	ZMB	3		YOMBE	0.39	COG	3
117	KUNG	0.10	BWA	2	228	ZAGHAWA	0.14	TCD	2
	KUNG	0.90	NAM	2		ZAGHAWA	0.86	SDN	2
118	KUNTA	0.85	MLI	2	229	ZEKARA	0.83	MAR	2
	KUNTA	0.15	DZA	2		ZEKARA	0.17	DZA	2
119	KWANGARE	0.84	AGO	2	230	ZIMBA	0.16	MWI	2
	KWANGARE	0.16	NAM	2		ZIMBA	0.84	MOZ	2

Appendix Table A reports the name of partitioned ethnic groups (as coded by Murdock (1959)) and the percentage of the historical homeland of the split ethnic groups that fall into more than one country. Section 2.1 gives details on our approach in identifying partitioned ethnicities. There are 230 partitioned ethnicities. In the empirical analysis we do not consider the Comorians, as the conflict databases do not cover the Comoros.