The Long-Run Effects of the Scramble for Africa*

Stelios Michalopoulos
Harvard University, Brown University, NBER and CEPR

Elias Papaioannou
London Business School, NBER and CEPR

March 24, 2015

Abstract

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

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

*We would like to thank the Editor and 4 referees for insightful comments and useful suggestions. We also thank Alberto Alesina, Maarten Bosker, Chris Blattman, Francesco Caselli, Jeremiah Dittmar, Joan Esteban, James Fenske, Fabrizio Dell' Acqua, Federico Finan, David Laitin, Horacio Larreguy Arbesú, Stathis Kalyvas, Sebastian Hohmann, Janina Matuszeski, Ted Miguel, Nathan Nunn, Gregorios Siourounis, Neils Weidman, Andreas Wimmer, and participants at the AEA meetings in Denver, Yale, UC Berkeley, Princeton, ALBA, UC Irvine, UC Merced, American University, Brown, the Institute for Economic Analysis, Autonoma University, George Washington University, CERGE-EI, Surrey, City University, the NBER Political Economy meetings, the CEPR meeting on the Political Economy of Conflict, the NBER Summer Institute Meetings on the Development of the American Economy and Income Distribution and Macroeconomics for useful comments and suggestions. All errors are our sole responsibility.

1 Introduction

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Our paper also contributes to the literature on the origins of civil conflict that mainly examines the role of country-level characteristics (see Collier and Hoeffler (2007) and Blattman and Miguel (2010) for reviews and Collier and Sambanis (2005) for case studies in Africa). Of most relevance are works that link a country's ethnic composition to civil war. While the correlation between ethnic fragmentation and civil war is weak (Fearon and Laitin (2003)), recent studies document interesting cross-country correlations associating various aspects of the societal structure with armed conflict. Montalvo and Reynal-Querol (2005) and Esteban, Mayoral, and Ray (2012) show a strong association between ethnic polarization and conflict, while Esteban and Ray (2008, 2011) emphasize the role of differences between and within ethnic groups. Huber and Mayoral (2014) show a link between within-group inequality and conflict. Wimmer, Cederman, and Min (2009) find that the likelihood of ethnic conflict increases when a large share of the population is excluded from power. We complement this research showing that ethnic partitioning affects both the likelihood of ethnic-based civil wars and political discrimination. We further show that the link between ethnic partitioning and civil conflict partially operates via discrimination.[4]

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

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

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

[5] 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 in the analysis and the overwhelming evidence of the African history on the arbitrariness of borders, our results suggest that the impact of unobservable factors are unlikely to be of first-order significance.

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

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

2 Historical Background

2.1 The Scramble for Africa

The "Scramble for Africa" starts in the 1860s when the French and the British begin the systematic exploration of Western Africa and sign bilateral agreements assigning spheres of influence. In the next 40 years, Europeans signed hundreds of treaties that divided the largely unexplored continent into protectorates, free-trade areas, and colonies. The event that stands for the partitioning of Africa is the conference that Otto von Bismarck organized in Berlin from November 1884 till February 1885. While the Berlin conference discussed only the boundaries of Central Africa (the Congo Free State), it came to symbolize the partitioning, because it laid down the principles that would be used among Europeans to divide the continent. The key consideration was to preserve the "status quo" preventing conflict among Europeans for Africa, as the memories of the European wars of the 18th-19th century were alive. As a result, in the overwhelming majority of cases, European powers drew borders without taking into account local conditions. African

leaders were not invited and had no say.[6] In many cases, European leaders were in such a rush that they didn't wait for the information arriving from explorers, geographers, and missionaries. As the British prime minister at the time Lord Salisbury put it, "*we have been engaged in drawing lines upon maps where no white man's feet have ever 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 accidental border design. First, at the time Europeans had little knowledge of local geographic conditions, as with the exception of few coastal areas, the continent was unexplored. There was a constant imperialist back and forth with European powers swapping pieces of land with limited (at best) idea of what they were worth of.[7] Second, Europeans were not drawing borders of prospective states, but of colonies and protectorates; clearly at the time none could foresee independence. Third, demarcation was poor.[8] Fourth, Europeans were unwilling to sacrifice their commitment not to go to war for any part of Africa and hence were reluctant to change colonial borders.[9] Fifth, as locals could freely move across colonial borders, African chiefs did not oppose much the colonial design, as little changed on the ground.[10] Wesseling (1996) summarizes: "*The partition of Africa was recorded by the Europeans on their maps, but the matter rested there for the time being....In Europe conquests preceded the drawing of maps; in Africa the map was drawn, and then it was decided what was going to happen. These maps did not therefore reflect reality but helped to create it.*"

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

[6] 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*".

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

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

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

[10] 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.*"

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

2.2 Channels and Case Studies

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

[11] 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.*"

[12] 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.

[13] 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.*"

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

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

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

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

[14] 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.).

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

3 Ethnic Partitioning and Border Artificiality

3.1 Identifying Partitioned Ethnic Groups

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

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

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

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

[17] 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.

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

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

groups. Appendix Table *A* lists partitioned ethnicities.

Our procedure identifies most major ethnic groups that have been split by African borders. For example, the Maasai are partitioned between Kenya and Tanzania (shares 62% and 38%), the Anyi between Ghana and the Ivory Coast (shares 58% and 42%), and the Chewa between Mozambique (50%), Malawi (34%), and Zimbabwe (16%). Other examples include the Hausa (split between Nigeria and Niger), the Ababda (split between Egypt and Sudan), the Ewe (split by the Togo-Ghana border), and the Bararetta clans (split between Kenya and Somalia). We also checked whether our coding is in line with Asiwaju (1985), who provides the only (to our knowledge) codification of partitioned ethnicities in Africa. Our strategy identifies almost all ethnic groups that Asiwaju (1985) lists as partitioned.[20]

Figure 1*a*

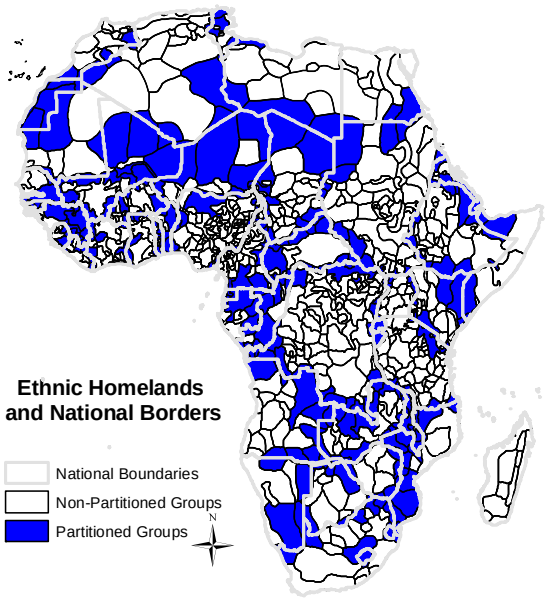

Figure 1*b*

3.2 Border Artificiality

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

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

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

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

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

[21] 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.

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

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

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

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

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

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

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

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

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

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

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

4 Ethnic Partitioning and Civil Conflict

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

4.1 Main Conflict Data

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

[23] 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).

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

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

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

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

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

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

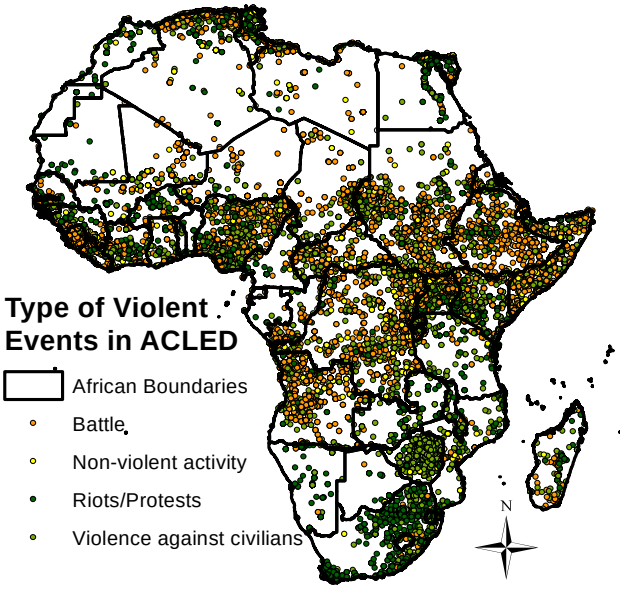

Figure 2*a*

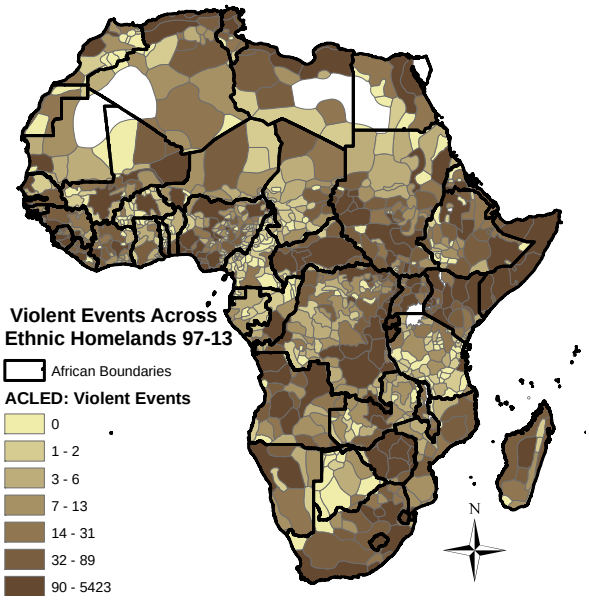

Figure 2*b*

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

4.2 Econometric Specification and Estimation

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

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

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

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

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

[25] 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.

[26] 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).

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

4.3 Ethnic Partitioning and Civil Conflict

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

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

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

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

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

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

4.4 Ethnic Partitioning and Conflict Intensity

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

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

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

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

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

Example[30] Senegal offers an illustration of our results. ACLED records 565 events across its 12 constituent ethnic homelands. In the isolated Casamance region in the South, where the Diola and the Banyun ancestral homelands are split by the colonial border, we observe 154 and 85 events, respectively.[31] This is 42.5% of all events (63% if we exclude the capital) though these two regions capture 11.2% of Senegal's area and only 6% of the country's population. Not only conflict is concentrated in Casamance but intensity is quite high. In these two split homelands, we observe 77 and 35 deadly events, out of a total of 182 deadly events (61.5% of all deadly events). The overwhelming majority of these events involve government troops (69 and 60 events) and/or rebels (114 in the homeland of the Diola and 63 in Banyun territory). And in both ethnic areas we observe conflict for 16 out of the 17 years between $1997 - 2013$, much longer than in all other ethnic regions (with the exception of the capital, the mean is 5). The duration of deadly events is 15 and 12 years for Diola and Banyun, while the mean for the other ten ethnic areas is 3.2 years. Conflict is also quite deadly. Out of a total of $1,210$ fatalities in Senegal, 540 (45%) and 354 (29%) have occurred in the two ethnic regions in Casamance.

4.5 Ethnic Partitioning and Type of Conflict

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

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

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

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

[31] 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 from the Portuguese side of the border (Tomas (2006)).

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

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

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

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

4.6 Sensitivity Checks

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

4.7 Heterogeneous Effects

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

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

5 Mechanisms and Channels

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

5.1 Hypotheses

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

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

5.2 Ethnic Partitioning and Conflict by Key Actors (ACLED)

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

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

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

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

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

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

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

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

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

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

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

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

5.3.1 Data

To shed further light on the link between ethnic partitioning and the type of conflict we used data from the Uppsala Conflict Data Program Geo-referenced Events Dataset (UCDP GED) that covers the period $1989-2010$ (Sundberg, Lindgren, and 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; this database identifies civil wars when conflict results in at least 25 fatalities in a given year. UCDP conflicts are grouped into three mutually exclusive categories.

(1) State-based armed conflict is defined as a *"contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths in one calendar year"*. Examples of state-based conflict where ethnic partitioning seems to have played a role include the fights between the Ethiopian government and rebels seeking self-determination in the Afar and the Ogaden region[35] and the fighting between Tuareg

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

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

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

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

5.3.2 Results

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

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

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

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

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

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

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

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

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

and the resulting insurgency in the region.]

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

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

6 Partitioning and Ethnic Power Relations

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

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

[38] 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 those

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

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

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

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

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

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

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

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

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

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

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

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

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

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

7 Ethnic Partitioning and Individual Well-Being

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

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

7.1 Data and Specification

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

Our empirical specification reads:

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

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

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

7.2 Baseline Estimates

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

[41] 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.

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

7.3 Ethnic Identity and Ethnic Homelands

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

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

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

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

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

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

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

8 Conclusion

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

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

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

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

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

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

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

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

[44] 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, The Netherlands.

Acemoglu, D., T. Reed, and J. A. Robinson (2014): "Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone," *Journal of Political Economy*, 122(2), 319–368.

Aghion, P., J. V. Reenen, and L. Zingales (2013): "Innovation and Institutional Ownership," *American Economic Review*, 103(1), 277–304.

Aker, J. C., M. W. Klein, S. A. O'Connell, and M. Yang (2010): "Borders, Ethnicity and Trade," Working Paper, NBER, No. 15960.

Alesina, A., W. Easterly, and J. Matuszeski (2011): "Artificial States," *Journal of the European Economic Association*, 9(2), 246–277.

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

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

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

Berman, N., M. Couttenier, D. Rohner, and M. Thoenig (2014): "This Mine is Mine! How Minerals Fuel Conflicts in Africa," working paper, Department of Economics, University of Lausanne.

Besley, T., and T. Persson (2011): *Pillars of Prosperity. The Political Economics of Development Clusters*. Princeton University Press, Princeton, NJ.

Besley, T., and M. Reynal-Querol (2014): "The Legacy of Historical Conflicts. Evidence from Africa," *American Political Science Review*, 108(2), 319–336.

Blattman, C., and E. Miguel (2010): "Civil War," *Journal of Economic Literature*, 48(1), 3–57.

Bloom, N., M. Schankerman, and J. V. Reenen (2013): "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 81(4), 1347–1393.

Bosker, M., and J. de Ree (2014): "Ethnicity and the Spread of Civil War," *Journal of Development Economics*, 108, 206–221.

Brecke, P. (1999): "Violent Conflicts 1400 A.D. to the Present in Different Regions of the World," Paper prepared for the 1999 Meeting of the Peace Science Society in Ann Arbor, Michigan.

Cameron, A. C., and P. K. Trivedi (2013): *Regression Analysis of Count Data Book*, vol. Econometric Society Monograph No. 53. Cambridge University Press, Cambridge, UK, 2nd edition edn.

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

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

Ciccone, A., and M. Jarocinski (2010): "Determinants of Economic Growth: Will Data Tell?," *American Economic Journal Macroeconomics*, 2(2), 222–246.

Cioffi-Revilla, C. (1996): "Origins and Evolution of War and Politics," *International Studies Quarterly*, 40(1), 1–22.

Collier, P. (2007): *The Bottom Billion*. Oxford 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, North Holland, Amsterdam and Oxford.

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

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

Djankov, S., and M. Reynal-Querol (2010): "The Causes of Civil Wars," mimeo UPF.

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

Easterly, W., and R. Levine (2012): "The European Origins of Economic Development," NBER Working Paper No. 18162.

Eck, K. (2012): "In Data We Trust? A Comparison of UCDP GED and ACLED Conflict Events," *Cooperation and Conflict*, 47(1), 124–141.

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

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

Esteban, J., and D. Ray (2008): "On the 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.

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

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

——— (2011): "Sons of the Soil, Migrants, and Civil War," *World Development*, 39(2), 199–211.

——— (2012): "How Persistent Is Armed Conflict?," mimeo, Stanford University.

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

Francois, P., I. Rainer, and F. Trebbi (2015): "How Is Power Shared In Africa?," *Econometrica*, Working Paper, NBER, No. 18425.

Fuchs-Schundeln, N., and T. A. Hassan (2015): "Natural Experiments in Macroeconomics," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and H. Uhlig, vol. 2.

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

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

Greene, W. H. (2005): *Econometric Analysis*. Prentice Hall, Upper Saddle River, NJ, fourth edition edn.

Guimaraes, P. (2008): "The Fixed Effects Negative Binomial Model Revisited," *Economic Letters*, 99(1), 63–66.

Harari, M., and E. La Ferrara (2014): "Conflict, Climate and Cells: A Disaggregated Analysis," mimeo MIT and Bocconi University.

Hausman, J., B. H. Hall, and Z. Griliches (1984): "Econometric Models for Count Data with an Application to the Patents-RD Relationship," *Econometrica*, 52(4), 909–938.

Hegre, H., and N. Sambanis (2006): "Sensitivity Analysis of Empirical Results on Civil War Onset," *Journal of Conflict Resolution*, 50(4), 508–535.

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

Horowitz, D. L. (1985): *Ethnic Groups in Conflict*. University of California Press, Berkeley and Los Angeles, California.

Huber, J. D., and L. Mayoral (2014): "Inequality, Ethnicity and Civil Conflict," Working Paper, Department of Political Science, Columbia University.

Huillery, E. (2009): "History Matters: The Long Term Impact of Colonial Public Investments in French West Africa," *American Economic Journal - Applied Economics*, 1(2), 176–215.

Humphreys, M., and H. A. Mohamed (2005): "Senegal and Mali: A Comparative Study of Rebellions in West Africa," in *Understanding Civil War: Evidence and Analysis Vol 1 Africa*, ed. by P. Collier, and N. Sambanis. The World Bank, Washington DC, USA.

Jedwab, R., and A. Moradi (2015): "The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa," *Review of Economics and Statistics*, forthcoming.

Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. E. Sachs, and J. Sachs (2004): "A Global Index of the Stability of Malaria Transmission," *American Journal of Tropical Medicine and Hygiene*, 70(5), 486–498.

MercyCorps (2011): "Uganda Conflict and Market Assessment - Karamoja," Report.

Meredith, M. (2005): *The State of Africa: A History of the Continent since Independence*. The Free Press, London, United Kingdom.

Michalopoulos, S. (2012): "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 102(4), 1508–1539.

Michalopoulos, S., and E. Papaioannou (2013): "Pre-colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, 81(1), 113–152.

——— (2014): "National Institutions and Subnational Development in Africa," *Quarterly Journal of Economics*, 129(1), 151–213.

Montalvo, J.-G., and M. Reynal-Querol (2005): "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review*, 95(3), 796–816.

Murdock, G. P. (1959): *Africa: Its Peoples and their Culture History*. McGraw-Hill.

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

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

——— (2013): "Historical Development," in *Handbook of Economic Growth, Volume 2*, ed. by P. Aghion, and S. Durlauf, vol. 2. Elsevier, Amsterdam, Netherlands.

Nunn, N., and D. Puga (2012): "Ruggedness: The Blessing of Bad Geography in Africa," *Review of Economics and Statistics*, 94(1), 20–36.

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

O'Brien, P. (1999): *Atlas of World History*. Oxford University Press, Oxford, United Kingdom.

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

Posner, D. N. (2005): *Institutions and Ethnic Politics in Africa*. Cambridge University Press, New York, NY.

Raleigh, C., A. Linke, and C. Dowd (2014): "Armed Conflict Location and Event Dataset (ACLED) Codebook Version 3," Working Paper, Centre for the Study of Civil War, International Peace Research Institute, Oslo (PRIO).

Ross, M. L. (2012): *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton University Press, Princeton, NJ.

Sundberg, R., M. Lindgren, and A. Padskocimaite (2010): "UCDP GED Codebook version 1.0-2011," Department of Peace and Conflict Research, Uppsala University.

Sundberg, R., and E. Melander (2013): "Introducing the UCDP Georeferenced Event Dataset," *Journal of Peace Research*, 50(4), 523–532.

Thomson, A. (2010): *An Introduction to African Politics*. Routledge, London and Nyew York.

Tomas, J. (2006): "The Traditional Authorities Cross the National Border: Opposing Views on the Role of the Religious Leaders of the Jola Huluf and Ajamaat of the Lower Casamance (1886-1909)," *Africana Studia*, 9(1), 73–97.

Touval, S. (1967): "The Organization of African Unity and African Borders," *International Organization*, 21(1), 102–127.

Vaughan, S. (2003): "Ethnicity and Power in Ethiopia," Ph.D. thesis, The University of Edingurgh.

Wantchekon, L., M. Klasnja, and N. Novta (2015): "Education and Human Capital Externalities: Evidence from Colonial Benin," *Quarterly Journal of Economics*, 130(2).

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

Wilkenfeld, J. K. B. J., D. Brancati, J. Fearon, T. Gurr, D. Laitin, A. Pate, and S. Saideman (2014): "A-MAR (All-Minorities at Risk): Addressing the Selection Bias Issue," mimeo.

Wimmer, A., L.-E. Cederman, and B. Min (2009): "Ethnic Politics and Armed Conflict. A Configurational Analysis of a New Global Dataset," *American Sociological Review*, 74(2), 316–337.

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

Table 1 - Border (Ethnic Partitioning) Artificiality

Panel A: Geography, Ecology, Natural Resources and Ethnic Partitioning

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

Table 1 - Panel A reports linear probability model (LPM) estimates associating ethnic partitioning (SPLIT) with geographical, ecological and natural resource variables. In all specifications the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary country (using the 2000 Digital Chart of the World). All specifications include a set of (five) region fixed effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. The Supplementary Appendix reports summary statistics for all variables. Standard errors in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 1 - Border (Ethnic Partitioning) Artificiality

Panel B: Historical (Pre-colonial) Features and Ethnic Partitioning

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

Table 1 - Panel B reports linear probability model (LPM) estimates associating ethnic partitioning (SPLIT) with historical variables capturing pre-colonial conflict, kingdoms, and slavery (in Panel B). In all specifications the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary country (using the 2000 Digital Chart of the World). All specifications include a set of (five) region fixed effects (constants not reported). The Data Appendix gives detailed variable definitions and data sources. The Supplementary Appendix reports summary statistics for all variables. Standard errors in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 2: Ethnic Partitioning and Civil Conflict.
Baseline Country Fixed-Effects Estimates

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

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

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 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	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**
Double-clustered s.e.	(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
Double-clustered s.e.	(0.4761)	(0.0518)	(0.2964)	(0.5468)	(0.0534)	(0.3744)
Log Likelihood	-1453.054	__	-1046.922	-528.002	__	-383.392
adjusted R-square	__	0.471	__	__	0.441	__
Panel B: One-Sided Violence against Civilian Population						
SPLIT (Partitioning)	0.3468	0.0269	0.2750*	0.3288	0.0331	0.2925
Double-clustered s.e.	(0.2416)	(0.0292)	(0.1474)	(0.2615)	(0.0404)	(0.2237)
SPIL (Adjacent Split)	0.4708	0.0829*	0.4935**	0.0901	0.0161	0.1659
Double-clustered s.e.	(0.4549)	(0.0481)	(0.2277)	(0.6886)	(0.0626)	(0.3416)
Log Likelihood	-1499.837	__	-1099.667	-556.790	__	-396.804
adjusted R-square	__	0.404	__	__	0.434	__
Panel C: Non-State Civil Conflict						
SPLIT (Partitioning)	-0.2087	-0.0459	0.026	-0.4122	-0.0351	-0.1797
Double-clustered s.e.	(0.4062)	(0.0317)	(0.3374)	(0.5178)	(0.0283)	(0.4917)
SPIL (Adjacent Split)	-0.8703	-0.0344	-0.5089	-0.6593	-0.0268	-0.6964
Double-clustered s.e.	(0.7193)	(0.0302)	(0.4423)	(0.8728)	(0.0415)	(0.7086)
Log Likelihood	-841.675	__	-644.791	-243.970	__	-199.677
adjusted R-square	__	0.320	__	__	__	__
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1212	1212	1212	579	579	579

The table reports estimates associating three types of civil conflict with ethnic partitioning at the country-ethnicity homeland level using data from the UCDP GED 1.5 project. Panel A gives estimates focusing on state conflict (where government forces, troops, and militias participate). Panel B gives estimates focusing on one-sided violence, mostly against the civilian population. Panel C gives estimates focusing on conflict between (at least) two non-state actors (where the government is not involved). The dependent variable in columns (1) and (4) is the total number of civil conflict incidents at each country-ethnic homeland over the period 1989-2010. These models are estimated with the negative binomial maximum likelihood model. The dependent variable in columns (2) and (5) is a dummy variable that takes on the value of one for country-ethnic homelands that have experienced each conflict type over the period 1989-2010 (and zero otherwise). The dependent variable in columns (3) and (6) is the number of years that each country-ethnic homeland has experienced each type of conflict over the period 1989-2010. These models are estimated with the negative binomial maximum likelihood model. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. SPIL –that captures spillovers- is the share of adjacent partitioned ethnic homelands to the total number of adjacent ethnic homelands.
The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country homeland to the national border; 61.3 kilometers). All specifications include a country fixed effects (constants not reported) and a rich set of control variables. The set of simple controls includes the log of land area, the log of population in 1960, an indicator for lakes and an indicator for rivers. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country and an indicator for country-ethnic areas that are by the sea coast. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, an oil field indicator, and an indicator for areas with major city in 1400. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

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

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

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

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

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

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

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

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	-0.3122**	-0.2078**	-0.2893***	-0.2440***	-0.3760*	-0.2908**	-0.2832**	-0.1659*	-0.1922*	-0.1484*	-0.3654*	-0.2410**
Double-clustered s.e.	(0.1236)	(0.0939)	(0.1058)	(0.0797)	(0.1931)	(0.1256)	(0.1340)	(0.0853)	(0.1007)	(0.0773)	(0.2009)	(0.1167)
Partitioning - Location	-0.2275**	0.041	0.0346	0.0983	-0.2873*	-0.0392	-0.1039	0.0729	0.0904	0.1241**	-0.1808	0.0003
Double-clustered s.e.	(0.1134)	(0.0811)	(0.1364)	(0.0815)	(0.1494)	(0.1037)	(0.0784)	(0.0628)	(0.0992)	(0.0630)	(0.1109)	(0.0819)
Non-Indigenous Indicator	0.1894**	0.1837***	0.1371*	0.2013***	0.2217*	0.2256***	0.1062	0.1047	0.0601	0.0903*	0.1479	0.1426*
Double-clustered s.e.	(0.0911)	(0.0684)	(0.0771)	(0.0671)	(0.1153)	(0.0784)	(0.0852)	(0.0696)	(0.0665)	(0.0526)	(0.1134)	(0.0831)
Partitioned Location & Identity					(0.1310)	(0.1687)					(0.1686)	(0.1527)
Double-clustered s.e.					(0.2297)	(0.1312)					(0.1798)	(0.1151)
Adjusted R-squared	0.032	0.167	0.053	0.162	0.032	0.167	0.151	0.255	0.133	0.214	0.152	0.256
Observations	88171	88171	44090	44090	88171	88171	88171	88171	44090	44090	88171	88171
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Location Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

Data Appendix

The Long-Run Effects of the Scramble for Africa

1 Ethnic Partitioning

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

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

2 Civil Conflict Variables (ACLED)

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

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

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

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

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

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

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

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

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

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

3 Civil Conflict Variables (UCDP)

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

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

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

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

Conflict Duration: For each type of conflict (state-driven, one-sided violence, and non-state-driven) we defined variables measuring the number of years that a country-ethnic area has experienced at least one (high geo-precision) conflict event over the period $1989-2010$. *Source: UCDP GED 1.5.*

4 Control Variables

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

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

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

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

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

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

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

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

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

Distance to the Sea: The geodesic distance of the centroid of the historical homeland of each ethnic group (or each country-ethnic area) from the nearest coastline, measured in 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).*

5 Pre-colonial Features

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

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

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

6 DHS Data

Composite Wealth Index: The wealth index is a composite measure of almost all household assets and utility services including country-specific items. The wealth index is calculated using easy-to-collect data on a household's ownership of selected assets, such as televisions and bicycles; materials used for housing construction; and types of water access and sanitation facilities. Generated with a statistical procedure known as principal components analysis, the wealth index places individual households on a continuous scale of relative wealth within a country. More details are available here: http://www.measuredhs.com/topics/Wealth-Index.cfm and here: http://www.measuredhs.com/pubs/pdf/CR6/CR6.pdf. *Source: Demographic and Health Surveys (http://www.measuredhs.com/).*

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

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

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

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

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

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

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

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

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

Distance to the National Border: The geodesic distance from the location (gps coordinates) of

each household to the nearest national border. *Source: Calculated using ArcGis.*

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

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

References

Besley, T., and M. Reynal-Querol (2014): "The Legacy of Historical Conflicts. Evidence from Africa," *American Political Science Review*, 108(2), 319–336.

Brecke, P. (1999): "Violent Conflicts 1400 A.D. to the Present in Different Regions of the World," Paper prepared for the 1999 Meeting of the Peace Science Society in Ann Arbor, Michigan.

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

Cioffi-Revilla, C. (1996): "Origins and Evolution of War and Politics," *International Studies Quarterly*, 40(1), 1–22.

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.

Michalopoulos, S. (2012): "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 102(4), 1508–1539.

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

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

O'Brien, P. (1999): *Atlas of World History*. Oxford University Press, Oxford, United Kingdom.