THE LONG-RUN EFFECTS OF THE SCRAMBLE FOR AFRICA

Stelios Michalopoulos
Elias Papaioannou

The Long-Run Effects of the Scramble for Africa
Stelios Michalopoulos and Elias Papaioannou
NBER Working Paper No. 17620
November 2011, Revised October 2013
JEL No. N17,N47,O10,Z10

<u>ABSTRACT</u>

We examine the long-run consequences of a neglected aspect of colonization, the artificial drawing of borders during the Scramble for Africa and uncover the following empirical regularities. First, apart from the land mass and water area, no other pre-colonial trait predicts a group's partitioning. Second, using georeferenced data on conflict we show that battles, violence against civilians and territorial changes are concentrated in the historical homeland of partitioned ethnicities. Third, we show that individuals identifying with split groups are on average poorer and less educated. The uncovered evidence brings in the foreground the violent repercussions of ethnic partitioning.

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

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

1   Introduction

The predominant explanations on the deep roots of contemporary African underdevelopment are centered around the influence of Europeans during the colonial period (Acemoglu *et al.* (2001, 2002, 2005)), but also in the centuries before colonization when close to 20 million slaves were exported from Africa (Nunn (2008), Nunn and Wantchekon (2011)). Yet in the period between the ending of the slave trades and the colonial rule, another major event took place in European capitals that according to the African historiography had malicious long-lasting consequences. The "Scramble for Africa" starts with the Berlin Conference of $1884 - 1885$ and is completed by the turn of the 20th century. In this brief period, Europeans partitioned Africa into spheres of influence, protectorates, colonies, and free-trade areas. The borders were designed in European capitals at a time when Europeans had barely settled in Africa and had little -if any- knowledge of 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 the African historiography (e.g., Asiwaju (1985); Dowden (2008); Wesseling (1996)) argues that the main channel of Europeans' influence on African development was not colonization per se, but the improper border design. As Herbst (2000) summarizes "*for the first time in Africa's history [at independence], territorial boundaries acquired salience...The boundaries were, in many ways, the most consequential part of the colonial state.*" The artificial borders fostered ethnic struggles, patronage politics, violence, 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)). African borders did not only split tribal homelands across countries, they also produced some of the largest and most heterogeneous countries in the world, such as the Democratic Republic of Congo and Sudan. Moreover, the Scramble for Africa created several hinterland states with challenging geographies that limit national governments' ability to broadcast power effectively (Herbst (2000)).

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

---

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

While there is little disagreement among historians that colonial (and hence post-independence) borders were drawn with little respect to the local political geography, we start our analysis establishing formally this thesis. We do so exploring whether partitioned ethnicities differ systematically from non-split groups across a variety of geographic, ecological, and pre-colonial social, economic, and cultural traits. With the sole exceptions of the land mass of the historical ethnic homeland and area under water, there are no significant differences between partitioned and non-split ethnicities with respect to geography (elevation, distance to the coast, land's agricultural quality), the disease environment (malaria), the presence of natural resources (diamond mines and oil fields). We further show that there are no systematic differences between split and non-split groups, across several pre-colonial ethnic-specific institutional, cultural, and economic features, such as the size of settlements, the type of the subsistence economy, etc. We also show that there is no association between ethnic partitioning and two distinct proxy measures of pre-colonial conflict, one based on pre-colonial wars mostly between African kingdoms and the other capturing conflict during the slave trades. Our large-scale econometric results thus support the claim of the African historiography that in the overwhelming majority of cases, Europeans did not take into account local political and geographic aspects while designing the borders.

We then employ the scramble for Africa as a quasi-natural experiment to assess the impact of ethnic partitioning on regional civil conflict. Using a new dataset that reports detailed geo-referenced information for $1997 - 2010$ on the exact location of more than $43,000$ incidents of political violence, including battles between government forces, rebel groups and militias, changes of territorial control, as well as violence against civilians (the latter includes murders, abductions, child soldiering raids, rapes, mutilations), we show that civil conflict is concentrated in the homelands of partitioned ethnicities. Our regional focus allows us to account for the numerous country-wide factors that interact with civil conflict, such as national institutions and politics, the type of colonization, fractionalization, etc. Furthermore, by moving the analysis from the national to the ethnic homeland level, we are able to condition on ethnic-family fixed effects and thus take into consideration cultural, genetic, social, and economic differences across the quite heterogeneous African ethnicities. We obtain similar results when we restrict estimation to ethnic homelands close to national borders. Our most conservative estimates suggest that civil conflict intensity is approximately 30% higher in areas where partitioned ethnicities reside as compared to the homelands of ethnic groups that have not been separated by the national borders. We further find that homelands of partitioned groups experience a 5% to 10% increased likelihood of a territorial control change between the government and rebel groups. It is not only army fighting that is concentrated in the homelands of partitioned groups. Violence against civilians is roughly 40% higher in regions where split groups reside. The evidence thus uncovers the on-going violent repercussions of the colonial border design.

Finally we complement the regional analysis providing micro-level evidence illustrating the adverse economic consequences of ethnic partitioning. Exploiting information from the Demographic and Health Surveys (DHS) spanning more than $85,000$ households in 20 African countries we show that individuals

identifying with partitioned groups have fewer household assets, poorer access to utilities, and worse educational outcomes, as compared to individuals from non-split ethnicities in the same country and even in the same village/town. The negative effect of ethnic partitioning on education, household wealth, and public goods is present for both old and young cohorts revealing the persistent deleterious consequences of the Scramble for Africa. Jointly these findings suggest that continuous discrimination against split groups in the provision of public goods (such as education and access to basic utilities) is a likely channel via which ethnic partitioning operates on underdevelopment and conflict.

1.1 Historical Background

The "Scramble for Africa" starts in the 1860s when the French and the British begin the systematic exploration of Western Africa and sign bilateral agreements assigning spheres of influence to each other. In the next 40 years, Europeans signed hundreds of treaties that partitioned the largely unexplored continent into protectorates, free-trade areas, and colonies. The event that stands for the partitioning of Africa is the conference that Otto von Bismarck organized in Berlin from November 1884 till February 1885. While the Berlin conference discussed only the boundaries of Central Africa (the Congo Free State), it came to symbolize the partitioning, because it laid down the principles that would be used among Europeans to divide the continent. 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.[2] 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 idiosyncratic border design. First, at the time Europeans had very 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.[3] Second, Europeans were

---

[2] 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*".

[3] 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*

not drawing borders of prospective states or -in many cases- even colonies (in most cases Europeans ruled Africa indirectly via protectorates where they had minimal presence). Third, while in most cases the treaties indicated that the exact boundaries would be set by special commissions, demarcation was poor. Fourth, Europeans were not willing to sacrifice their commitment not to go to war for any part of Africa and hence were reluctant to change colonial borders.[4] In many cases London and Paris turned down requests from local administrators to redraw the border because it did not coincide with a physical boundary or because an ethnic group was split. Fifth, as there was an implicit agreement between Europeans that ethnicities could freely move across colonial borders, African chiefs did not oppose the colonial design, as little changed on the ground.[5] Wesseling (1996) summarizes: "*The partition of Africa was recorded by the Europeans on their maps, but the matter rested there for the time being....In Europe conquests preceded the drawing of maps; in Africa the map was drawn, and then it was decided what was going to happen. These maps did not therefore reflect reality but helped to create it.*"

The other major event in recent African history, namely independence, occurred at a speed that not even the key protagonists expected (Herbst (2000)). The independence of Northern African countries in the 1950s was soon followed by Ghana's and Guinea's independence in 1957 and in 1958, respectively. By the end of 1966, 40 countries had become independent. While at the time, many proposed changing the borders, African leaders and leaving Europeans did not touch this issue. The leaders of the newly-crafted African states believed that nation building and industrialization would sideline ethnic divisions. Moreover, national leaders feared that a border realignment would threaten their position whereas Europeans' main objective was to maintain the special rights and corporate deals with their former colonies, and, as such, they were also reluctant to open the border issue.[6]

1.2 Channels and Case Studies

Irredentism, secession, and autonomy The literature has stressed the impact of ethnic partitioning on generating irredentist demands, as split ethnicities may want to unify with their peers across the border.[7] In line with this argument, Wimmer, Cederman, and Min (2009) estimate that 20% of all

---

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

[4]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.*"

[5]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.*"

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

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

civil wars in Africa have a secessionist demand.[8] 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 tribes, 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 tribes (Somalia, North Kenya, Southern Ethiopia, Djibouti, and Eritrea); three long-lasting wars have been partly driven by the desire of Somalis in Ethiopia, Djibouti, and Kenya to become part of Somalia (Meredith (2005)). In our sample that covers the period $1997 - 2010$, the bulk of battles and violent events against civilians have taken place in the Ogaden region in Southern Ethiopia where Somali tribes reside. Specifically, in Ethiopia (that in Murdock's map comprises 48 ethnic homelands), 33% of a total of 961 battles between government forces, rebel groups and militias as well as 19% of 295 violent events against civilians occurred in the Ogaden region where the partitioned Afar and the other Somali clans are located (while only 7% of Ethiopia's total population resides in these regions). Similarly, in the initial years after independence Kenya experienced substantial conflict in the Northern Frontier District as Somalis were fighting (*shifta*) for annexation to Somalia (Touval (1967)). Another example is the Tuaregs who in the early 1990s declared autonomy both in Niger and in Mali, with some of their leaders envisioning a unified Tuareg state. Moreover, even in cases that obtaining autonomy 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)).

Spillovers (Shelter-Seeking and Cross-Border Migration) Groups in Africa often become targets of the central governments. The latter attempt to suffocate ethnicities by seizing property and imposing high taxation in the activities of specific ethnic groups (Bates (1981)). What is different between split and non-split groups is that partitioned ones can seek shelter within their homeland on the other side of the border. Moreover, split ethnicities can re-organize, obtain arms, and get assistance from their co-ethnics across the national boundary both when they are on the defence and when they attack. Similarly, partitioned groups are frequently used by African governments to destabilize neighboring countries, promoting their interests. An example is the case of the Ewe, a 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. This escalated ethnic tensions between the Ewe, the Ashanti, and the Akan, in Ghana leading to conflict in the subsequent 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" was supported by the neighboring Guinea-Bissau, where the Diola exert significant political influence (Humphreys

---

more than one North African state."

[8]Civil wars with a secession demand are almost absent in Central and South America. Besides African regions secession-driven civil conflicts are also found in the Middle East, India, and the Caucasus.

and Mohamed (2005)). In Senegal Murdock (1959) maps 12 ethnic homelands. In our sample 40% of a total of 198 battles and 40% of 140 violent events against civilians have taken place in the homeland of the partitioned Diola, where only 5% of Senegal's population resides.[9]

Spillovers may also occur because of population displacement, as quite often refugee flows change the ethnic composition in adjacent countries igniting conflict. This has been the case for the Alur-land. The Alur had been split between the Belgian Congo and the British Protectorate of Uganda during the late phase of the scramble for Africa $(1910-1914)$. When Mobutu Sese Seko initiated the subjugation of several minority groups in Zaire, a large portion of the Alur were pushed to their homeland in Uganda. This in turn generated opposition from the Buganda in Southern Uganda leading to conflict (Asiwaju (1985)). In line with this case, Fearon and Laitin (2011) report that 31% of all civil wars (and 57% of all ethnic wars) involve "*members of a regional ethnic group that considers itself to be the indigenous sons-of-the-soil and recent migrants from other parts of the country*".[10]

Other Aspects of the Scramble for Africa Besides ethnic partitioning, the artificial border design has contributed to underdevelopment and conflict via other channels. In particular, 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 sea that affect development. For example, Herbst (2000) argues that civil conflict is more pervasive in large African countries that find it hard to broadcast power across their territories (see Michalopoulos and Papaioannou (2013a) for empirical evidence). 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, ethnic partitioning, we are able to account for these other aspects of European's influence with the inclusion of country fixed effects that absorb all time-invariant, country-specific characteristics.[11]

1.3   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 formation of early institutions (e.g., Acemoglu, Johnson, and Robinson (2005)), infrastructure (e.g., Huillery (2009), Jedwab and Moradi (2012)), and human capital (e.g., Glaeser, La Porta, de Silanes, and Shleifer (2004) and Easterly and Levine (2012)). We emphasize a neglected aspect of colonization, the

---

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

[10]Fearon and Laitin (2011) classification lists eight conflicts in African (26% of all wars) that involved indigenous versus migrant conflict (e.g., Tuaregs in Mali in 1989; Senegal in 1989, involving Diolas in the Casamance region, Darfur in 2003 and Zimbabwe in 1972). Overall Asia is the continent with the most indigenous-versus-migrant conflict.

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

drawing of political boundaries in the end of the 20th century that resulted in a large number of partitioned ethnicities. As such our work is mostly related to Alesina, Easterly, and Matuszeski (2011) who show that "artificial states" with straight borders and where a significant part of the population resides in more than one country, under-perform economically compared to more "organic" countries.

Our work also fits a growing strand of literature that examines the roots of African underdevelopment and conflict. Nunn (2008), Nunn and Puga (2012), and Nunn and Wantchekon (2011) show that the slave trades have crucially shaped African development mostly by spurring ethnic conflict and by lowering trust. Gennaioli and Rainer (2006, 2007) and Michalopoulos and Papaioannou (2013b) show that deeply-rooted pre-colonial ethnic institutions correlate significantly with contemporary economic development. A closely related study is that of Besley and Reynal-Querol (2012), who use geo-referenced data on pre-colonial conflict and show that the latter is associated with a greater prevalence of contemporary conflict.

Our work also contributes to the literature on the origins of civil conflict that mainly examines the role of country-level characteristics, such as income and natural resources (see Collier and Hoeffler (2007) and Blattman and Miguel (2010) for reviews and Collier and Sambanis (2005) for case studies in Africa). Of most relevance are works that link a country's ethnic composition to civil war. While the correlation between ethnic fragmentation and civil war is weak (Fearon and Laitin (2003)), recent studies document interesting cross-country correlations associating various aspects of the societal structure with armed conflict. Montalvo and Reynal-Querol (2005) and Esteban, Mayoral, and Ray (2012) show a strong negative association between ethnic polarization and conflict. Wimmer, Cederman, and Min (2009) find that the likelihood of ethnic conflict increases when a large share of the population is excluded from power. Englebert, Tarango, and Carter (2002) show a positive cross-country correlation between proxy measures of suffocation and dismemberment and violence, secession attempts, border disputes, and warfare.[12]

The correlations found in studies linking cross-country variation in border features and the distribution of ethnicities to development proxies (income or conflict) are informative; yet they cannot be easily interpreted (see Blattman and Miguel (2010) for a discussion). The main endogeneity concern is that the process of border drawing is usually an outcome of state formation that affects directly economic performance and conflict. Moreover, as the recent literature on state capacity shows, nation building, development, and conflict are all inter-linked; and jointly driven by hard-to-account-for factors related to societal structure, geography, and historical legacies (see Besley and Persson (2011)). Thus, issues of selection, reverse causation, and omitted variables are non-negligible in cross-country works. Likewise, due to measurement error in the main independent variables, multi-colinearity, and the limited degrees of freedom, the correlations documented in cross-country studies are often quite sensitive even to small permutations

---

[12]Our paper also makes contact with works in political science and security studies on nationalism. Using mostly case-studies and narrative arguments, this body of research examines the origins and consequences of ethnic identification, nationalism, and irredentism (e.g. Fearon and Laitin (2000), Brancati (2006)), as well as the pros and cons of regional separation as a solution to ethnic conflict (e.g. Kaufman (1996, 1998) , Fearon (2004), Englebert (2009)). Sambanis (2000), Chapman and Roeder (2007), and Sambanis and Schulhofer-Wohl (2009) estimate cross-country regressions examining whether regional separation results in long-lasting peace.

and data revisions (see Hegre and Sambanis (2006) and especially Ciccone and Jarocinski (2010), who show that the identification of the correlates of cross-country growth are fragile even to minimal data revisions).

By exploiting variation across ethnic homelands, we account for many of the shortcomings of existing cross-country works. First, by showing that there are no systematic differences in geographic, economic, institutional, and cultural characteristics between partitioned and non-split ethnicities, our analysis offers large-scale econometric evidence supportive to the African historiography on the accidental nature of most African borders. Moreover, our results that on average there is no association between ethnic partitioning and various proxies of pre-colonial conflict and the plethora of case studies documenting the idiosyncratic nature of African borders enables the causal interpretation of the estimates.

Second, the use of information on the spatial distribution of ethnicities in the end of 19th century, well before the current national boundaries came into effect, alleviates concerns related to migratory flows ignited by the border design. Since borders were drawn by Europeans, we are able to focus on cases where country boundaries were not the result of political, economic, or military developments; thus selection and reverse causation issues are secondary -if not absent.

Third, our analysis at the ethnic homeland is conceptually appealing in the context of Africa, where ethnic identification is especially strong and conflict has a strong ethnic component. In their synthesis of the case-study evidence on conflict in Africa and the results of cross-country regressions, Collier and Sambanis (2005) note "*the country-year is not the appropriate unit of observation to study such wars. Instead it would be more appropriate to focus on the ethnic group or we should analyze patterns of violence in a geographical region that does not necessarily correspond to predefined national boundaries. With current data limitations, however, it may not be feasible to adjust the unit of analysis problem.*"

Fourth, the use of micro-level data on well-being (from the DHS) enables us to further account for local and individual features, study inertia, and shed light on the channels. Our findings that the negative repercussions of ethnic partitioning apply to old and young cohorts alike illustrate its persistent effect, while our result that individuals identifying with a partitioned group have lower education and public goods even when compared to other individuals in the same village/town suggests that partitioned groups face discrimination and marginalization.

Structure In the next section we first discuss how we identify partitioned ethnic groups and present the geo-referenced civil conflict data. We then report descriptive statistics illustrating the significant differences in the likelihood and intensity of armed conflict between partitioned and non-split ethnic homelands. In Section 3 we examine whether there are systematic differences between partitioned and non-partitioned ethnicities with respect to an array of geographic and historical features that may independently affect conflict. Section 4 reports our estimates on the effect of ethnic partitioning on various aspects of civil conflict. Section 5 presents the micro-level results linking education and public goods on ethnic partitioning at the individual level. In Section 6 we summarize discussing possible avenues for future research.

2 Data

2.1 Identifying Partitioned Ethnic Groups

We identify partitioned groups projecting contemporary borders, as portrayed in the 2000 Digital Chart of the World on George Peter Murdock's Ethnolinguistic Map (1959) that depicts the spatial distribution of African ethnicities at the time of European colonization in the mid/late 19th century (Figure 1a).[13] Murdock's map divides Africa into 843 ethnic regions. The mapped ethnicities correspond roughly to levels $7 - 8$ of the Ethnologue's language family tree. 8 areas (all in the Sahara) are "uninhabited upon colonization" and are therefore not considered in our analysis. We also drop the Guanche, a small group in the Madeira Islands that is currently part of Portugal. Out of a total of 834 ethnicities in Murdock's Map, the homeland of 358 groups falls into more than one 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 Niger and only 0.5% falls into Algeria. 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 231 ethnicities with at least 10% of their historical homeland falling into more than one contemporary state (Figure 1b). Appendix Table A lists partitioned ethnicities. When we use a broader threshold of 5% we identify 267 partitioned groups. In our analysis we also exclude 8 regions where population according to the earliest post-independence census is zero. Thus, in our baseline sample we have a total of 826 populated ethnic areas of which 230 are partitioned.[14]

Our procedure identifies most major ethnic groups that have been split by African borders. For example, the Maasai are partitioned between Kenya and Tanzania (shares 62% and 38%, respectively), the Anyi between Ghana and the Ivory Coast (shares 58% and 42%, respectively), and the Chewa between Mozambique (50%), Malawi (34%), and Zimbabwe (16%). Other examples include the Hausa (split between Nigeria and Niger), the Ababda (split between Egypt and Sudan), and the Bararetta Somali clans (split between Kenya and Somalia). We also checked whether our coding of partitioned ethnicities is in line with Asiwaju (1985), who provides the only (to our knowledge) codification of partitioned ethnicities in Africa. Our strategy identifies almost all ethnic groups that Asiwaju (1985) lists as partitioned. For robustness we also construct a continuous index of partitioning in the spirit of the ethnic/linguistic fragmentation indicators (e.g., Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003)).[15]

---

[13]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.

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

[15]We prefer the binary index of partitioning for several reasons. First, all studies in African historiography suggest that what matters for civil conflict is whether an ethnicity has been partitioned or not rather than the degree of the split. Second, there is no clear reason why conflict propensity should monotonically increase with the degree of partitioning. Third, as Murdock's map certainly contains noise, this will be reflected more clearly in the continuous measure (as compared to the binary index). Nevertheless, to show that our results are not sensitive to the index of partitioning in the Appendix we report specifications
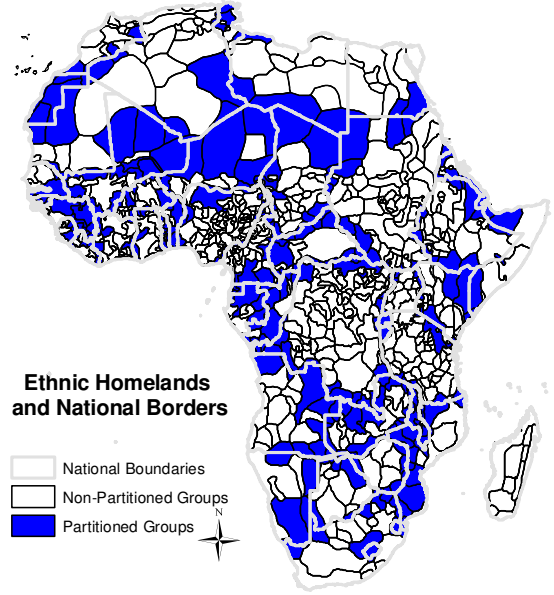
Figure **1a**

Figure **1b**

2.2   Civil Conflict Data

Our data for conflict come from the Armed Conflict Location and Event Dataset (Raleigh, Linke, and Dowd (2012b)) that provides geo-referenced information on the location and some key characteristics of violent events across all African countries since 1997.[16] A nice feature of this dataset (compared to the standard Correlates of War and PRIO data on civil wars) is the reporting of incidents of violence against the civilian population. This is quite important as Africa is plagued by prolonged civil warfare that the standard data sources of civil war miss (as they focus on battles). Not only violence against civilians, such as child soldiering raids, rapes, abductions, and burning is recurrent, but these incidents are deadly, economically harmful, and devastating for the victims and the local community.

The ACLED database is detailed listing 43, 271 incidents of political violence over the period 1997−2010. Political violence is understood as the use of force by a group with a political purpose or motivation. Conflict groups (actors) include rebels, militias, governments, and organized political or ethnic groups that interact violently over issues of political authority, such as territorial control, government control, access to resources, etc. ACLED categorizes armed conflict into 8 types. (1) Battles without change of control; (2) Battles where rebel groups gain control of the location; (3) Battles where the government regains control of a location; (4) Headquarter of base establishments, where rebel groups establish their base; (5) Non-violent conflict events where rebel groups, militias or government forces proceed in non-violent actions (without

---

with the continuous index.

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

active fighting) that are, however, within the context of an on-going civil conflict (e.g., recruitment drives and incursions); (6) Riots and protests; (7) Violence again civilians, where armed groups (rebels, militias or government forces) attack unarmed civilians; (8) Non-Violent transfer of control.

The data are based on a quite diverse set of sources.[17] For almost all countries data come from more than ten different sources, while for the most war-prone ones we have data from around twenty sources. This diversity assuages concerns of systematic biases in reporting from government controlled actors and media. The data are mostly based on international sources, such as the BBC (close to $10,000$ incidents), Reuters (close to $5,000$ incidents), the Associated Press (around $1,500$ incidents), and the Agence France Press (around $3,500$ incidents). A considerable fraction of the data (around 10%) comes from 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.[18] Even in cases of data coming from local sources (around 25% of the sample), most incidents come from pan-African news agencies, such as the *All Africa* network and independent –rather than government run/controlled- newspapers.

Our benchmark index of civil conflict is the count of all types of armed conflict. We also examine the effect of ethnic partitioning on (i) the total number of battles, (ii) battles that resulted in a change of territorial control, and (iii) violent events against civilians. Battles among armed forces account for 43.2% of all incidents; out of a total of $18,705$ battles, $2,324$ resulted in a change of territorial control. The dataset records information on $15,844$ cases of violence against civilians (36.6% of all incidents).

Examples of battles between armed actors include the fights between the Lord's Resistance Army (LRA), the Sudanese People's Liberation Army, and Uganda's People Defence Force (UPDF); the constant fighting between the Rwandan forces against the FDLR (Forces démocratiques de libération du Rwanda) Hutu rebels in Rwanda and in Eastern Congo; and the battles between Kikuyu rebel groups against Maasai militias. Battles result usually in casualties; for example, the brief 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. Battles resulting in territorial change of control are usually more devastating involving both a higher number of casualties and ambushes against the civilian population. For example, ACLED reports that in August 1997 when government forces of the Democratic Republic of Congo (DRC) retook control of the town of Watsa, close to the border with Uganda, where the partitioned Alur reside, there were 800 casualties.

---

[17]In parallel work Besley and Reynal-Querol (2012) and Harari and Ferrara (2012) also use the ACLED to study the role of other than ethnic partitioning correlates of conflict.

[18]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 a lot of incidents are reported by the Zimbabwe Human Rights NGO Forum, a coalition of nineteen human rights NGOs in Zimbabwe that get the data from their representatives on the ground. Likewise, 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 conflict events. Similarly, in Kenya many incidents are based on reports from the local Peace and Development Network Trust, a local NGO, cofounded by Oxfam that monitors peace.

Violent events against civilians include the raids of the Janjaweed militias against civilian population in the Darfur region in Western Sudan; the assaults and tortures of the Central Intelligence Organization in Zimbabwe; the killings of civilians in Northern and Western Rwanda by the Interahamwe Hutu ethnic militias (that are raiding from their bases in Eastern Congo and Uganda); and the killings, abductions, rapes, and terrorist activities of militias in the DRC. Violent events include also the burning of churches, hostage-taking and child-soldiering raids by rebels in Nigeria and in Sierra Leone. Going over the narratives of each event reveals that they may be also quite devastating. For example, according to the ACLED brief event description in a single day in Eastern Congo in May 1997 "*ADLF rebels moved in and took control of Mbandaka slaughtering* 200 *Rwandan Hutu refugees*".[19]
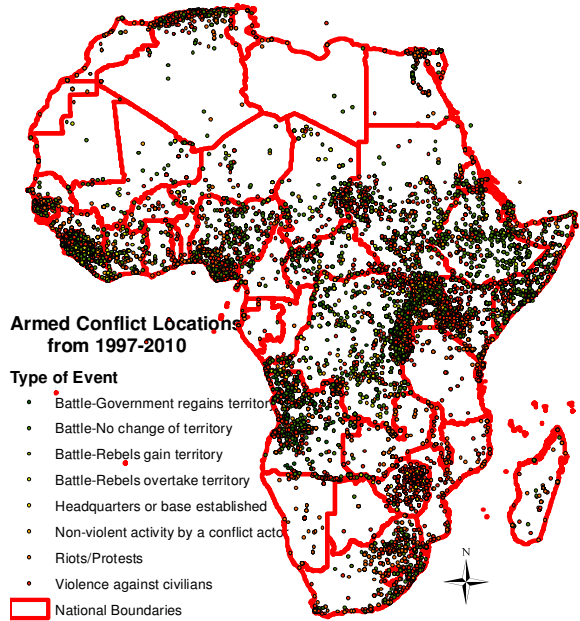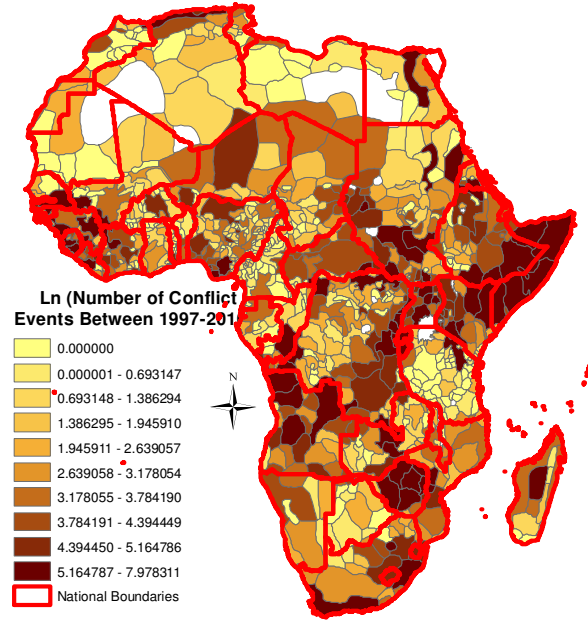
2.2.1 Data Patterns

Figure **2a**

Figure **2b**

Figure 2*a* maps the spatial distribution of all conflict events over the period 1997 − 2010. There is significant heterogeneity in the incidence of political violence. There are numerous events in Central Africa, mostly in Eastern Congo, Rwanda, Burundi, Sudan, and Uganda. In Western Africa, conflict and political violence are mostly present in Nigeria and in Sierra Leone. Violence is also pervasive in Somalia, Ethiopia, and Zimbabwe. In contrast there are few events in Botswana, Zambia, Tanzania, Namibia, and Gabon. There is also considerable variation within countries. For example, while conflict incidence in Tanzania is low, there are quite a few violent events along the border with Kenya and Rwanda. Likewise, most of the conflict in Senegal is concentrated on the Southern region of Casamance, while most conflict in Angola is close to the northern border with Congo and in the Cabinda enclave.

---

[19]These numbers are estimates and not reported for all events.

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

We examine separately the effect of partitioning on the number of battles and violence against civilians, as these are the dominant and most harmful types of conflict. The correlation between battles and violence against civilians in high, but far from perfect (0.60). For example, in Zimbabwe we observe many violent events against civilians and very few battles. Conversely in Ethiopia and Sudan we predominantly observe conflict between the government and rebels rather than unilateral violence against civilians. In the DRC and in Uganda the numbers of battles go hand in hand with civilian violence. The correlation between battles resulting in a territorial change and the total number of battles is 0.59 and with violent events against civilians is 0.22, suggesting that we can examine the effect of ethnic partitioning not only on overall political violence but also on its main components.

2.2.2 Summary Statistics

Table 1 reports descriptive/summary statistics for the main outcome variables across the 826 ethnic homelands.[20] In Panel A we report summary statistics (mean, median, and standard deviation) across all ethnic homelands, while in Panel B we report statistics for homelands close to the national border (using the median value of distance from the centroid of a group; 102 kilometers). This helps us isolate the role of ethnic partitioning from an overall border effect (which, nonetheless, may still be driven by partitioning itself).

Civil Conflict Incidents: All Types Three-fourths of all ethnic areas have experienced at least one conflict event over the period $1997-2010$ (column (1)). 84% of partitioned ethnicities experienced some conflict, while the likelihood of a civil conflict incidence for non-partitioned ethnicities is 11 percentage points lower (73%). When we focus on groups close to the border, we observe similar differences; on average 65% of non-partitioned ethnic homelands experienced a conflict, while 83% of split groups suffered from some type of conflict. Partitioned groups have also experienced more violent events. On average partitioned ethnicities experienced 64.75 incidents, while for the rest 47.6 incidents are recorded. This difference is not statistically significant because there are some extreme cases both across partitioned and non-split ethnic groups (see Appendix Table 2). To account for outliers we exclude ethnic homelands where capitals fall (in (4) and (5)) and homelands where the number of armed conflicts exceeds the $99th$ percentile (in (6)-(7)). The differences are now statistically significant. There are also large differences between partitioned (median=13) and non-split ethnicities across the median value of conflict events (13 compared to 3). The differences in conflict intensity between partitioned and non-split ethnic groups are also sizable when we focus on areas close to the border (Panel B). While the average (median) number of

---

[20]In Supplementary Appendix Table 2 we report summary statistics for all outcome and control variables.

all civil conflict incidents for partitioned ethnicities is 66 (12), for non-split ones the average (median) is 32.3 (1).[21]

Battles   On average 59% of all ethnic homelands has experienced at least one battle between government forces, rebels, or militias. The corresponding likelihood for partitioned and non-split groups in the full sample is 63.5% and 57.7%, respectively. When we restrict estimation to areas close to the national border, the difference between partitioned and non-split groups in the probability of experiencing at least one battle is 15%. On average partitioned ethnic homelands have experienced ten more battles as compared to non-split groups (29.9 versus 19.8); and while due to outliers this difference is not statistically significant, once we exclude the top 1% of the conflict distribution or regions where capitals fall, the difference becomes significant. Likewise, the median value of battles for non-split ethnic groups close to the national border is zero, while the corresponding median value for partitioned ethnic homelands is 3.

Territorial Control Change   The ACLED database also reports battles that resulted in territorial change of control. Focusing on such battles is interesting as in these cases the local population is likely to be more dramatically affected; moreover, these incidents may reflect more accurately irredentist demands and state collapse. The likelihood that a battle resulting in a change of territorial control for partitioned ethnic homelands is 36%, while the corresponding likelihood for non-split groups is 22.7%. This pattern suggests that partitioned ethnic groups are more likely to be traumatized as control oscillates between the government and rebel forces.

Violence against Civilians   The summary statistics of violence against civilians also reveal large and significant differences between partitioned and non-split groups. The likelihood that a partitioned ethnicity has experienced at least one violent event against the civilian population is 0.70, while the corresponding likelihood for not split ethnicities is 0.57. The difference is larger when we focus on ethnic homelands close to the national border (0.24). On average partitioned ethnic homelands experience 24 violent incidents against the civilian population, while the average for non-split ethnicities is 17.6. The median value of violence against civilians across partitioned ethnic homelands is three times the median value across non-split groups (3 versus 1) independently on whether we examine all ethnic homelands or we limit our attention to those close to the national border.

3   Border Artificiality

The African historiography provides ample evidence arguing that, in the majority of cases, Europeans did not consider ethnic politics and local geography in the design of colonial borders. The latter being mostly

---

[21]The results are similar if we use a narrower threshold of distance to the national border to identify ethnic homelands close to the national border. For example when we use the 25% percentile of distance to the border (45 km), the average (median) number of civil conflict incidents for partitioned ethnicities is 47 (8.5) while for non-split ethnicities 4.6 (1).

drawn before Europeans settled in Africa. In a few instances, nevertheless, Europeans did try taking into account political geography, as, for example, in Swaziland, Burundi, and in some parts of Uganda. Moreover, in two cases (Cameroon-Nigeria; Ghana-Togo) there were referenda on the redrawing of borders at independence. Yet what is key for establishing causality is not that all borders are idiosyncratically drawn (though many were); what is needed for causal inference is that there are no major differences between partitioned and non-split ethnic homelands across factors that may independently affect conflict. Thus in this section we examine in detail whether there are significant differences between the two sets of ethnicities across observable factors that may shape the propensity and intensity of conflict.

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

$$SPLIT_i\ (FRAC_i) = a_r + X_i'\Psi + e_i. \tag{1}$$

The dependent variable, $SPLIT_i$, equals one when at least 10% of the historical homeland of an ethnic group i has been partitioned into more than one contemporary state. For robustness, we also show results using the continuous measure of partitioning, $FRAC_i$. X_i is a vector of geographic, ecological, natural resource variables and proxies of pre-colonial conflict and development at the ethnicity level. Appendix Table 2 reports summary statistics for all variables. In all specifications we include region-specific constants (a_r) to account for the somewhat different timing and patterns of colonization.

3.1 Geographical, Ecological, and Natural Resource Measures

Table 2 reports probit (maximum-likelihood) marginal effects, where the dependent variable is the benchmark binary partitioning index ($SPLIT$), and LS estimates with the continuous measure of partitioning ($FRAC_i$) as the dependent variable (The results are similar if we estimate linear probability or logit models).

In Panel A we explore the role of geographic, ecological, and natural resources. Specifications (1)-(2) show that ethnic groups spanning large territories in the pre-colonial period were more likely to be partitioned. This is consistent with the narratives describing the idiosyncratic border drawing. The estimates further show that ethnicities residing in areas with larger water bodies were more likely to find themselves split by the political boundaries. This result is again in accord with the historical evidence that Europeans attempted to use natural barriers while delineating spheres of influence.

In columns (3)-(4) we augment the specification with an index reflecting the average land quality for agriculture and average elevation. We also add the respective standard deviations to proxy for ruggedness and the variance of land quality. Examining the association between partitioning and land characteristics is important because cross-country works and regional studies show that conflict is more prevalent in mountainous terrains (e.g., Fearon and Laitin (2003), Buhaug and Rod (2006), Cervellati, Sunde, and Valmori (2011)). All four geographic features enter with insignificant estimates.[22]

---

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

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

In columns (7)-(8) we include indicators identifying areas with diamond mines and petroleum fields. While in the initial phase of colonization Europeans were mostly interested in agricultural goods and minerals, adding these two indicators allows us to investigate whether partitioned groups differ from non-partitioned ones in terms of natural resources that are correlated with conflict both across and within countries (see Ross (2012) for a thorough review).[23] There are no systematic differences between the two groups of ethnic homelands.

3.2   Pre-colonial Conflict

While at the time of border design Europeans had limited understanding of the local political geography, it is necessary to examine the association between ethnic partitioning and pre-colonial conflict, as recent cross-country works (Fearon and Laitin (2012)) and cross-regional studies in Africa show a significant legacy of conflict from the pre-colonial period to contemporary times (Besley and Reynal-Querol (2012), Nunn and Wantchekon (2011)). In Table 2 -Panel B we thus examine the association between ethnic partitioning and two proxies of pre-colonial conflict, one based on wars between large pre-colonial African kingdoms and one reflecting conflict during the slave trade period.

Using historical data from Brecke (1999) and Cioffi-Revilla (1996), Besley and Reynal-Querol (2012) show that contemporary conflict is higher in regions that suffered from pre-colonial wars (such as the 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). Specifications (1)-(2) show the lack of a systematic association between ethnic partitioning and pre-colonial violence, as reflected by an indicator that takes the value one for ethnic homelands that experienced conflict over the period $1400 - 1700$. Similarly, specifications (3)-(4) show that ethnic partitioning is not related to distance to the centroid of the closest pre-colonial conflict (the results are similar with log distance). In line with the historical accords, these simple specifications reveal that Europeans did not take into account pre-existing animosity between African ethnicities when designing colonial borders. These results show that ethnic partitioning captures source of contemporary conflict distinct to that emphasized by Besley and Reynal-Querol (2012). And while there is some precision error on the exact locations of pre-colonial conflicts, the data have signal, as Besley and Reynal-Querol (2012) do find a robust positive association between

---

the role of land quality estimating various model permutations. Overall the correlation between partitioning and land quality is weak at best; in most specifications the estimate is statistically indistinguishable from zero. Even in the models where some index of land's quality enters with a significant estimate, the economic magnitude is tiny. Moreover, in our analysis of the effect of ethnic partitioning on conflict we are reporting specifications accounting for land quality (and numerous other geographic factors), showing that this has no impact on the coefficient of the partitioning index.

[23]We also repeated estimation using data on "old" diamond mines (discovered before 1970), finding similar results.

pre-colonial conflict and civil wars after African independence.

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

3.3 Pre-colonial Political and Economic Development

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

In columns (6)-(7) we proxy the pre-slave trade level of economic development using an indicator variable that equals one when a city with population exceeding $20,000$ people in 1400 AD was present in the historical homeland and zero otherwise (using data from Chandler (1987)). There is no evidence that ethnicities with historical urban centers were differentially treated during the early stage of colonization when borders were drawn in European capitals. In columns (7)-(8) we regress the partitioning measures on the average distance of each ethnic group to the main European exploration routes. This helps shed light on whether Europeans used the (limited) knowledge they had on local political geography. The coefficient is small and statistically indistinguishable from zero.

3.4 Pre-colonial Ethnic Traits

In Panel D we examine whether other ethnic-specific pre-colonial institutional, cultural, and economic traits correlate with partitioning, using the rich information provided in Murdock's (1967) Ethnographic

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

In columns (1)-(2) we use an ordered variable, ranging from 0 to 7 with higher numbers indicating more complex and more densely populated local communities. There is no association between partitioning and pre-colonial settlement pattern. Building on the insights of unified growth theory (Galor (2011)) that development was higher in areas with intensive use of agriculture in pre-industrial time in (3)-(4) we proxy for economic performance at the time of colonization with an index measuring the importance of agriculture in subsistence, failing again to detect a systematic association. In columns (5)-(8) we investigate whether Europeans took into account the degree of political centralization of the African ethnicities while designing the borders. Following Gennaioli and Rainer (2006, 2007), in (5) and (6) we proxy political centralization with an indicator variable that equals zero when Murdock assigns an ethnicity either as "*stateless*" or "*a petty chiefdom*" (e.g., Xam or the Tiv); and becomes 1 when the ethnicity is part of either a "*large paramount chiefdom*" or a "*large state*" (e.g., Ganda and Zulu). Since institutional and economic development in pre-industrial societies goes in tandem with class stratification (Diamond (1997)), in (7)-(8) we examine the association between ethnic partitioning and a stratification index that ranges from zero, for societies without any class distinctions, to four, for groups with significant class and wealth distinctions. There are no systematic differences between partitioned and non-split ethnicities along these proxies of political centralization.[24]

Summary These results are consistent with the historical narrative on the arbitrary design of African borders. Yet, the results do not imply 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 ethnic homelands and non-split ethnic regions across observable characteristics that may independently affect conflict.[25]

4   Partitioning and Civil Conflict

4.1   Econometric Specification

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

---

[24] In the Supplementary Appendix we provide further evidence on the lack of systemic association between ethnic partitioning and various other measures of pre-colonial societal traits.

[25] Note that the explanatory power of the models is poor. Mc Fadden's pseudo-R^2 (that compares the log likelihood value of the constant-only model with that of the full specification) is low across all permutations, at most 0.07. The probit specifications perform quite poorly in predicting which ethnicities have been partitioned. For example, the specification with all the geographical, ecological, and natural resource measures in Table 2-Panel A (not reported) predicts correctly $(G(X_i'\Psi + Z_i'\Theta + a_j) > 0.5)$ only 29 out of the 230 partitions with the benchmark index.

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

The dependent variable, $y_{i,c}$, reflects civil conflict in the historical homeland of ethnic group i in country c. In the country-fixed-effects specifications (with a_c), each partition of group i is assigned to the corresponding country c. For example, conflict in the part of the Lobi in Ivory Coast is assigned to Ivory Coast, while conflict in Lobi's homeland in Burkina Faso is assigned to Burkina Faso. The coefficient γ on $SPLIT$ captures the effect of ethnic partitioning on civil conflict. Given the lack of systemic correlation between the partitioning index and various historical, ecological, and geographical variables that previous works show that correlate with conflict the coefficient captures the local average treatment effect of ethnic partitioning. Vector $X'_{i,c}$ includes geographical, ecological, and other controls at the country-ethnicity level.

As the dependent variable is a count, we estimate negative binomial models with maximum likelihood (Wooldridge (2002)).[26] The negative binomial model accounts for the many zeros, as well as for the fact that there are a few extreme observations in the right tail of the distribution of the dependent variable. For robustness we also report log-linear LS specifications taking the log of one plus the respective civil conflict measure as the dependent variable.[27] To further account for outliers, we report specifications excluding homelands hosting capital cities or groups where the dependent variable exceeds the top 1%. In all specifications we cluster standard errors at the country level and at the ethnic-family level using method of Cameron, Gelbach, and Miller (2011) that accounts for spatial correlation and arbitrary residual correlation within each country and within each ethnic family.[28]

4.2   Cross-Sectional Estimates

We start our analysis estimating the relationship between partitioning and civil conflict across the 826 ethnic homelands, without conditioning on country fixed effects. Table 3 reports the results. While these estimates have many shortcomings, it is useful examining the cross-sectional patterns before moving to the country fixed effect results. In column (1) we simply control for log population at the ethnic homeland level using the first post-independence census (for most countries in the 1960s or 1970s), the log of land area, and the log of area under water, the only variables found to correlate with partitioning. The coefficient on the partitioning index is positive (0.76) and highly significant. Adding region constants (in (2)) has little effect on the estimate. In (3) we augment the specification with the distance to (i) the national border,

---

[26]Due to overdispersion in the dependent variable, specification tests reject the Poisson, favoring the negative binomial model. For example across all specifications in Tables $3-4$ the x^2 value 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 is preferred.

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

[28]Cameron, Gelbach, and Miller (2011) explicitly cite spatial correlation as an application of the multi-way clustering method. See Spolaore and Wacziarg (2009) and Nunn and Wantchekon (2011) for analogous applications of the multi-way clustering method in accounting for spatial correlation. Murdock (1959) assigns the 834 ethnic groups into 96 ethnolinguistic clusters/families. In the Sensitivity Analysis Section we also report spatial lag specifications, finding similar results. We also estimated standard errors using the method of Conley (1999) to account for spatial dependence of an unknown form, finding similar (and if anything less conservative) errors.

(ii) the sea coast, and (iii) the capital city. We also include an indicator for homelands where capitals fall. Overall, distance to the sea enters with a positive and significant estimate suggesting that there is less conflict in areas closer to the coast. Distance to the capital enters with a positive estimate suggesting that there is more conflict in regions further from the capitals, though the coefficient is not always significant. Distance to the border enters with a negative though insignificant coefficient. As violent events against civilians, riots, and protests often take place in the capitals, the capital city indicator enters with a positive and highly significant coefficient. In spite of the inclusion of these significant covariates, the partitioning indicator drops only slightly (0.674) and retains significance at the 99% level.

Column (4) includes controls reflecting geography (land quality for agriculture, elevation, malaria) and natural resources (indicators for diamond mines and oil deposits). The magnitude 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 ethnic regions where capitals fall. This has little effect on the ethnic partitioning index. The most conservative estimate implies that partitioned ethnicities experience an increase of approximately 162 log points in the number of civil conflict incidents. This translates into an 85% increase in civil conflict activity ($exp(0.62) - 1 = 0.85$) in areas where partitioned ethnicities reside (as compared to the homelands of non-split ethnicities). The effect of ethnic partitioning on civil conflict is quantitatively as strong as the effect of the petroleum indicator that enters with a positive and significant coefficient.

In (7)-(12) we restrict estimation to ethnic areas close to the national border. This allows us to compare civil conflict intensity between partitioned ethnicities and other at-the-border ethnic groups that were not directly affected by the artificial border design. We now have a more balanced sample with 213 partitioned ethnicities and 200 non-split groups.[29] Across all permutations the coefficient on the partitioning index is positive and highly significant, reassuring that our estimates in the full sample are not capturing an overall border effect (which however by itself could reflect the impact of ethic split).

4.3   Within-Country Analysis

Baseline Estimates   In Table 4 we report our baseline country-fixed-effects specifications associating civil conflict across ethnicity-country homelands with partitioning. Columns (1)-(6) report estimates in the full sample, while columns (7)-(12) present results across homelands whose centroid is close to the national border (using the median value of distance to the national border, which at the ethnicity-country homeland level is 61 $km.$). The coefficient on the ethnic partitioning index in (1) and (2) is positive and more than two standard errors larger than zero. The estimate in (2) implies that on average civil conflict intensity is higher in homelands of partitioned groups by approximately 60% ($exp(0.47) - 1 = 0.60$). In column (3) we control for distance to the national border, distance to the sea coast, distance to the capital,

---

[29]When we restrict estimation to homelands where the centroid falls within the median value of distance to the national border, we lose information from 17 split groups, because the centroid of the partitioned homeland is more than 103 kilometers from the border. The results are intact if we include these groups in the estimation.

and the capital city dummy. The coefficient, if anything, increases in absolute value, and becomes more precisely estimated. Conditioning on the rich set of controls and accounting for outliers either by excluding observations where capitals fall or by dropping areas where the dependent variable exceeds the top 1% has no effect on the estimated magnitude. In columns $(7) - (12)$ we restrict estimation across ethnic areas that are close to the national border. Across all specifications the coefficient on ethnic partitioning is positive (around 0.85) and highly significant. Although the coefficient in areas close to the national border is somewhat larger compared to the estimate in the full sample, a Hausman-Chow type test reveals that these differences are not statistically significant.

Ethnic Partitioning and Type of Conflict In an effort to shed some light on the mechanism at work, in Table 5 we examine the effect of ethnic partitioning on the different types of conflict.[30] We start our analysis focusing on battles between armed groups. The coefficients on ethnic partitioning in (1)-(2) imply that fighting between government forces, militias, and rebel groups is more pervasive in the historical homelands of partitioned groups; the estimate suggests that on average partitioned groups experience approximately 80% ($exp(0.60) - 1 = 0.82$) more battles as compared to non-split ethnic groups. Limiting our focus to ethnic areas close to the national border has little effect on the estimate.

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

Another interesting dimension of the ACLED data is the reporting of incidents, mostly battles, between government forces, rebel groups, and militias resulting in territorial changes. Examining instances of territorial changes is useful as they reflect -almost by definition- the government's lack of monopoly of violence. In columns (5), (6), (11), and (12) we report linear probability models associating partitioning with the likelihood that a change in territorial control occurs in an ethnic homeland. The estimates show that partitioned ethnic homelands are more likely to swing between different control groups. The coefficient in (6) and (12) implies that there is a $4\% - 5\%$ higher likelihood that a battle resulting in a change of territorial control occurs in the homeland of a partitioned ethnicity. This effect is far from being small, as in the country-ethnic homeland sample, the overall likelihood that a territorial change takes place is around 20%. The results in Table 5 are in accord with the arguments put forward by the African historiography suggesting that areas of partitioned ethnic groups are tightly contested between the government, ethnic

---

[30] In the Supplementary Appendix we report analogous cross-sectional specifications at the ethnic homeland level.

militias, and rebel groups.[31]

LS Specifications In Table 6, columns (1)-(3) and (6)-(8) we report LS specifications using the natural log of one plus the total number of conflict incidents as the dependent variable. The estimate on the partitioning index is positive and highly significant both in the full sample and when we restrict estimation to areas close to the border.[32] The estimates in the full sample imply that conflict is approximately 20% higher in the homelands of split groups, as compared to groups that were not split by the border. Columns (4) and (9) report linear probability models where the dependent variable is a dummy identifying areas that have experienced some civil conflict. While by solely looking at the "extensive" margin, we do not exploit the richness of the data, we account for the non-linear nature of the dependent variable. Moreover, these specifications shed light on the margin that ethnic partitioning matters. The estimate implies that there is an 8% higher likelihood that a partitioned group will suffer some conflict. We also estimated linear probability models using as the cut-off the median number of conflicts; as the median is 2, these specifications assign regions that experienced just one incident in the group of ethnicities with no conflict. (So this also accounts for potential misclassification and noise). The highly significant coefficient on the partitioning index implies that compared to non-split groups, the likelihood that a partitioned ethnic homeland will experience more than one conflict incident is 10% higher.

Counterfactual We performed an out-of-sample exercise to get a rough estimate of the overall impact of partitioning in Africa, under the (heroic) assumptions that borders do not split any group and that there are no spillovers. We first estimate the baseline model (without and with country fixed effects) in the group of non-split groups (680 country-ethnicity observations). Second, we obtained out-of-sample predictions of conflict for the group of partitioned ethnicities (502 country-ethnicity observations). Third, we compared the predictions with the actual number of incidents. The forecast model -that does not take into account the role of partitioning- predicts approximately 4, 300 incidents in the simple specification and 6, 265 incidents when we include country fixed effects. Yet in the homelands of split groups we observe 11, 731 incidents, implying that at least 5, 500 (if not 7, 500) incidents may be attributed to ethnic partitioning. Since overall we observe 43, 271 incidents, this calculation suggests that conflict would be 12.5% to 17.5% lower in the absence of partitioning. When we perform the out-of-sample exercise with the linear probability specification, the model predicts that 55% of partitioned ethnic groups would have experienced at least one conflict. In practice, however, 66.5% of country-ethnicity homelands experienced conflict. Thus ethnic partitioning explains approximately 10% of overall conflict incidence.

---

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

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

4.4 Example: Conflict in East-Central Africa

East-Central Africa, one of the most conflict-prone regions in the world, offers an illustration of our results. Let us start from Tanzania, a country with little conflict; in the 69 ethnic regions of Tanzania there have been 175 incidents over the period $1997 - 2010$. The mean (median) conflict per ethnic homeland is 2.5 (0). Most conflict (19 incidents) occurs at the border with Rwanda where the partitioned Rundi tribes reside. While in the Rundi homeland only 1.3% of Tanzania's total population lives (in both 1960 and in 2000), the share of conflict is 27.5%. Conflict also appears in the nearby homeland of the (ethnically similar to the Rundi) Ha, where militias based in nearby Rwanda and Burundi raid against the civilian population. The share of total conflict occurring in the Ha homeland is 27.5% (19 incidents) while the share of Tanzania's population living in this region is 4.5%. Interestingly, there is zero conflict in the non-split homelands of the Bende and the Fipa, although both groups reside at the border with the Democratic Republic of Congo, the country with the highest conflict intensity in Africa. This is because lake Tanganyika serves as the natural border of the two countries.

Focusing now on the northern border of Tanzania with Kenya, there is some conflict (10 incidents, 15%) in the homeland of the partitioned Maasai, while the fraction of Tanzania's population living in Maasai homeland is just 3%. Going over the history reveals that both the Tanzanian and the Kenyan governments have consistently tried to suffocate this semi-nomadic tribe by confiscating their land, destroying and burning their villages, and trying to move them outside their ancestral homeland. In contrast, there is no conflict in the Eastern part of the Tanzania-Kenya border, where the non-split Pare group resides.

Focusing now on the Democratic Republic of Congo (DRC), there are $4,333$ conflict events across the 102 ethnic regions (mean=26; median=4). In the homelands of the three partitioned Ruanda ethnicities (of the Interlacustrine Bantu - Rwanda family) we observe 946 incidents (more than 20% of all conflict), while the share of population residing in the homelands of the Hunde, the Konjo, and the Rwanda is around 6%; and in the two adjacent non-split, but ethnically similar Rwanda groups of the Hunde and Toro we have 409 and 27 events, respectively. So, despite the small share of DRC's population living in the Hunde homeland (is around 2.2%), we observe close to 10% of conflict. While the origins of conflict in Eastern Congo are complex and by no means solely driven by partitioning, it involves ethnic militias (such as the FDLR) constantly moving across the border between Rwanda, Burundi, and Uganda to re-group, re-organize and re-arm.[33] In contrast, there is no conflict at the (non-split) Holoholo ethnic homeland on the Western (Zairian) bank of the Tanganyika lake; yet in the adjacent to the Holoholo homeland in the South, in the Ta(a)bwa region we observed 91 conflict events (2%), while the share of DRC's population in

---

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

this region is just 0.8%. This is not surprising since, although the Eastern border of Congo with Tanzania is organic (Tanganyika lake), the Southern one with Zambia follows a straight (latitudinal) line that splits the homeland of the Ta(a)bwa almost equally between D.R.C. and Zambia.
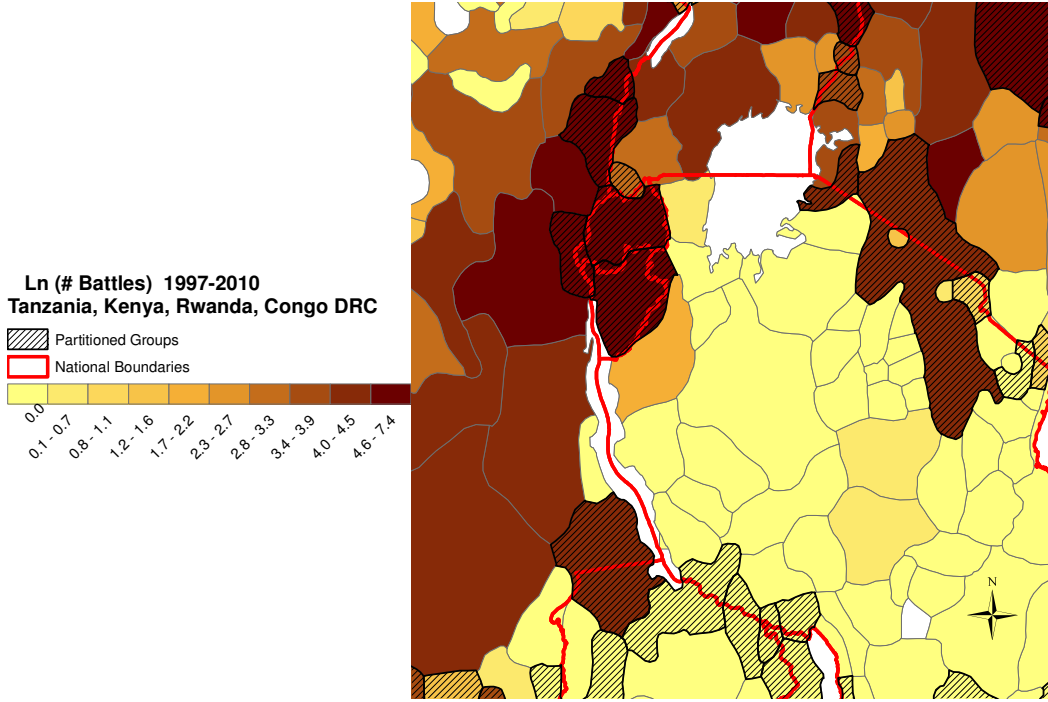

Figure 3

So, while conflict in Eastern-Central Africa has multiple, complex, and inter-linked origins, border artificiality seems to be part of the explanation. First, most militias move across the border as their co-ethnics on the other side offer shelter or/and military assistance. Second, partitioned ethnic groups are used by governments in nearby countries to intervene in other nations. Third, conflict gets attenuated in areas with organic borders, showing *inter alia* that ethnic partitioning is related to conflict.

4.5 Further Evidence and Sensitivity Analysis

Role of Unobservables In Table 7, columns (1)-(4) we augment the specification with a cubic polynomial in latitude and longitude of the centroid of an ethnic group in each country (see Dell (2010)).[34] This allows accounting for unobservable characteristics that vary smoothly in space. Not only the coefficient on the partitioning index remains virtually unaffected, but it retains significance at the 99% confidence level. In columns (5)-(8) we include ethnic-family fixed effects to account both for local conditions and broad cultural, institutional, and other hard-to-observe ethnic-family factors. This robustness check is particularly interesting, as recent works show that ethnic-specific factors, related to history, culture, the type of political and economic organization, and genetics affect crucially development (see Nunn (2012)

---

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

and Spolaore and Wacziarg (2013)). Examples of ethnic families include the Bedouin Arabs, the Tuareg, and the Southwestern Bantu. The estimates suggest that, even when we solely examine within-country, within-ethnic-family variation, civil conflict is significantly more pervasive in border areas belonging to partitioned ethnicities. In columns (9)-(12) we report specifications with both ethnic-family fixed effects and the third-order polynomial in latitude and longitude; while we may be over-fitting, the coefficient on the ethnic partitioning index retains its significance. The fact that the coefficient on the index remains stable across these restrictive specifications implies that it is highly unlikely that unobservable features -correlated with both partitioning and conflict- can explain the strong negative association (see Altonji, Elder, and Taber (2005)).

Regional Economic Development Given the significant negative association between conflict and economic development (see Collier and Hoeffler (2007) and Blattman and Miguel (2010) for reviews and Miguel, Satyanath, and Sergenti (2004) for a seminal contribution on identification), it is useful assessing whether ethnic partitioning affects conflict conditional on income or whether the partitioning-conflict nexus operates via income. In Table 8 we report specifications associating battles, violence against civilians, and territorial change of control with ethnic partitioning, conditioning on the log of per capita GDP at the ethnic homeland level using data from the G-Econ project (Nordhaus and Chen (2009)). While in some specifications income enters with a significant negative coefficient, this has no effect on the estimate on the split indicator that retains its significance. The coefficient on the ethnic partitioning index is almost identical to the analogous estimates in Table 5. Moreover, ethnic partitioning itself is not systematically linked to regional income (results not shown).[35] This set of results, therefore, suggests that the effect of ethnic partitioning on conflict does not operate through an *overall* decline in economic performance of partitioned areas. Nevertheless, as we show below the non-effect of partitioning on average regional development masks considerable heterogeneity with individuals ethnically identifying with split groups underperforming economically.

Other Robustness Checks We have performed numerous perturbations to investigate the robustness our results. For brevity we report and discuss these sensitivity checks in the on-line Supplementary Appendix. Specifically, we show that results are similar when: (1) We focus on (relatively more disruptive) conflicts that have resulted in death casualties. (2) We use events reported solely by international media and the United Nations. (3) We use the continuous index of partitioning ($FRAC$). (4) We use a binary index of partitioning using a 5% threshold to identify split groups. (5) To account for different colonial and post-independence policies we drop iteratively ethnic homelands from each of the five main African regions. (6) We estimate spatial models that account for spill-overs and for factors in nearby ethnic regions.

---

[35]We also estimated specifications proxying regional income with satellite images on light density at night (as in Michalopoulos and Papaioannou (2013a,b)) at night finding similar results. Overall ethnic partitioning is negatively associated with luminosity at the country-ethnic homeland level, though the correlation is insignificant in most (though not all) specifications.

(7) We account for conflict spillovers from regions in the same country or/and the same ethnolinguistic cluster. (8) We control for the historical legacy of violence in the pre-colonial period and the slave trade period. (9) We exclude from estimation ethnic homelands that have been affected by pre-colonial conflict. (10) We control for contemporary (log) population density. (11) We repeat the analysis using *Ethnologue's* database that portrays the spatial distribution of linguistic groups in the mid 1990s.

5 Micro-level Evidence

5.1 Motivation

The evidence so far suggests that conflict and political violence are concentrated in the historical homelands of partitioned groups. This pattern is quite robust to an array of sensitivity checks and it operates even when we condition on proxies of *average* regional income. Nevertheless, the narrative of the African historiography suggests that in several instances split ethnicities face discrimination, as the central government tries to suffocate them, by relocation, property confiscation, etc. (e.g., Herbst (2000), Asiwaju (1985)). To assess the validity of this channel in this Section we move beyond the location-based ethnicity-country level of analysis and exploit individual-level variation in economic conditions and ethnic affiliation.

We employ micro-level data from the Demographic and Health Surveys (DHS) to explore how individuals belonging to partitioned groups fare across various dimensions compared to other citizens in the same country (and even in the same region). Exploiting individual-level information has some straightforward advantages that nicely complement the analysis at the ethnic homeland level. First, we can directly test whether people identifying with split ethnicities under-perform compared to individuals from non-split groups using direct measures of well-being and self-reported ethnic affiliations. Second, we can control for a host of individual level characteristics, so as to better isolate the effect of ethnic partitioning. Third, since we observe people residing both in their group's historical homeland and outside, we can assess whether ethnic partitioning has negative repercussions for all residents of partitioned homelands (i.e., irrespective of their ethnic background) or whether the negative effects are concentrated on individuals identifying with partitioned groups; if the latter is true then this is consistent with ethnic-based discrimination. Fourth, we can shed light on persistence by quantifying the effects of partitioning across different cohorts.

5.2 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 country surveys with gps (global positioning system) coordinates and information on ethnic identification. Our sample comprises 20 countries and covers $86,582$ male respondents.[36] We focus on three outcome variables. First, we use the DHS composite wealth

---

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

index (range from 0 to 5) that reflects household's access to basic public goods (such as electricity, toilet facilities, clean piped drinking water) and economic status, including ownership of selected assets (such as television and radio). Second, we use an ordered education variable that assigns the respondent's level of formal schooling into six categories; no education, incomplete primary, complete primary, incomplete secondary, complete secondary and higher. Third, we use a binary index of literacy. The Data Appendix gives detailed variable definitions, while the Supplementary Appendix reports summary statistics.

Our empirical specification reads:

$$y_{i,e,r,c} = a_c + \beta Split_e + X'_{i,e,r,c}\Phi + Z'_{i,r,c}\Gamma + \zeta_{i,e,r,c}. \tag{3}$$

The dependent variable, $y_{i,e,r,c}$, reflects economic conditions, education and literacy at the individual level; each individual i belongs to (self-identifies with) an 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 (10) five-year age-brackets fixed effects, 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). All specifications include country-specific constants, a_i, that account for factors common to all individuals in a given country that may affect economic development. Moreover, the country constants capture survey differences across countries.[37] $Split_e$, the variable of interest, is an indicator that takes on the value of one if individual i identifies with an ethnicity that has been split across different countries. To construct this dummy variable we link the ethnic information from the DHS to the ethnic groups in Murdock's map.

5.3 Results

5.3.1 Baseline Estimates

Table 9 columns (1)-(6) report the baseline country fixed effects estimates linking the composite DHS wealth index with ethnic partitioning.[38] The coefficient on the partitioning index in column (1) that only includes a set of country constants is negative and highly significant, implying that individuals that 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 sea coast. We also include an indicator for households residing in capital cities ("location controls").[39] In spite of the inclusion of these significant covariates, the

---

2007, and Zambia in 2007.

[37] For example, the DHS manual explicitly states that the composite wealth index is meaningful only for comparison across individuals in the same country; moreover, the various surveys were conducted in different years.

[38] Following Nunn and Wantchekon (2011) to account for spatial correlation and the fact that the split indicator takes on the same value for split and non-split groups, respectively, we cluster standard errors at the ethnicity level and the ethnic homeland (regional/level). Conley standard errors are similar.

[39] 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 suggesting that the provision of public goods is lower as one approaches the national border. The capital city dummy enters with a positive coefficient, implying that living conditions and education are higher for households residing in the capitals.

coefficient on the ethnic partitioning index retains significance at standard confidence levels. The estimate implies that the composite wealth index is -0.21 points lower (approximately 0.2 standard deviations) 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 0.08, 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 empirical approach and micro-level data from the 2005 Afrobarometer Surveys. Another way to assess the economic importance of ethnic partitioning on individual living conditions is to compare its magnitude with another significant covariate. The "standardized" beta coefficient on the capital city indicator -that is a highly significant correlate of the composite wealth index- is around 0.1; this implies that the impact of ethnic partitioning is comparable to that of residing at 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 in our sample (80 kilometers). In all specifications the coefficient on the ethnic partitioning index is negative and significant at the 99% confidence level. Moreover, the estimate is similar to the analogous estimates in the full sample. In columns (7)-(12) we report the results with education as the dependent variable.[40] The estimates imply that -conditional on the location of each respondent and various individual characteristics- people from partitioned ethnic groups are on average less educated than individuals from non-split groups. The standardized "beta" coefficient on the ethnic partitioning index in these specifications is around 0.05, similar to the "beta" coefficient on the capital dummy. The results are similar when we use literacy as the dependent variable (see Supplementary Appendix).

5.3.2 Channels

These estimates reveal that partitioning has negative repercussions on the well-being of the descendants of split ethnicities. Our estimates in Section 4 showed that conflict is concentrated in the homelands of partitioned ethnicities. The natural question that emerges from this set of results is whether ethnic partitioning decreased living conditions and education across split homelands (i.e., residents in split homelands are poorer, less educated and have lower access to public goods irrespective of their ethnic affiliation) or whether it is the individuals belonging to split ethnicities that experience disproportionately lower standards of living. The African historiography stresses that it is individuals belonging to split groups that suffer the consequences of partitioning, so we now explore this discrimination-related hypothesis.

In the DHS we observe individuals from partitioned and from non-split ethnicities residing both inside and outside their ancestral homelands. Specifically, close to 60% (22, 306 out of a total of 37, 831) of individuals from partitioned ethnic groups reside in their historical homeland ("indigenous"), while 40% are found outside their ancestral homeland ("movers"). Table 10 thus reports estimates where we augment the

---

[40]For clarity and to minimize the "incidental parameter" problem we report OLS estimates. Ordered probit and ordered logit estimates that explicitly account for the nature of the outcome variable yield similar estimates.

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 (and zero otherwise). Doing so allows disentangling the importance of ethnically identifying with a split group from the role of residing in the homeland of a partitioned ethnicity. Note that for individuals residing in their ancestral homelands the two partitioning indexes (origin-based and location-based) coincide. Since most individuals (61%) reside in their historical homeland, the two variables are positively correlated (0.40). Thus in these specifications we obtain identification from "movers" (i.e., those living in a location different from their ancestors). In column (1) both the origin-based split indicator and the location-based one enter with negative and significant estimates (−0.31 and −0.24, respectively). In (2) we control for individual characteristics. This has little effect on the estimates of both partitioning indicators. In column (3) we augment the specification with the distance of each household to the capital city, the sea coast, and the national border, as well as the capital dummy. The coefficient on the location-based measure of partitioning drops and becomes statistically indistinguishable from zero; in contrast, the estimate on the ethnicity-based partitioning index retains its statistical and economic significance.

In columns (4)-(6) we restrict estimation to households residing close to the national borders. The coefficient on identity-based ethnic partitioning remains negative and is quite stable (ranging from −0.31 to −0.35); this suggests that even when we focus on border areas and control for numerous individual and location features, households belonging to partitioned ethnic groups have on average worse living conditions and lower access to utility services as compared to households identifying with non-partitioned groups. The results are similar when we use the education index as the dependent variable (in columns (7)-(12)). The identity-based partitioning index enters all permutations with a significantly negative estimate (around −0.20), implying that individuals from partitioned ethnicities have on average a lower level of formal education as compared to individuals from non-split ethnicities. In contrast, the location-based index of partitioning is not systematically linked to education.[41]

To sharpen the intuition that ethnic partitioning is deleterious for individuals who ethnically identify with split groups, in Table 11 we restrict estimation to ethnic areas of partitioned groups. The estimate on the ethnic partitioning index in columns (1)-(3) is negative and highly significant implying that households of partitioned ethnicities have lower access to public goods and worse living conditions than households from non-split groups. The results in columns (4)-(6) and (7)-(9) further show that -on average- individuals from split ethnicities have fewer years of education and are more likely to be illiterate.

In line with the African historiography these findings suggest that individuals identifying with split ethnicities have fewer economic opportunities. In an effort to further push on the ethnic discrimination explanation, we repeated estimation exploiting variation solely within each enumeration area (village/town/city). Estimating enumeration-area-fixed effects specifications allows controlling for (observed

---

[41]In the Supplementary Appendix we report similar specifications including a set of ethnic homeland fixed effects, so as to account for homeland-specific features. The coefficient on the identity-based ethnic partitioning index is negative and significant in all specifications.

and unobserved) local factors at an extremely fine level; as such the ethnic partitioning indicator isolates the impact of respondent's ethnic origin. To have reasonable within village/city variation we focus on enumeration areas with 5 or more (interviewed) households (though this has no effect on the estimates). Table 12 reports the within-enumeration-area estimates. These extremely restrictive specifications -that include more than 4,500 city/town constants- assess whether differences in wealth and education across ethnic lines are related to ethnic partitioning among individuals residing in the same location. In spite of the efficiency loss (as most of the variation in both the explanatory variables and the ethnic partitioning index is absorbed by the village fixed effects) the coefficient on the ethnic partitioning indicator is negative and significant with all three outcome variables. These estimates point out that even when we compare respondents currently residing in the same village, those of a partitioned ethnic background have lower access to utility services, are less educated, and more likely to be illiterate.

5.3.3 Persistence

The results linking partitioning to conflict and education/public goods point out that the Scramble for Africa has had sizable long-run effects on economic development. Since the early/mid-nineties, however, many African countries have made considerable steps towards democratic institutions, while over the past decade growth has been high across the continent. Hence, it is quite important especially from a policy standpoint to assess whether the recent economic and political modernization efforts have been associated with convergence in economic conditions between split and non-split groups.

To shed light on this question we exploit the fact that the DHS report respondents' age cohort and repeat estimation separately for "old" and "young" cohorts. Table 13 reports the findings. For brevity we report results only with the composite wealth index. In odd-numbered columns we restrict estimation to "young" respondents, while in even-numbered columns we restrict estimation to "old" respondents; as a cut-off value we use those born before and after 1977 which is the median date of birth across respondents (the pattern is similar when we use 1975, 1980, or 1985). The results suggest that the negative effect of ethnic partitioning is quite strong both across "young" and "old" cohorts. Moreover, the estimate on the ethnic partitioning index is almost identical in the two sub-samples, revealing that the adverse effects of partitioning are quite persistent.

6 Conclusion

This study examines the consequences of a neglected aspect of colonization, the artificial drawing of political boundaries among European powers in the end of the 19th century, which in the eve of African independence led to the partitioning of several ethnicities across the 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 with various geographic, ecological, and ethnic-specific pre-colonial characteristics. With the sole

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

Second, we examine the effect of ethnic partitioning on civil conflict, as this has been hypothesized to be the major consequence of the scramble for Africa. Our analysis is based on regional data spanning the universe of African ethnic areas. We exploit a new geocoded dataset that reports information on more than $43,000$ conflict events over the period $1997 - 2010$. The database is quite useful in examining the long-run impact of ethnic partitioning, as it reports both the precise location of battles between government forces, militias, and rebel groups, as well as the incidents involving violence against civilians. This is key as due to data limitations most previous works have not examined this quite damaging and economically important aspect of conflict. Exploiting within-country variation and focusing on ethnic homelands is appropriate for Africa since ethnicity is salient and conflicts have usually an ethnic dimension. We find that battles between armed groups, as well as violence against the civilian population are concentrated in the homelands of partitioned ethnicities. Moreover, these territories are intensely contested among the different conflict actors resulting in a significantly higher likelihood of a change in territorial control. These results are robust to different estimation techniques, alternative classifications of partitioned ethnicities, and accounting for ethnic family features.

Third, using data from the Demographic and Health Surveys spanning more than $85,000$ respondents in 20 African countries we show that individuals identifying with partitioned groups have fewer household assets, poorer access to utilities, and worse educational outcomes, as compared to individuals from non-split ethnicities. This pattern is not due to a generalized decline in standards of living of households residing in split homelands; rather it is driven by the poorer economic performance of individuals of split ethnic identification. We obtain the same results exploiting within-village variation. Finally, the empirical analysis reveals that the repercussions of ethnic partitioning do not attenuate over time shedding light on the persistent legacy of the scramble for Africa on increased political violence and underdevelopment.

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 economic performance. 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.

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., AND J. A. ROBINSON (2012): *Why Nations Fail? The Origins of Power, Prosperity, and Poverty*. Crown Publishers, New York, NY.

ALESINA, A., A. DEVLEESCHAUWER, W. EASTERLY, S. KURLAT, AND R. WACZIARG (2003): "Fractionalization," *Journal of Economic Growth*, 8(2), 155–194.

ALESINA, A., W. EASTERLY, AND J. MATUSZESKI (2011): "Artificial States," *Journal of the European Economic Association*, 9(2), 246–277.

ALTONJI, J. G., T. E. ELDER, AND C. R. TABER (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Jounal of Political Economy*, 113(1), 151–184.

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.

BESLEY, T., AND T. PERSSON (2011): *Pillars of Prosperity. The Political Economics of Development Clusters*. Princeton University Press, Princeton, NJ.

BESLEY, T., AND M. REYNAL-QUEROL (2012): "The Legacy of Historical Conflicts. Evidence from Africa," mimeo, London School of Economics and UPF.

BLATTMAN, C., AND E. MIGUEL (2010): "Civil War," *Journal of Economic Literature*, 48(1), 3–57.

BOSKER, M., AND J. DE REE (2010): "Ethnicity and the Spread of Civil War," CEPR Discussion Paper No. 8055.

BRANCATI, D. M. (2006): "Decentralization: Fueling the Fire or Dampening the Flames of Ethnic Conflict and Secessionism?," *International Organization*, 60(2), 651–685.

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

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

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

CERVELLATI, M., U. SUNDE, AND S. VALMORI (2011): "Disease Environment and Civil Conflict," IZA Discussion Paper No. 5614.

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

CHAPMAN, T., AND P. G. ROEDER (2007): "Partition as a Solution to Wars of Nationalism: The Importance of Institutions," *American Political Science Review*, 101(4), 677–691.

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

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

COLLIER, P. (2007): *The Bottom Billion*. Oxford 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.

DELL, M. (2010): "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 78(6), 839–857.

DIAMOND, J. (1997): *Guns, Germs, and Steel. The Fates of Human Societies*. W.W. Norton Co, New York, NY.

DJANKOV, S., AND M. REYNAL-QUEROL (2010): "The Causes of Civil Wars," mimeo UPF.

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

EASTERLY, W., AND R. LEVINE (2012): "The European Origins of Economic Development," NBER Working Paper No. 18162.

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

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

ESTEBAN, J., L. MAYORAL, AND D. RAY (2012): "Ethnicity and Conflict: An Empirical Study," *American Economic Review*, 102(4), 1310–1342.

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

FEARON, J. D. (2004): "Separatist Wars, Partition, and World Order," *Security Studies*, 13(4), 394–415.

FEARON, J. D., AND D. D. LAITIN (2000): "Violence and the Social Construction of Ethnic Identity," *International Organization*, 54(4), 845–877.

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

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

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

GALOR, O. (2011): *Unified Growth Theory*. Princeton University Press, Princeton, NJ.

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

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

GLAESER, E. L., R. LA PORTA, F. L. DE SILANES, AND A. SHLEIFER (2004): "Do Institutions Cause Growth?," *Journal of Economic Growth*, 9(3), 271–303.

HARARI, M., AND E. L. FERRARA (2012): "Conflict, Climate and Cells: A Disaggregated Analysis," mimeo MIT and Bocconi University.

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

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

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

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

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

JEDWAB, R., AND A. MORADI (2012): "Colonial Investments and Long-Term Development in Africa: Evidence from Ghanaian Railways," mimeo George Washington University and Sussex University.

KAUFMAN, C. (1996): "Possible and Impossible Solutions to Ethnic Civil Wars," *International Security*, 20(4), 133–175.

———— (1998): "When All Else Fails: Separation as a Remedy for Ethnic Conflicts, Ethnic Partitions and Population Transfers in the Twentieth Century," *International Security*, 23(2), 120–156.

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.

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 (2013a): "National Institutions and Sub-national Development in Africa," Quarterly Journal of Economics, forthcoming.

———— (2013b): "Pre-colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, 81(1), 113–152.

MIGUEL, E., S. SATYANATH, AND E. SERGENTI (2004): "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," *Journal of Political Economy*, 112(4), 725–753.

MONTALVO, J.-G., AND M. REYNAL-QUEROL (2005): "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review*, 95(3), 796–816.

MURDOCK, G. P. (1959): *Africa: Its Peoples and their Culture History.* McGraw-Hill, New York, NY.

NORDHAUS, W. D., AND X. CHEN (2009): "Geography: Graphics and Economics," *The B.E. Journal of Economic Analysis and Policy, Contributions*, 9(2), article 1.

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

——— (2012): "Culture and the Historical Process,," *Economic History of Developing Regions*, 27(S1), 108–126.

——— (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.

PRUNIER, G. (2009): *Africa's World War. Congo, the Rwandan Genocide, and the Making of a Continental Catastrophe*. Oxford University Press, Oxford, United Kingdom.

RALEIGH, C., D. CUNNINGHAM, L. WILHELMSEN, AND N. P. GLEDITSCH (2006): "PRIO Conflict Sites Codebook, 1946-2005," Centre for the Study of Civil War, Peace Research Institute in Oslo.

RALEIGH, C., A. LINKE, AND C. DOWD (2012a): "Armed Conflict Location and Event Dataset (ACLED) Codebook Version 2," Working Paper, Centre for the Study of Civil War, International Peace Research Institute, Oslo (PRIO).

——— (2012b): "Armed Conflict Location and Event Dataset (ACLED) Codebook Version 2," Working Paper, Centre for the Study of Civil War, International Peace Research Institute, Oslo (PRIO).

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

SAMBANIS, N. (2000): "Partition as a Solution to Ethnic War: An Empirical Critique of the Theoretical Literature," *World Politics*, 52(2), 437–483.

SAMBANIS, N., AND J. SCHULHOFER-WOHL (2009): "What's In A Line? Is Partition the Solution to Civil War?," *International Security*, 34(2), 82–118.

SILVA, J. M. S., AND S. TENREYRO (2006): "The Log of Gravity," *Review o Economics and Statistics*, 88(4), 641–658.

Spolaore, E., and R. Wacziarg (2009): "The Diffusion of Development," *Quarterly Journal of Economics*, 124(2), 469–529.

———— (2013): "How Deep Are the Roots of Economic Development?," *Journal of Economic Literature*, 51(2).

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

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

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

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

7 Data Appendix

7.1 Main Outcome and Control Variables

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

Continuous Measure of Partitioning (FRAC): The index reflects the probability that a square kilometer of an ethnic area falls to a different country than the rest of the historical ethnic homeland. Computed similarly to the Herfindahl index. *Source: Calculated intersecting Murdock's (1959) ethnic map of Africa with the Digital Chart of the World (DCW) shapefile. The latter contains the polygons delineating the international boundaries in 2000.*

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

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

Violence against Civilians: Total number of violent events against civilians at the ethnic homeland (in each country for partitioned ethnicities). Violence against civilians occurs when any armed/violent group attacks unarmed civilians. Rebels, governments, militias, rioters can all commit violence against civilians. We aggregate the data at the ethnic homeland level and at the country-ethnic homeland level. See Section 2 for details. *Source: ACLED.*

Territorial Change of Control: Indicator that takes on the value of one if a battle resulting in change of territorial control takes place at the historical homeland on an ethnic group / country-ethnicity.

Source: ACLED.

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

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

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

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

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

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

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

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

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

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

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

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

Income per capita: Log of per capita income (GDP) in 2000 at the country-ethnic homeland level. *Source: Nordhaus and Chen (2009) G-Econ Database. available at: http://gecon.yale.edu/*

Latitude: Latitude of the centroid of each ethnic group, constructed using ArcGIS Software.

Longitude: Longitude of the centroid of each ethnic group, constructed using ArcGIS Software.

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

7.2 Pre-colonial Features

Distance to Explorer's Routes: The geodesic distance of the centroid of each group to the nearest route of the principal European explorers. *Source: The "Century Atlas, Africa" digitized by Nunn and Wantchekon (2011).*

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

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

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

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

Settlement Pattern: Ordered variable ranging from 0 to 7 quantifying "settlement pattern of each group". 0 indicates fully nomadic (migratory) groups, 1 indicates semi-nomadic, 2 indicates semi-sedentary, 3 identifies groups that live in compact and impermanent settlements, 4 indicates societies those in neighborhoods of dispersed family homes, 5 indicates for groups in separated hamlets forming a single community, 6 indicates societies living in compact and relatively permanent settlements, and 7 denotes the groups residing in complex settlements. *Source: Murdock (1967); variable code v30.*

Political Centralization: The binary index is constructed using Murdock's (1967) Jurisdictional Hierarchy beyond Local Community $0-4$ index that indicates the number of jurisdictional levels (political complexity) in each society above the local level. The political centralization index takes the value 0 if the

Jurisdictional Hierarchy beyond Local Community variable equals 0 or 1 (when the society is classified as either stateless or forming a small chiefdom). The index takes on the value 1 if the Jurisdictional Hierarchy beyond Local Community variable equals 2, 3, and 4 (when the society is classified as being part of large paramount chiefdom or a large state). This aggregation follows Gennaioli and Rainer (2006, 2007). *Source: Murdock (1967).*

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

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

7.3 DHS Data

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

Education: 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/)*

Literacy: A dummy (indicator) variable that takes on the value of one when the household head can perfectly read and understand a simple sentence presented to him by the DHS team in his mother tongue and zero otherwise (i.e. when the individual cannot read at all or when the individual can only partially read the sentence). *Source: Demographic and Health Surveys. http://www.measuredhs.com/)*

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.

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

Panel A: All Ethnic Homelands

	Indicator Likelihood	Number of Incidents		Excluding Capitals		Excluding Outliers (top 1%)	
	mean (1)	mean (2)	median (3)	mean (4)	median (5)	mean (6)	median (7)
All Types of Conflict Events							
all ethnic homelands (N=826)	0.762	52.386	4.000	35.202	4.000	36.584	4.000
non-partitioned ethnic groups (N=596)	0.732	47.616	3.000	30.495	3.000	30.460	3.000
partitioned ethnic groups (N=230)	0.839	64.748	13.000	47.872	10.000	52.297	13.000
difference	**0.11**	**17.13**	**10.00**	**17.38**	**7.00**	**21.84**	**10.00**
difference (p-value)	*(0.00)*	*(0.28)*	*(0.00)*	*(0.02)*	*(0.00)*	*(0.00)*	*(0.00)*
Battles							
all ethnic homelands (N=826)	0.593	22.645	1.000	16.781	1.000	14.920	1.000
non-partitioned ethnic groups (N=596)	0.577	19.837	1.000	14.551	1.000	13.074	1.000
partitioned ethnic groups (N=230)	0.635	29.926	2.500	22.787	2.000	21.123	2.000
difference	**0.06**	**10.09**	**1.50**	**8.24**	**1.00**	**8.05**	**1.00**
difference (p-value)	*(0.13)*	*(0.20)*	*(0.01)*	*(0.05)*	*(0.01)*	*(0.01)*	*(0.01)*
Battles with Change in Territory							
all ethnic homelands (N=826)	0.265	2.810	0.000	2.460	0.000	2.096	0.000
non-partitioned ethnic groups (N=596)	0.227	2.183	0.000	1.752	0.000	1.558	0.000
partitioned ethnic groups (N=230)	0.365	4.448	0.000	4.355	0.000	3.498	0.000
difference	**0.14**	**2.26**	**0.00**	**2.60**	**0.00**	**1.94**	**0.00**
difference (p-value)	*(0.00)*	*(0.01)*	*(0.00)*	*(0.00)*	*(0.00)*	*(0.00)*	*(0.00)*
Violence against Civilians							
all ethnic homelands (N=826)	0.608	19.180	1.000	12.196	1.000	12.227	1.000
non-partitioned ethnic groups (N=596)	0.572	17.648	1.000	10.502	1.000	10.381	1.000
partitioned ethnic groups (N=230)	0.700	23.157	3.000	16.758	2.000	17.004	3.000
difference	**0.13**	**5.51**	**2.00**	**6.26**	**1.00**	**6.62**	**2.00**
difference (p-value)	*(0.00)*	*(0.43)*	*(0.00)*	*(0.03)*	*(0.00)*	*(0.01)*	*(0.00)*

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

Panel B: Ethnic Homelands close to the National Border

	Indicator Likelihood	Number of Incidents		Excluding Capitals		Excluding Outliers	
	mean (1)	mean (2)	median (3)	mean (4)	median (5)	mean (6)	median (7)
All Types of Conflict Events							
all ethnic homelands (N=413)	0.741	49.660	4.000	38.280	3.000	36.643	4.000
non-partitioned ethnic groups (N=200)	0.645	32.300	1.000	29.270	1.000	24.704	1.000
partitioned ethnic groups (N=213)	0.831	65.967	12.000	47.249	10.000	47.957	12.000
difference	**0.19**	**33.67**	**11.00**	**17.98**	**9.00**	**23.25**	**11.00**
difference (p-value)	*(0.00)*	*(0.05)*	*(0.00)*	*(0.10)*	*(0.00)*	*(0.00)*	*(0.00)*
Battles							
all ethnic homelands (N=413)	0.545	23.194	1.000	18.031	1.000	16.186	1.000
non-partitioned ethnic groups (N=200)	0.465	15.700	0.000	13.867	0.000	11.402	0.000
partitioned ethnic groups (N=213)	0.620	30.235	3.000	22.178	2.000	20.724	2.500
difference	**0.15**	**14.53**	**3.00**	**8.31**	**2.00**	**9.32**	**2.50**
difference (p-value)	*(0.00)*	*(0.10)*	*(0.00)*	*(0.17)*	*(0.00)*	*(0.02)*	*(0.00)*
Battles with Change in Territory							
all ethnic homelands (N=413)	0.266	3.281	0.000	2.911	0.000	2.496	0.000
non-partitioned ethnic groups (N=200)	0.170	2.135	0.000	1.633	0.000	1.303	0.000
partitioned ethnic groups (N=213)	0.357	4.357	0.000	4.183	0.000	3.616	0.000
difference	**0.19**	**2.22**	**0.00**	**2.55**	**0.00**	**2.31**	**0.00**
difference (p-value)	*(0.00)*	*(0.04)*	*(0.00)*	*(0.01)*	*(0.00)*	*(0.02)*	*(0.00)*
Violent Indcidents							
all ethnic homelands	0.574	17.910	1.000	13.860	1.000	12.616	1.000
non-partitioned ethnic groups	0.450	11.390	0.000	10.786	0.000	7.525	0.000
partitioned ethnic groups	0.690	24.033	3.000	16.919	2.000	17.393	3.000
difference	**0.24**	**12.64**	**3.00**	**6.13**	**2.00**	**9.87**	**3.00**
difference (p-value)	*(0.00)*	*(0.07)*	*(0.00)*	*(0.14)*	*(0.00)*	*(0.00)*	*(0.00)*

The table reports summary statistics and test of means and medians for the ACLED civil conflict (outcome) variables employed in the empirical analysis at the ethnic homeland level. Panel A reports test of means/medians at the full sample. Panel B reports test of means/medians across ethnic homelands close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland to the national border; 102 kilometers). Column (1) reports the likelihood that a conflict (all conflict incidents, battles, battles resulting in a territorial change, and violence against the civilian population) affects an ethnic homeland. Columns (2)-(3) report the mean and the median value for each type of conflict, respectively. Columns (4)-(5) report the mean and the median value for each type of conflict excluding ethnic regions where capital cities fall. Columns (6)-(7) report the mean and the median value for each type of conflict excluding ethnic regions where the respective variable exceeds the 99th percentile. For each variable the table reports the mean/median value using all ethnic homelands, partitioned ethnicities and non-partitioned ethnicities. The table also reports the mean and median difference and the p-value of mean-median equality between the group of partitioned and non-partitioned ethnicities. The Data Appendix gives detailed variable definitions and data sources.

Table 2 - Border Artificiality

Panel A: Geographical, Ecological and Natural Resources Features

	SPLIT (1)	FRAC (2)	SPLIT (3)	FRAC (4)	SPLIT (5)	FRAC (6)	SPLIT (7)	FRAC (8)
Land Area under Water	0.3219***	0.0613***	0.3891***	0.0683***	0.3449***	0.0670***	0.3352***	0.0623***
	(0.0955)	(0.0125)	(0.1012)	(0.0117)	(0.0953)	(0.0122)	(0.0998)	(0.0134)
Land Area	0.0869	0.0149*	0.1051*	0.0171**	0.0938	0.0167**	0.0697	0.0131*
	(0.0567)	(0.0080)	(0.0600)	(0.0086)	(0.0583)	(0.0084)	(0.0542)	(0.0071)
Elevation			-0.0623	-0.0209				
			(0.1834)	(0.0293)				
St. Dev. Elevation			-0.0001	0.0000				
			(0.0006)	(0.0001)				
Suitability for Agriculture			0.4494	0.0621				
			(0.3328)	(0.0491)				
St. Dev. Suit. Agricult.			0.8556	0.0672				
			(0.7386)	(0.0859)				
Malaria Stability Index					0.1250	0.0292		
					(0.2297)	(0.0409)		
Distance to the Coast					-0.0001	0.0000		
					(0.0002)	(0.0000)		
Diamond Mine Indicator							0.1626	0.018
							(0.1802)	(0.0287)
Oil Indicator							0.0081	0.0026
							(0.1696)	(0.0351)
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.050	__	0.057	__	0.051	__	0.051	__
Adjusted R-squared	__	0.082	__	0.090	__	0.089	__	0.083
Observations	826	826	826	826	826	826	826	826

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

Table 2 - Border Artificiality

Panel B: Pre-colonial Conflcit

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

Table 2 - Panel B reports probit marginal effects (in odd-numbered columns) and OLS estimates (in even-numbered columns) associating ethnic partitioning with proxy measures of pre-colonial conflcit. In odd-numbered specifications, the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary countries. In even-numbered columns, the dependent variable is a continuous index of ethnic partitioning that reflects the probability that a randomly chosen pixel of the historical homeland of an ethnic group falls into a different country. All specifications include a set of region fixed effects (constants not reported). In columns (1)-(4) we use data on pre-colonial conflict (in the period 1400-1700) from Besley and Reynal-Querrol (2012) and Brecke (1999). In columns (5)-(8) we use data on enslavement during the African slave trades from Nunn (2008) and Nunn and Watcekon (2011). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 2 - Border Artificiality

Panel C: Pre-colonial Statehood & Development

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

Table 2 - Panel C reports probit marginal effects (in odd-numbered columns) and OLS estimates (in even-numbered columns) associating ethnic partitioning with geographical, ecological and natural resource variables. In odd-numbered specifications, the dependent variable is an indicator that equals one when at least 10% of the historical ethnic homeland (as portrayed in Murdock's (1959) Ethnolinguistic map) falls to more than one contemporary countries. In even-numbered columns, the dependent variable is a continuous index of ethnic partitioning that reflects the probability that a randomly chosen pixel of the historical homeland of an ethnic group falls into a different country. All specifications include a set of region fixed effects (constants not reported). Data on pre-colonial statehood (large kingdoms and empires) are taken from Besley and Reynal-Querrol (2012) and O' Brian (1999). Data on major cities before the slave trades (in 1400) come from Chandler (1987). Data on the pincipal European explore's routes come from Nunn (2009). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 2 - Border Artificiality

Panel D: Pre-colonial Ethnic Features (using data from Murdock (1967))

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

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

Table 3: Ethnic Partitioning and Civil Conflict
Cross-Sectional Estimates

	All Ethnic Homelands						Ethnic Homelands Close to the National Border					
	All Observations				Excl. Outliers	Excl. Capitals	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SPLIT (Partitioning)	0.7590***	0.7250***	0.6740***	0.6207***	0.7093***	0.6813***	0.7076***	0.7136***	0.9580***	0.8492**	0.9430***	0.8812**
Double-clustered s.e.	(0.2928)	(0.2735)	(0.2338)	(0.2234)	(0.2150)	(0.2317)	(0.2473)	(0.1798)	(0.2979)	(0.3316)	(0.2970)	(0.3608)
Log Likelihood	-3221.79	-3204.03	-3159.39	-3151.76	-3095.09	-2847.00	-1593.4	-1578.59	-1560.3	-1552.09	-1498.35	-1422.98
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	826	826	826	826	821	779	413	413	413	413	408	393

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the ethnic homeland level. The dependent variable is the total number of civil conflict incidents at each ethnic homeland over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. The specifications in columns (2)-(6) and (7)-(12) include a set of region fixed effects (constants not reported). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population approximately in 1960. The set of location controls includes the distance of the centroid of each ethnic homeland from the capital, from the sea coast, from the national border, and an indicator that takes on the value one if a capital city falls in the historical homeland. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The specifications in columns (5) and (11) exclude ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (6) and (12) exclude ethnic homelands where capital cities fall. The specifications in columns (7)-(12) focus on ethnic areas close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland to the national border; 102 kilometers). The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 4: Ethnic Partitioning and Civil Conflict
Country-Fixed-Effects Estimates

	All Ethnicity-Country Homelands						Ethnicity-Country Homelands Close to the National Border					
	All Observations				Excl. Outliers	Excl. Capitals	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SPLIT (Partitioning)	0.4929***	0.4748***	0.6731***	0.6185***	0.6284***	0.6171***	1.0208***	0.8465***	0.9258***	0.8502***	0.8502***	0.8388***
Double-clustered s.e.	(0.1559)	(0.2073)	(0.1977)	(0.1876)	(0.1876)	(0.1829)	(0.1767)	(0.3061)	(0.3143)	(0.3145)	(0.3140)	(0.3154)
Log Likelihood	-3942.45	-3708.6	-3615.94	-3603.19	-3498.16	-3340.02	-1556.06	-1419.28	-1393.15	-1384.94	-1377.37	-1322.22
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	1182	1182	1182	1182	1170	1139	576	576	576	576	575	565

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnicity homeland level. The dependent variable is the total number of civil conflict incidents at each ethnic homeland within a country over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. The specifications in columns (2)-(6) and (7)-(12) include country fixed effects (constants not reported). The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country homeland to the national border; 62 kilometers). The specifications in columns (5) and (11) exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (6) and (12) exclude country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 5: Ethnic Partitioning and Civil Conflict
Country Fixed Effects Estimates with Various Measures of Civil Conflict

	All Ethnicity-Country Homelands						Ethnicity-Country Homelands Close to the National Border					
	Battles		Violence		Territorial Change		Battles		Violence		Territorial Change	
	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SPLIT (Partitioning)	0.6051***	0.5909***	0.4943***	0.5029***	0.0376*	0.0396*	0.8442***	0.8500***	0.8895***	0.8937***	0.0507*	0.0502*
Double-clustered s.e.	(0.1420)	(0.1426)	(0.1621)	(0.1743)	(0.0199)	(0.0209)	(0.2510)	(0.2537)	(0.2439)	(0.2488)	(0.0280)	(0.0293)
Log Likelihood	-2630.41	-2432.38	-2502.75	-2299.75	__	__	-995.072	-946.214	-938.915	-891.698	__	__
Adjusted R-squared	__	__	__	__	0.455	0.446	__	__	__	__	0.466	0.457
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1139	1182	1139	1182	1139	576	565	576	565	590	565

The table reports Negative Binomial Maximum Likelihood (ML) estimates (in columns (1)-(4) and (7)-(10) and linear probability (LS) estimates associating various aspects of civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in columns (1), (2), (7) and (8) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3), (4), (9), and (10) is the number of violent events against civilian populations. The dependent variable in columns (5), (6), (11) and (12) is a dummy variable that equals one if a battle resulting in territorial change of control has taken place and zero otherwise. All specifications include country fixed effects (constants not reported). The specifications in columns (7)-(12) focus on ethnicity-country areas close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland within a country to the national border; 62 kilometers). The specifications in even-numbered columns exclude country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 6: Ethnic Partitioning and Civil Conflict
OLS Specifications

	All Ethnicity-Country Homelands					All Ethnicity-Country Homelands Close to the National Border				
	OLS ln(1+All Civil Conflict Events)			OLS Linear Probabily		OLS ln(1+All Civil Conflict Events)			OLS Linear Probabily	
	All Obs	No Capitals	No Outliers	All Obs	All Obs	All Obs	No Capitals	No Outliers	All Obs	All Obs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SPLIT (Partitioning)	0.1912**	0.1984**	0.1869**	0.0824***	0.0993***	0.3087***	0.3131***	0.3281***	0.0801*	0.1016*
Double-clustered s.e.	(0.0953)	(0.0916)	(0.0908)	(0.0314)	(0.0336)	(0.1179)	(0.1158)	(0.1197)	(0.0459)	(0.0553)
Adjusted R-squared	0.617	0.593	0.576	0.439	0.447	0.603	0.595	0.574	0.465	0.451
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1170	1139	1182	1182	590	575	565	576	576

The table reports OLS estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnicity level. The dependent variable in columns (1)-(3) and (6)-(9) is the log of one plus the total number of civil conflict incidents in an ethnic region within a country over the period 1997-2010. The dependent variable in columns (4) and (8) is a dummy variable that takes on the value of zero if no conflict has taken place during 1997-2010. The dependent variable in columns (5) and (10) is an indicator that equals zero if there have been at most two conflict incidents (2 is the median value of all conflict incidents over the 1997-2010) in an ethnicity-country. The specifications in columns (6)-(10) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country to the national border; 62 kilometers). The specifications in columns (2) and (7) exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (3) and (8) exclude country-ethnic homelands where capital cities fall. All specifications include country fixed effects (constants not reported). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7: Ethnic Partitioning and Civil Conflict
Accounting for Unobservables

	Latitude-Longitude Polynomial				Ethnic Family Fixed Effects				Latitude-Longitude Polynomial & Ethnic Family Fixed Effects			
	All Ethnicity-Country Areas		Border Areas		All Ethnicity-Country Areas		Border Areas		All Ethnicity-Country Areas		Border Areas	
	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals	All Obs	No Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SPLIT (Partitioning)	0.6298***	0.6182***	0.7750**	0.7692**	0.5568***	0.5562**	0.7434**	0.7200*	0.5354**	0.5110**	0.6207*	0.5770*
Double-clustered s.e.	(0.1923)	(0.1941)	(0.3046)	(0.3068)	(0.2182)	(0.2335)	(0.3792)	(0.3814)	(0.1984)	(0.2074)	(0.3346)	(0.3335)
Log Likelihood	-3582	-3317.33	-1354.67	-1293.51	-3467.48	-3209.94	-1279.21	-1219.501	-3451.45	-3195.16	-1269.921	-1210.46
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1139	576	565	1182	1139	576	565	1182	1139	576	565

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnic homeland level. The dependent variable is the total number of civil conflict incidents in each ethnic homeland in each country over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. All specifications include country fixed effects (constants not reported). The specifications in columns (3), (4), (7), (8), (11), and (12) focus on areas close to the national border (using as a cutoff the median distance from the centroid of each country-ethnic homeland to the national border; 62 kilometers). The specifications in even-numbered columns exclude country-ethnic homelands where capital cities fall. The specifications in (1)-(4) and (9)-(12) include a cubic polynomial in latitude and longitude. The specifications in (5)-(12) include a set of ethnic family fixed effects (constants not reported). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 8: Ethnic Partitioning and Civil Conflict, Conditioning on Economic Development
Country Fixed Effects Estimates with Various Measures of Civil Conflict

	All Ethnicity-Country Homelands						Ethnicity-Country Homelands Close to the National Border					
	Battles		Violence		Territorial Change		Battles		Violence		Territorial Change	
	All Obs (1)	No Capitals (2)	All Obs (3)	No Capitals (4)	All Obs (5)	No Capitals (6)	All Obs (7)	No Capitals (8)	All Obs (9)	No Capitals (10)	All Obs (11)	No Capitals (12)
SPLIT (Partitioning)	0.6050***	0.6001***	0.5054***	0.5226***	0.0329*	0.0375*	0.8433**	0.8465**	0.8909***	0.8919***	0.0503*	0.0498
Double-clustered s.e.	(0.1734)	(0.1786)	(0.1384)	(0.1476)	(0.0193)	(0.0204)	(0.3370)	(0.3403)	(0.2318)	(0.2361)	(0.0293)	(0.0306)
Log GDP p.c.	-0.3487**	-0.4679***	0.0773	-0.0087	-0.0571***	-0.0659***	-0.4362	-0.6028**	0.0024	-0.1144	-0.03	-0.0277
Double-clustered s.e.	(0.1585)	(0.1723)	(0.1999)	(0.1891)	(0.0158)	(0.0184)	(0.2709)	(0.2783)	(0.3776)	(0.3870)	(0.0302)	(0.0287)
Log Likelihood	-2618.39	-2420.26	-2457.95	-2255.25	__	__	-994.321	-945.006	-938.857	-891.605	__	__
Adjusted R-squared	__	__	__	__	0.457	0.447	__	__	__	__	0.466	0.457
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1167	1124	1167	1124	1167	1124	575	564	575	564	575	564

The table reports Negative Binomial Maximum Likelihood (ML) estimates (in columns (1)-(4) and (7)-(10) and linear probability (LS) estimates associating various aspects of civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in columns (1), (2), (7) and (8) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3), (4), (9), and (10) is the number of violent events against civilian populations. The dependent variable in columns (5), (6), (11) and (12) is a dummy variable that equals one if a battle resulting in territorial change of control has taken place and zero otherwise. All specifications include country fixed effects (constants not reported). The specifications in columns (7)-(12) focus on ethnicity-country areas close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland within a country to the national border; 62 kilometers). The specifications in even-numbered columns exclude country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. In all specifications we control for the log of regional GDP per capita in 2000 using data from Nordhaus and Chen (2012) G-Econ project. 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 9: The Long-Run Effects of Ethnic Partitioning on Individual Well-Being and Education. DHS Data
Baseline Estimates

	All Observations (households)						Observations (households) close to the National Border					
	Composite Wealth Index			Education			Composite Wealth Index			Education		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Ethnic Partitioning	-0.3887***	-0.3306***	-0.2062**	-0.3241**	-0.2335**	-0.1560*	-0.3074***	-0.2974***	-0.2660***	-0.1848*	-0.1611*	-0.1392*
Double-clustered s.e.	(0.1173)	(0.1028)	(0.0986)	(0.1348)	(0.0985)	(0.0842)	(0.1102)	(0.1016)	(0.0844)	(0.1071)	(0.0928)	(0.0786)
Adjusted R-squared	0.025	0.064	0.161	0.143	0.220	0.254	0.056	0.082	0.154	0.120	0.184	0.210
Observations	88124	86574	86574	88123	86573	86573	44525	43503	43503	44524	43502	43502
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 a composite wealth index (in columns (1)-(3) and (7)-(9)) and an education index (in columns (4)-(6) and (10)-(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. 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 (7)-(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 10: The Long-Run Effects of Ethnic Partitioning on Individual Well-Being and Education. DHS Data Channels; Location and Identity

	All Observations (households)						Observations (households) close to the National Border					
	Composite Wealth Index			Education			Composite Wealth Index			Education		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Partitioning - Identity	-0.3099**	-0.2629**	-0.2151**	-0.3100***	-0.2998***	-0.2790***	-0.2880**	-0.2120**	-0.1751**	-0.2042*	-0.1802*	-0.1635**
Double-clustered s.e.	(0.1292)	(0.1142)	(0.1017)	(0.1155)	(0.1068)	(0.0857)	(0.1390)	(0.1043)	(0.0880)	(0.1114)	(0.0965)	(0.0807)
Partitioning - Location	-0.2425**	-0.2110**	0.0348	0.0109	0.0104	0.0574	-0.1111	-0.067	0.0747	0.0834	0.0808	0.1072
Double-clustered s.e.	(0.1167)	(0.1066)	(0.0792)	(0.1380)	(0.1299)	(0.0783)	(0.0838)	(0.0677)	(0.0612)	(0.1072)	(0.0915)	(0.0635)
Adjusted R-squared	0.0300	0.0673	0.1607	0.0564	0.0818	0.1541	0.1440	0.2200	0.2544	0.1207	0.1841	0.2106
Observations	88124	86574	86574	44525	43503	43503	88123	86573	86573	44524	43502	43502
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 a composite wealth index (in columns (1)-(3) and (7)-(9)) and an education index (in columns (4)-(6) and (10)-(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 (range from 0 to 5) education index 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 partitioned ethnicity and zero otherwise.

The location based ethnic partitioning index takes on the value of one for individuals that reside in ethnic homelands that have been partitioned by the national border and zero otherwise.

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 (7)-(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 11: The Long-Run Effects of Ethnic Partitioning. Further Evidence on Channels
Looking within Partitioned Ethnic Homelands

	Composite Wealth Index			Education			Literacy		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnic Partitioning	-0.3410***	-0.3350***	-0.3220***	-0.2322**	-0.2169**	-0.2064**	-0.0650**	-0.0570**	-0.0526**
Double-clustered s.e.	(0.1180)	(0.1100)	(0.1001)	(0.1075)	(0.0957)	(0.0904)	(0.0293)	(0.0265)	(0.0239)
Adjusted R-squared	0.060	0.087	0.138	0.127	0.188	0.204	0.253	0.281	0.290
Observations	39215	38371	38371	39215	38371	38371	39215	38371	38371
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Location Controls	No	No	Yes	No	No	Yes	No	No	Yes

The table reports OLS estimates associating a composite wealth index (in columns (1)-(3)), education (in columns (4)-(6)), and literacy (in columns (7)-(9)) with ethnic partitioning at the individual level. All specifications are estimated in the sample of individuals residing in homelands of ethnicities that have been systematically partitioned by the national border.

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 (range from 0 to 5) education index 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 literacy indicator takes the value of one when a respondent can read a sentence that has been presented to him by the DHS team in the respondent's language; the index equals zero when the respondent can neither read at all or can only read parts of the sentence.

The ethnic partitioning index takes on the value of one for individuals that identify with a partitioned ethnicity and zero otherwise.

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 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 12: The Long-Run Effects of Ethnic Partitioning.
Enumeration-area Fixed Effects Estimates

	Composite Wealth Index		Education		Literacy	
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic Partitioning	-0.0670***	-0.0655***	-0.0938**	-0.0877**	-0.0215**	-0.0188**
Double-clustered s.e.	(0.0208)	(0.0209)	(0.0368)	(0.0349)	(0.0094)	(0.0090)
Adjusted R-squared	0.692	0.691	0.462	0.484	0.408	0.422
Observations	86600	85075	86599	85074	86600	85075
Enumeration Area Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	No	Yes	No	Yes

The table reports OLS estimates associating a composite wealth index (in columns (1)-(2)), education (in columns (3)-(4)), and literacy (in columns (5)-(6)) 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 literacy indicator takes the value of one when a respondent can read a sentence that has been presented to him by the DHS team in the respondent's language; the index equals zero when the respondent can neither read at all or can only read parts of the sentence. The ethnic partitioning index takes on the value of one for individuals that identify with a partitioned ethnicity and zero otherwise.

All specifications include a vector of enumeration area (village/town/city) 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 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 13: The Long-Run Effects of Ethnic Partitioning.
Examining Persistence

	All Observations				Observations close to the National Border			
	Young	Old	Young	Old	Young	Old	Young	Old
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Baseline Ethnic Partitioning - Identity	-0.2171**	-0.1935**	-0.2347**	-0.1938**	-0.2500***	-0.2755***	-0.2679***	-0.2843***
Double-clustered s.e.	(0.1066)	(0.0936)	(0.1089)	(0.0972)	(0.0876)	(0.0831)	(0.0875)	(0.0858)
Ethnic Partitioning - Location			0.0639	0.0009			0.0738	0.0419
Double-clustered s.e.			(0.0803)	(0.0812)			(0.0787)	(0.0853)
Adjusted R-squared	0.160	0.164	0.161	0.164	0.154	0.161	0.154	0.161
Observations	44546	42028	44546	42028	22392	21111	22392	21111
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports OLS estimates associating a composite wealth index with ethnic partitioning at the individual level. Odd-numbered columns report estimates restricting estimation to individuals born after 1977 (young). Even-numbered columns report estimates restricting estimation to individuals born before (or on) 1977 (old).
The composite wealth index (dependent variable) is calculated by the DHS team in each country via a principal component method using easy-to-collect data on a household's ownership of selected assets (e.g., televisions and bicycles), materials used for housing construction and public good access (e.g., type of water access, electrification, and sanitation).
The ethnic partitioning index takes on the value of one for individuals that identify with a partitioned ethnicity and zero otherwise.
The location based ethnic partitioning index takes on the value of one for individuals that reside in ethnic homelands that have been partitioned by the national border.
All specifications include a vector of country ethnic homeland fixed effects (constants not reported). All specifications include a vector of 10 age-bracket fixed effects, a vector of 6 marital-status fixed effects, and a vector of 7 religion fixed effects (individual controls). All specifications also include as controls 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 (location controls).
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

On-line Supplementary Appendix

(not for publication)

The Long-Run Effects of the Scramble for Africa[*]

Stelios Michalopoulos
Brown University and NBER

Elias Papaioannou
London Business School, NBER and CEPR

September 23, 2013

Abstract

The analysis in the Supplementary Appendix is structured into five sections. Section 1 gives descriptive evidence and summary statistics. Section 2 reports additional results showing that ethnic partitioning is not systematically related to various ethnic-specific pre-colonial economic, social, and political traits using data from Murdock's Ethnographic Atlas. Section 3 reports numerous sensitivity checks illustrating the impact of ethnic partitioning on conflict. In Section 4 we explore potential heterogeneous effects of ethnic partitioning on conflict. Section 5 reports further sensitivity checks on the negative effect of identifying with a partitioned ethnicity using micro-level data on well-being and education from the Demographic and Health Surveys.

---

1 Descriptive Evidence and Summary Statistics

Appendix Table 1 gives summary statistics for the main variables employed in the empirical analysis. In Panel A we tabulate summary statistics at the ethnic homeland level, which is the unit of analysis in Section 3 that examines the correlates of ethnic partitioning. In Panel B we tabulate summary statistics at the country-ethnic homeland level, which is the unit of analysis in Section 4 that assesses the impact of ethnic partitioning on conflict.

Appendix Table 2 gives the correlation structure of the main outcome variables capturing conflict, namely the total number of conflict events, battles between armed forces, battles resulting in territorial change, and violence against the civilian population.

Appendix Figures 1 and 2 plot the number of battles and an indicator that identifies ethnic homelands where a battle resulted in a territorial change, respectively. Appendix Figure 3 portrays the spatial distribution of violence against civilians.
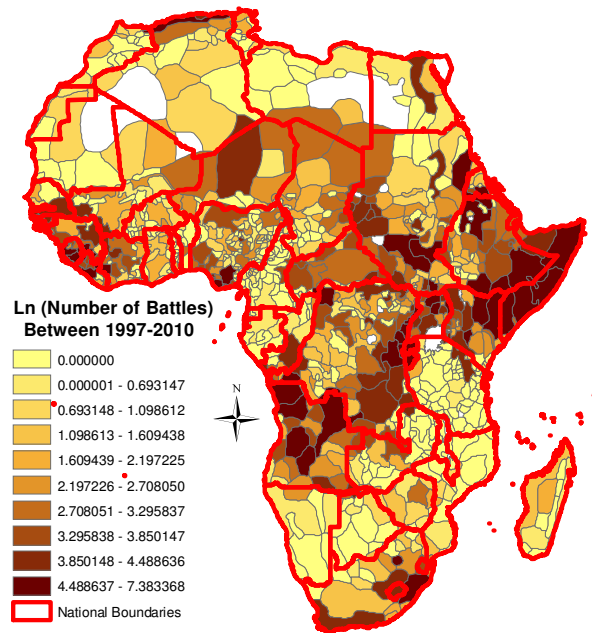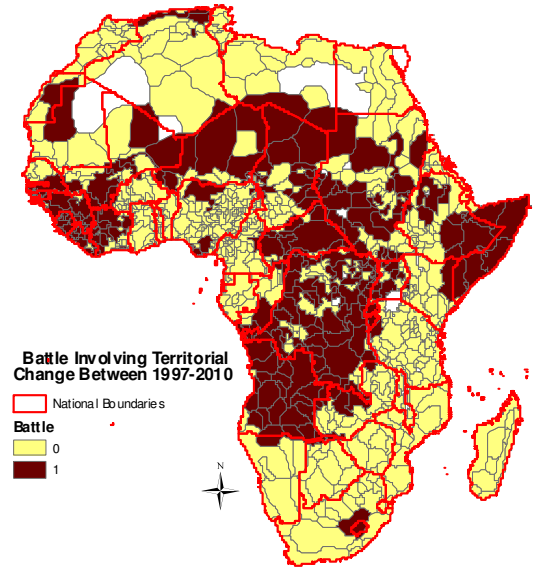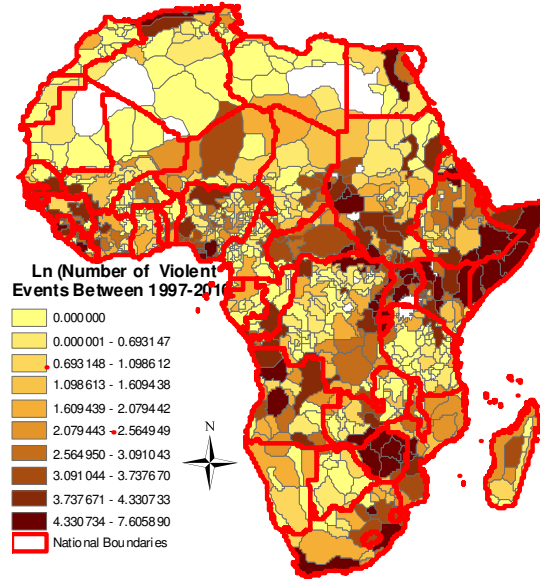

Figure **1**

Figure **2**

Figure **3**

2 Sensitivity Analysis. Border Artificiality

The Role of Ethnic-Specific Pre-colonial Traits   In Appendix Table 3 we associate ethnic partitioning with numerous ethnic-specific variables from Murdock (1967), that measure at the time of colonization the dependence of the economy on agriculture, fishing, hunting, the type of family organization, the presence of rules for inheritance, the role of clans, and other traits. Specifications in column A are simple unconditional models. Specifications in column B include a set of region fixed effects. There are no significant differences between split and

non-partitioned ethnicities. And while these variables do contain noise (and thus attenuation is a valid concern), measurement error cannot fully explain the lack of significance, as many works document robust associations between several societal traits in Murdock's Ethnographic Atlas and contemporary outcomes (e.g., Fenske (2012b), Gennaioli and Rainer (2007), Fenske (2012a), Alesina, Giuliano, and Nunn (2013)).

Geographic Features and Ethnic Partitioning. Country-Fixed-Effects Estimates Since in our empirical analysis linking conflict to ethnic partitioning we mostly explore within-country variation, one would also like to know whether post-independence there are systematic differences between partitioned ethnicities and non-split groups in the same country. In Appendix Table 4 we report country fixed effects specifications associating geographical, ecological, and natural resource features with ethnic partitioning. In this case the unit of analysis becomes an ethnicity-country observation assigning each partition to the corresponding country. The evidence suggests that the two sets of ethnic groups are comparable. Interestingly, the positive correlation between surface, water area and partitioning turns now insignificant. This is because after partitioning both the surface area and the area under water of split groups within a country are similar to those of non-partitioned ethnicities.

3 Sensitivity Analysis. Ethnic Partitioning and Conflict

Deadly Conflicts First, we repeated estimation focusing on deadly incidents. While the ACLED does not report an official number of casualties, going over the description of each event we can identify deadly incidents. As it is interesting to investigate whether ethnic partitioning is linked to casualties, we re-run our baseline specification focusing on deadly incidents. Appendix Table 5 gives the results. The coefficient on the ethnic partitioning index is positive and significant across all perturbations suggesting that not only partitioning is associated with a higher incidence of conflict, but those violent events (mostly battles and violence against civilians) were also deadly.

Data Sources Second, to minimize concerns related to data quality we classified each of the approximately 43,000 events based on the original source (international newspaper, global news agency, NGOs, United Nations, academic bulletins, local sources) and then re-estimated the empirical specifications using events reported solely by international sources, such as global news networks (e.g., Associated Press, Reuters, BBC), non-African newspapers (e.g., the Guardian, New York Times, 24 Horas), international NGOs (e.g., the Human Rights Watch, Amnesty International), and the United Nations. Appendix Table 6-Panel *A* reports

the results. In all specifications the coefficient on the ethnic partitioning indicator is positive and statistically significant at the 99% confidence level. We also estimated models using only those events reported by NGOs and the various reports of the United Nations teams. Appendix Table 6-Panel B reports the results. While the number of events drops dramatically (as we use just 15% of the data) and we lose countries which are not covered either by NGO's or UN reports, this minimizes concerns of any media-specific bias. In spite of the efficiency loss, we obtain similar results.

Alternative Measures of Partitioning   Third, we repeated estimation using alternative measures of ethnic partitioning. Appendix Table 7 gives the results. In columns (1)-(3) and (7)-(9) we associate conflict with a continuous index of partitioning ($FRAC$) that in the spirit of ethnolinguistic fragmentation measures reflects the probability that a randomly chosen pixel (area) of the historical homeland of an ethnic group falls into a different country (e.g., Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003)). The ethnic partitioning measure enters with a positive and highly significant coefficient, implying that a higher degree of ethnic partitioning is associated with a higher likelihood of civil conflict. In columns (4)-(6) and (10)-(12) we regress civil conflict on a binary index of ethnic partitioning using a 5% threshold to identify split groups. This has little effect on our baseline results. In all permutations the binary index of ethnic partitioning enters with a highly significant positive estimate.

Location   Fourth, we estimated models dropping iteratively ethnic homelands from each of the five African regions so as to investigate whether the results are driven by a particular part of the continent. Appendix Table 8 reports the results. In (1)-(2) we exclude North Africa to account for the fact that Europeans had contacts with the northern part of the continent since the ancient times. In (3)-(4) we drop Southern African countries. In columns (5)-(6) we drop Western African countries because some of the contemporary African borders in this region correspond to internal administrative borders of the Federation of the French West Africa. In (7)-(8) and (9)-(10) we exclude ethnic areas in East Africa and Central Africa, respectively. This allows us to examine the robustness of our results to influential observations, as the most deadly and prolonged conflicts have taken place in Ethiopia, Somalia, Sudan, Rwanda, and the Democratic Republic of Congo. Moreover, since most large pre-colonial African states were present in Ethiopia and in other parts of Eastern Africa (Besley and Reynal-Querol (2012)), we further account for the role of pre-colonial conflict and political development. The results show that the strong positive effect of ethnic partitioning on civil conflict is not driven by a particular region.[1]

---

[1] The same applies if we drop iteratively countries.

Spatial Effects Fifth, we estimated spatial models to account for spill-overs and for factors in nearby ethnic regions. Appendix Table 9 reports spatial lag specifications (estimated by maximum-likelihood). Since spatial lag models are somewhat sensitive to the details of the specification, we report estimates using two different types of (inverse) weighting matrixes (linear in distance and quadratic in distance) and three different types of models. Accounting for spatial spill-overs in conflict as well as the effect of the independent variables in the homelands of neighboring groups (in the Durbin and Generalized Spatial models) has little effect on the coefficient of ethnic partitioning that retains its economic and statistical significance.

Spillovers Sixth, while in the presence of externalities our estimates will be lower bounds of the effect of partitioning, in Appendix Table 10 we further explore the impact of spatial spill-overs running specifications linking conflict with ethnic partitioning conditioning on the total number of conflicts in the same country (netting out conflict in each ethnic homeland) and conditioning on total conflict of each ethnolinguistic family (netting out conflict of each ethnic homeland). The cross-sectional estimates reveal that conflict is higher when an ethnic homeland falls in a conflict-prone country and when there is a lot of conflict in groups from the same family. These results are quite interesting as they show that conflict and violence against civilians spread within country and along ethnic family lines. Yet, this has little impact on the ethnic partitioning index that retains its economic and statistical significance.

Accounting for Pre-colonial Conflict Seventh, we estimated empirical models linking contemporary conflict to ethnic partitioning, controlling for the historical legacy of violence. Appendix Table 11 reports the results. Following Besley and Reynal-Querol (2012), in Panel A we augment the empirical specification with the distance of the centroid of each ethnic homeland to the center of the closest pre-colonial conflict. Given the lack of association between pre-colonial conflict and ethnic partitioning (see Table 2 in the main part of the paper), it should come at no surprise that the estimate on the partitioning index retains its economic and statistical significance across all permutations. Since many wars before colonization took place in the territories of large centralized kingdoms, in Panel B we include in the specification a dummy variable that takes on the value of one if an ethnic homeland was part of a large pre-colonial kingdom. The coefficient on the ethnic partitioning index is positive and highly significant. In Panel C, we control for the impact of conflict during the slave trades augmenting the specification with the estimates of Nunn (2008) and Nunn and Wantchekon (2011) on enslavement at the ethnicity level. The coefficient on the partitioning index retains intact, further reassuring that it does not capture conflict during the pre-colonial period.

To further minimize concerns that ethnic partitioning captures the legacy of pre-colonial

5

conflict, in Appendix Table 12 we report estimates excluding all ethnic homelands that experienced significant pre-colonial conflict. The results are intact. Likewise, in Appendix Table 13 we drop from the estimation all ethnic areas that were part of large pre-colonial states, kingdoms, and empires. In spite of the significant drop in the number of observations, the coefficient on the ethnic partitioning index retains its economic and statistical significance.

Contemporary Population Density Eighth, we estimated models linking conflict to ethnic partitioning conditioning on contemporary (log) population density. Because civil conflict affects population dynamics and following the recommendation of Angrist and Pischke (2008) in the main part of the paper we control for (log) population at the ethnic homeland level using the earliest census. Yet, as the results reported in Supplementary Appendix Table 14 show using contemporary population estimates has little impact on the estimate on the ethnic partitioning index.

Alternative Mapping of Ethnic Homelands and Ethnic Partitioning Ninth, we repeated the analysis using *Ethnologue's* database that reports the spatial distribution of linguistic groups in the early/mid 1990s. *Ethnologue* explicitly maps linguistic homelands within each country making the identification of partitioned groups straightforward. Moreover, using a contemporary dataset is useful because it contains less error than Murdock's pre-colonial map. The disadvantage is that the current location of ethnic groups is likely to have been affected by the border drawing, local violence, colonization, as well as numerous country-level characteristics that affect conflict. *Ethnologue* includes information on 2405 linguistic groups in Africa out of which 821 are mapped in more than one countries. Partitioned groups are 48% likely to experience at least one conflict compared to 33% for non-partitioned groups. Moreover, conditional on having at least one conflict split groups experience an average of 61 incidents whereas non-split ones register 34 conflict events. Overall, partitioned groups have suffered an average of 30 conflict incidents in-between $1997 - 2010$ whereas groups located in a single country have experienced on average 10 conflict events. All differences are statistically significant at conventional levels.

Appendix Table 15 reports cross-sectional and within-country specifications using the *Ethnologue* data. Panel A reports negative binomial maximum likelihood estimates; Panel B reports analogous LS specifications using the log of one plus the number of all civil conflict events as the dependent variable. Columns (1)-(6) include all linguistics groups whereas in columns (7)-(12) we focus on groups whose centroid's distance to the national boundary is less than the median distance (86 kilometers). The coefficient on the ethnic partitioning index is positive and highly significant across all permutations. The most conservative estimate implies

6

that conflict intensity is approximately 30% ($exp(0.267) - 1 = 0.30$) higher in the contemporary homelands of partitioned groups.

Riots   The arguments put forward by the African historiography stress the role of ethnic partitioning in spurring civil conflict; hence, if this hypothesis is correct one should not be able to detect a significant association between partitioning and riots and protests, that are mostly observed in capitals. This is what we show in Supplementary Appendix Table 16. Across all permutations the ethnic partitioning index enters with a small and statistically indistinguishable from zero estimate.

Cross-Sectional Estimates   Finally, in Appendix Table 17 we report cross-sectional estimates associating each type of civil conflict with ethnic partitioning. So, these estimates are similar to the ones in Table 3, where the dependent variable was the total number of conflict incidents in each ethnic region. Ethnic partitioning is systematically positively linked both to battles (Panel A) and violence against the civilian population (Panel B).

4   Heterogeneous Effects

Our results suggest that ethnic partitioning has had a strong positive effect on civil conflict, on average. Yet it may be the case that ethnic partitioning is especially harmful in certain environments. Based on previous works on the origins of African development and the historical narratives of the influence of the Scramble for Africa on conflict, in Appendix Tables 18 and 19 we explore potential heterogeneous effects of partitioning. Searching for heterogeneous effects is also helpful as it sheds light on how and under which circumstances ethnic partitioning spurs conflict.

4.1   Ethnic Composition, Size, and Landlocked Countries

The Scramble for Africa has led to the creation of some very large and heterogeneous states. Moreover, the border design resulted in Africa having the most landlocked countries across all continents. In Appendix Table 18 we examine whether these aspects of the Scramble for Africa interact with ethnic partitioning. In columns (1)-(3) we investigate whether the impact of ethnic partitioning differs with respect to the degree of ethnic, linguistic, or religious fractionalization; we do so by augmenting the baseline specification with an interaction term between the $SPLIT$ index and a dummy identifying countries scoring above the median on the Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003) fractionalization measures. The estimates show that the impact of ethnic partitioning is quite homogeneous and does not depend

7

on the degree of fractionalization. The results are similar when we group countries based on the degree of cultural fractionalization (using data from Fearon (2003)), ethnic polarization (using the index of Montalvo and Reynal-Querol (2005)), or ethnic segregation (see Alesina and Zhuravskaya (2011) and Matuszeski and Schneider (2006)).

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

4.2   The Role of National Institutions

In Appendix Table 19 we examine whether the negative consequences of ethnic participating are mitigated in relatively more (institutionally and economically) advanced countries. In column (1) we use a composite rule of law index that reflects the quality of property rights and legal institutions (using data from World Bank's Governance Matters Project). The interaction term enters with a negative coefficient, hinting that sound national institutions may attenuate the impact of partitioning; yet the estimate is not significant at standard confidence levels (the results are similar with other proxies of legal quality). In column (2) we allow the effect of ethnic partitioning to differ for democratic countries, using a binary democracy variable from Przeworski, Alvarez, Cheibub, and Limongi (2000). The interaction enters with a significantly negative coefficient suggesting that democratic regimes are better at dealing with the adverse consequences of ethnic partitioning. The results are similar when we use alternative sources to identify democratic countries (Freedom House, Polity Project). In column (3) we further explore the role of democracy interacting the partitioning index with a measure of democratic capital (from Persson and Tabellini (2009)) that reflects counties' experience with representative rule (rather than the degree of democracy in a given year). The interaction is negative and significant, further showing that democratic rule tends to reduce the negative impact of ethnic partitioning.

While these results are supportive of the idea that national institutions and democracy, in particular, may be beneficial by lowering the propensity of conflict emerging from ethnic

partitioning, they cannot be casually interpreted because conflict may be both a cause and a consequence of institutional quality. Moreover, both civil war and institutional quality are a function of economic development, as well as other factors (related to geography, culture, colonization, etc.). In line with this, specification (4) shows that the positive effect of partitioning on conflict is lower in more economically advanced countries. In columns (5)-(6) we jointly include in the specification the interaction between ethnic partitioning and democracy and the interaction of ethnic partitioning with economic development. Both interactions enter with significantly negative coefficients, implying that the attenuating effect of democracy on conflict works on top of the role of economic development. In columns (7)-(8), following Acemoglu, Johnson, Robinson, and Yared (2008), we use executive constraints at independence (in the 1960s) and income per capita in 1961 to identify institutionally and economically advanced countries at the time of de-colonization. The results suggest that partitioned groups found in countries that at independence were relatively more institutionally advanced experience today less conflict, implying that inclusive institutions may be helpful in lowering conflict caused by ethnic partitioning. If civil conflict and institutional development are inter-liked, then these results suggest ethnic partitioning may have contributed to the emergence and persistence of a conflict-driven poverty trap, where partitioning has particularly strong effects in under-developed countries, which in turn impedes the consolidation of inclusive institutions

5 Further Evidence. Micro-level Based Evidence with DHS Data

5.1 Summary Statistics

Appendix Table 20 gives summary statistics for the variables employed in Section 5; all data come from the Demographic and Health Surveys, spanning 20 African countries, namely: The countries and interview years are Benin in 2001, Burkina Faso in 2010, Central African Republic in 1994, Ethiopia in 2011, Ghana in 2008, Guinea in 2005, Kenya in 2008, Mali in 2006, Mozambique in 2011, Malawi in 2010, Namibia in 2000, Niger in 1998, Senegal in 2010, Sierra Leone in 2008, Togo in 1998, Uganda in 2011, the Democratic Republic of Congo in 2007, and Zambia in 2007.

5.2 Sensitivity Analysis

Results with Literacy In Appendix Table 21 we report linear probability estimates associating an indicator (dummy) variable of literacy with ethnic partitioning. The estimates reported in columns (1)-(3) and (7)-(9) are similar to the baseline specifications with the composite wealth index and education reported in Table 10. Across all specifications the coefficient

on the ethnic partitioning index that identifies individuals that self-identify with partitioned ethnic groups is negative and highly significant. The estimate implies that individuals from ethnicities that have been systematically split by the artificial drawing of African borders are on average $4\% - 7\%$ less likely to be able to perfectly read a simple sentence, as compared to individuals from non-partitioned ethnicities. This effect is far from being small, as on average only 56% of respondents is literate.

In columns (4)-(6) and (10)-(12) we augment the specification with the location-based measure of partitioning that takes on the value of one for individuals residing in the ancestral homeland of split groups. So these empirical models are similar to the ones reported in Table 11. The baseline ethnic origin-based partitioning index enters with a negative and highly significant estimate (around -0.20), implying that conditional on various individual and location characteristics, individuals from partitioned ethnicities are on average much more likely to be illiterate as compared to people from non-split groups. In contrast -and as in the estimates in Table 11- literacy is not systematically to the regional-based of partitioning. Quantitatively, the estimate in the last specification implies that even when we zoom in border areas and condition on numerous individual and location characteristics, household heads of a partitioned ethnic background have a 4% lower likelihood of being literate, as compared to individuals from non-split ethnicities.

Ethnic-homeland Fixed Effects Estimates   In Appendix Table 22 we estimate similar to Table 12 specifications restricting estimation to partitioned ethnic homelands; yet rather than conditioning on country constants, we include a set of country-ethnic homeland fixed effects, so as to further account for location-specific characteristics. This allows to isolate the effect of a respondents' ethnic origin on the three outcome variables. In all specifications the coefficient on the ethnic partitioning index is negative and significant. These results complement our analysis at the ethnic homeland level, by showing that when we focus within split homelands, unlike households from a non-split ethnic background, it is the indigenous ones (i.e., those that belong to the partitioned group in place) that underperform economically. This evidence points out that ethnic partitioning has disproportionately adversely affected the welfare of individuals belonging to split groups exacerbating economic inequality along ethnic lines.

6 Data Sources and Variable Definitions

6.1 Pre-colonial Ethnic Features from Murdock (1967)

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

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

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

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

Agriculture Type: $0-4$ scale index reflecting the type of agriculture. The index equals 0 when there is "no agriculture"; 1 when there is "causal agriculture"; 2 when there is "extensive or shifting agriculture"; 3 when there is "intensive agriculture"; and 4 when there is "intensive irrigated agriculture." *Source: Murdock (1967); variable code v28.*

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

Binary Class Stratification: The dummy stratification index equals zero when Murdock's variable equals zero indicating "absence of significant class distinctions among freemen, ignoring variations in individual repute achieved through skill, valor, piety, or wisdom," and one when Murdock's class stratification measure equals 1, 2, 3, or 4. The construction of this variable follows Gennaioli and Rainer (2006, 2007). *Source: Murdock (1967); variable code v67.*

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

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

Jurisdictional Hierarchy of Local Community: Ordered variable ranging from 0 to

2 reflecting the hierarchy of local community organization. A zero score indicates the theoretical minimum of two (e.g.„ family and band), while a score of 2 indicates the theoretical maximum of four levels (e.g.„ nuclear family, extended family, clan barrio, village levels). *Source: Murdock (1967); variable code v32.*

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

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

6.2   County-level Variables Used in Heterogeneous Effects Section

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

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

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

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

Ethnolinguistic Polarization: Index of ethnolinguistic polarization that achieves a maximum score when a country is occupied by two groups of the same population. *Source: Montalvo and Reynal-Querol (2005).*

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

and Zhuravskaya (2011).

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

Democracy: Binary index of political regime (autocracy and democracy), based on the conduct of contested elections in 2000. *Source: Przeworski, Alvarez, Cheibub, and Limongi (2000); updated series available:*

http://politics.as.nyu.edu/object/przeworskilinks.html.

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

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

References

ACEMOGLU, D., S. JOHNSON, J. A. ROBINSON, AND P. YARED (2008): "Income and Democracy," *American Economic Review*, 98(3), 808–842.

ALESINA, A., A. DEVLEESCHAUWER, W. EASTERLY, S. KURLAT, AND R. WACZIARG (2003): "Fractionalization," *Journal of Economic Growth*, 8(2), 155–194.

ALESINA, A., AND E. ZHURAVSKAYA (2011): "Segregation and the Quality of Government in a Cross-Section of Countries," *American Economic Review*, 101(5), 1872–1911.

ALESINA, A. F., P. GIULIANO, AND N. NUNN (2013): "On the Origins of Gender Roles: Women and the Plough," *Quarterly Journal of Economics*, 128(2), 469–530.

ANGRIST, J., AND J.-S. PISCHKE (2008): *Mostly Harmless Econometrics*. Princeton University Press, Princeton, NJ.

BESLEY, T., AND M. REYNAL-QUEROL (2012): "The Legacy of Historical Conflicts. Evidence from Africa," mimeo, London School of Economics and UPF.

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

FEARON, J. D. (2003): "Ethnic Structure and Cultural Diversity by Country," *Journal of Economic Growth*, 8, 195–222.

FENSKE, J. (2012a): "African polygamy: Past and present," mimeo, Oxford University.

———— (2012b): "Ecology, Trade, and States in Pre-colonial Africa," mimeo, Oxford University.

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

HESTON, A., R. SUMMERS, AND B. ATEN (2011): "Penn World Table Version 7.0," Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

KAUFMANN, D., A. KRAAY, AND M. MASTRUZZI (2005): "Governance Matters IV: Governance Indicators," *Policy Research Working Paper Series, The World Bank*.

MARSHALL, M. G., K. JAGGERS, AND T. R. GUR (2011): *Polity IV Project. Political Regime Characteristics and Transitions, 1800-2010. User's Manual.* Center for Systemic Peace.

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

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

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

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.

PERSSON, T., AND G. TABELLINI (2009): "Democractic Capital: The Nexus of Political and Economic Change," *American Economic Journal: Macroeconomics*, 1(1), 88–126.

PRZEWORSKI, A., M. ALVAREZ, J. CHEIBUB, AND F. LIMONGI (2000): *Democracy and Development: Political Institutions and Well-Being in the World 1950-1900.* Cambridge University Press, Cambridge, United Kingdom.

Appendix Table 1: Summary Statistics

	Obs.	mean	st. dev.	min	p25	median	p75	p99	max
Panel A: Ethnic Homeland Level									
All Civil Conflict Incidents	826	52.39	195.69	0.00	1.00	4.00	34.00	580.00	2916.00
Battles	826	22.65	94.44	0.00	0.00	1.00	13.00	313.00	1608.00
Battles with Territorial Change	826	19.18	96.20	0.00	0.00	1.00	9.00	231.00	2009.00
Violent Events against Civilians	826	0.27	0.44	0.00	0.00	0.00	1.00	1.00	1.00
Population at Independence	826	328,991	1,063,156	57	41,810	118,160	303,659	2,912,382	25,700,000
Land Area	826	34.17	59.18	0.24	6.16	14.48	36.07	286.33	604.90
Land Area under Water	826	0.86	2.26	0.00	0.01	0.17	0.68	10.87	27.66
Catital City Indicator	826	0.06	0.23	0.00	0.00	0.00	0.00	1.00	1.00
Distance to the Capital City	826	141.15	126.00	0.00	44.87	102.25	204.98	536.73	636.87
Distance to the Sea Coast	826	499.98	371.34	11.31	255.59	391.81	629.83	1673.47	1846.93
Distance to the National Border	826	597.49	432.10	0.22	208.79	554.35	918.12	1609.77	1721.30
Mean Elevation	826	0.62	0.44	0.00	0.30	0.49	0.94	1.84	2.17
Land Suitability for Agriculture	826	0.41	0.24	0.00	0.26	0.42	0.57	0.93	0.98
Malaria Stability Index	826	0.75	0.36	0.00	0.58	0.98	1.00	1.00	1.00
Oil Indicator	826	0.13	0.40	0.00	0.00	0.00	0.00	2.00	4.00
Diamond Mine Indicator	826	0.12	0.33	0.00	0.00	0.00	0.00	1.00	1.00
Panel B: Country-Ethnic Homeland Level									
All Civil Conflict Incidents	1182	33.63	139.63	0.00	0.00	2.00	18.00	438.00	2888.00
Battles	1182	14.54	67.09	0.00	0.00	0.00	6.00	238.00	1608.00
Battles with Territorial Change	1182	1.94	7.88	0.00	0.00	0.00	0.00	1.00	100.00
Violent Events against Civilians	1182	12.25	72.43	0.00	0.00	1.00	5.00	182.00	2009.00
Population at Independence	1182	222,093	880,884	17	15,528	62,931	186,570	2,143,565	25,600,000
Land Area	1182	22.41	40.68	0.11	2.81	8.33	23.29	216.23	493.82
Land Area under Water	1182	0.51	1.39	0.00	0.00	0.10	0.40	6.18	21.63
Catital City Indicator	1182	0.04	0.19	0.00	0.00	0.00	0.00	1.00	1.00
Distance to the Capital City	1182	0.51	0.36	0.01	0.26	0.41	0.63	1.70	1.88
Distance to the Sea Coast	1182	0.60	0.43	0.00	0.23	0.56	0.93	1.63	1.74
Distance to the National Border	1182	0.11	0.12	0.00	0.02	0.06	0.16	0.52	0.64
Mean Elevation	1182	0.62	0.44	0.00	0.29	0.49	0.95	1.91	2.18
Land Suitability for Agriculture	1182	0.41	0.24	0.00	0.25	0.42	0.57	0.94	0.98
Malaria Stability Index	1182	0.72	0.34	0.00	0.50	0.89	1.00	1.00	1.00
Oil Indicator	1182	0.09	0.28	0.00	0.00	0.00	0.00	1.00	1.00
Diamond Mine Indicator	1182	0.09	0.29	0.00	0.00	0.00	0.00	1.00	1.00

The table reports descriptive statistics for all variables employed in the empirical analysis. Panel A reports summary statistics for all control variables at the ethnic homelad level. Panel B reports summary statistics for all control variables at the country-ethnicity sample. The Data Appendix gives detailed variable definitions and data sources. The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 2: Correlation Structure of the Main Conflict Variables

Panel A: Ethnic Homeland Level

All types of ethnic conflict	1			
Battles	0.8733*	1		
Violent events against civilians	0.9049*	0.6068*	1	
Battles resulting is change of territory (Indicator)	0.2529*	0.3164*	0.1422*	1

Panel B: Country-Ethnicity Homeland Level

All types of ethnic conflict	1			
Battles	0.8268*	1		
Violent events against civilians	0.8839*	0.4965*	1	
Battles resulting is change of territory (Indicator)	0.2462*	0.3208*	0.1209*	1

The table reports the correlation structure between all ACLED civil conflict measures. Panel A reports the correlogram at the ethnic homeland level (*N=826*); Panel B reports the correlogram at the country-ethnicity level (*N=1182*). * indicates statistical significance at the 95% confidence level. The Data Appendix gives detailed variable definitions and data sources.

Appendix Table 3 - Ethnic Partitioning as a Function of Economic, Social, Cultural, and Political Ethnic Using Data from Murdock's Ethnographic Atlas

	Specification A Unconditional Relationship		Specification B Conditional Relationship	
	Additional Variable (1)	Obs. (2)	Additional Variable (3)	Obs. (4)
Gathering	0.0128 (0.0524)	487	-0.0682 (0.0486)	487
Hunting	0.0415 (0.0483)	487	0.0388 (0.0452)	487
Fishing	0.0234 (0.0556)	487	-0.004 (0.0518)	487
Milking	-0.0116 (0.0552)	452	0.0118 (0.0377)	452
Alternative Argiculture Dep.	-0.0381 -0.0283	452	-0.0231 -0.0322	452
Polygyny	-0.0174 (0.0453)	478	-0.0051 (0.0503)	478
Clan Communities	0.0197 (0.0609)	396	0.0385 (0.0614)	396
Complex Settlements	0.0209 (0.0607)	451	0.017 (0.0528)	451
Binary Class Stratification	-0.0158 (0.0495)	394	-0.0754 (0.0457)	394
Jurisdictional Hierarchy of Local Community	0.0221 -0.0395	440	0.0061 -0.0319	440
Elections	-0.0663 (0.0822)	344	-0.0542 (0.0903)	344
Inheritance Rule for Property	-0.0821 (0.0927)	374	0.0211 (0.0898)	374

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

Appendix Table 4: Ethnic Partitioning and Geographic Characteristics within Countries

	Dependent variable is:								
	Log Land Area	Log Area under Water	Mean Elevation	Land Suitability	Malaria Stability	Distance to the Sea	Diamond Indicator	Petroleum Indicator	Major City in 1400
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
SPLIT (Partitioning)	-0.1748	-0.0190	0.0179	0.0001	-0.0091	0.0203	-0.0216	-0.0126	-0.0003
Double-clustered s.e.	(0.1092)	(0.0255)	(0.0369)	(0.0194)	(0.0221)	(0.0324)	(0.0224)	(0.0190)	(0.0097)
adjusted R-squared	0.223	0.255	0.575	0.418	0.640	0.620	0.340	0.187	0.063
Mean Dependent Variable	2.0620	0.2706	0.6168	0.4079	0.7209	0.6042	0.0854	0.0947	0.0270
Observations	1182	1182	1182	1182	1182	1182	1182	1182	1182
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports OLS estimates associating various geographical, ecological, and natural resource characteristics with ethnic partitioning within countries. The unit of analysis is an ethnic territory in a country (ethnicity-country). SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. All specifications include country fixed effects (constants not reported). The dependent variable in column (1) is the log of a country-ethnicity's region surface area; in column (2) is the log of (1 + surface area under water); column (3) is average elevation; in column (4) is an index capturing land's (soil) suitability (quality) for agriculture; in column (5) is the average value of a malaria stability index; in columns (6) is the distance of the centroid of each country-ethnic region to the nearest sea coast; in column (7) is a binary index that takes on the value of one if a diamond mine is present; in column (8) is a binary index that takes on the value of one if an oil/petroleum field is present; and in column (9) a binary index that takes on the value of one if a major city was present before European's arrival in Africa (in 1400). The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 5: Ethnic Partitioning and Deadly Civil Conflict

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

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in all columns is the total number of deadly conflict incidents at the country-ethnic homeland level. To identify deadly incidents we go over the event description searching for keywords "kill, death, died". In columns (1)-(4) we estimate the specifications in the full sample. The specification in column (5) excludes country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specification in column (6) excludes country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 6: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Distinguishing between Conflict Incident Sources

	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)

Panel A: Data from All International Sources

SPLIT Indicator (Partitioning)	0.4234***	0.4439**	0.6224***	0.5565***	0.5995***	0.5531***
	(0.1657)	(0.1852)	(0.1859)	(0.1732)	(0.1664)	(0.1760)
Log Likelihood	-3382.90	-3183.73	-3107.80	-3095.23	-2991.64	-2854.01
Observations	1182	1182	1182	1182	1170	1139

Panel B: Data only from NGOs and United Nation's Reports

SPLIT Indicator (Partitioning)	0.9258***	0.5588***	0.7778***	0.7943***	0.7830***	0.8183***
	(0.2995)	(0.2149)	(0.2779)	(0.2636)	(0.2757)	(0.2864)
Log Likelihood	-1330.075	-1164.565	-1125.269	-1120.363	-1028.719	-995.075
Observations	1038	1038	1038	1038	1026	1005
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in all columns is the total number of conflict incidents at the country-ethnic homeland level. In Panel A we use conflict events reported by international media, news agencies, NGOs, and the United Nations. In Panel B we use conflict events reported by NGOs and the United Nations. In columns (1)-(4) we estimate the specifications in the full sample. The specification in column (5) excludes country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specification in column (6) excludes country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 7: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis. Alternative Measures of Ethnic Partitioning

	All Ethnicity-Country Homelands						All Ethnicity-Country Homelands Close to the Border					
	All Obs	No Outliers	No Capitals	All Obs	No Outliers	No Capitals	All Obs	No Outliers	No Capitals	All Obs	No Outliers	No Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
FRAC	0.7463**	0.6824*	0.7422**				0.8122*	0.687	0.8115**			
Double-clustered s.e.	(0.3707)	(0.3883)	(0.3724)				(0.4340)	(0.4237)	(0.4337)			
SPLIT-ALT				0.4974***	0.5119***	0.5178***				0.8614***	0.8412***	0.8612***
Double-clustered s.e.				(0.1621)	(0.1728)	(0.1637)				(0.2377)	(0.2377)	(0.2374)
Log Likelihood	-3612.47	-3349.05	-3507.77	-3608.11	-3344.23	-3502.91	-1392.54	-1329.67	-1384.94	-1387.22	-1324.50	-1379.63
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1139	1170	1182	1139	1170	576	565	575	576	565	575

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnic homeland level. The dependent variable is the total number of civil conflict incidents at each ethnic homeland in each country over the period 1997-2010. FRAC is a continuous measure of ethnic partitioning that reflects the probability that a randomly chosen pixel (area) of the historical homeland of an ethnic group falls into a different country. SPLIT-ALT is an indicator variable that identifies partitioned ethnicities as those with at least 5% of the historical homeland falling into more than one contemporary country. All specifications include country fixed effects (constants not reported). The specifications in columns (7)-(12) focus on country-ethnic areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country to the national border; 62 kilometers). The specifications in columns (2), (5), (8) and (11) exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (3), (6), (9), and (12) exclude country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 8: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis. Excluding Each Time a Different African Region

	Excluding									
	North		South		West		East		Central	
	All	Border	All	Border	All	Border	All	Border	All	Border
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SPLIT (Partitioning)	0.5593*	0.7990*	0.6595***	0.8090**	0.5548***	1.2937***	0.7961***	0.8755**	0.4311**	0.4296**
Double-clustered s.e.	(0.3213)	(0.3844)	(0.1183)	(0.3297)	(0.1974)	(0.2514)	(0.1507)	(0.4904)	(0.1371)	(0.2000)
Log Likelihood	-3416.15	-1354.6	-3143.06	-1237.22	-2617.48	-894.705	-2558.18	-1026.96	-2635.13	-993.766
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1104	554	1040	508	808	350	893	456	883	436

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnic homeland level. The dependent variable is the total number of civil conflict incidents in an ethnic homeland within a country over the period 1997-2010. In columns (1)-(2) we exclude ethnicity-country observations that fall in North Africa. In columns (3)-(4) we exclude observations that fall in South Africa. In columns (5)-(6) exclude observations that fall in West Africa. In columns (7)-(8) we exclude observations that fall in East Africa. In column (9)-(10) we exclude observations that fall in Central Africa The regional classification follows Nunn (2007). All specifications include country fixed effects (constants not reported). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 9: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Accounting for Spillovers with Spatial Models

Weighting Matrix	Linear in Eucledian Distance			Quadratic in Eucledian Distance		
Spatial Model Type	Simple	Durbin	Generalized	Simple	Durbin	Generalized
	(1)	(2)	(3)	(4)	(5)	(6)
SPLIT Indicator (Partitioning)	0.1891**	0.1754**	0.1822**	0.1844**	0.1582**	0.1685**
	(0.0765)	(0.0757)	(0.0765)	(0.0756)	(0.0777)	(0.0781)
Log Likelihood	-1750.65	-1670.10	-1744.80	-1740.92	-1671.16	-1736.70
Observations	1182	1182	1182	590	590	590
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

The table reports spatial auto-regressive model maximum-likelihood estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnicity level. The dependent variable is the log of one plus the total number of civil conflict incidents in an ethnic region within a country over the period 1997-2010. Columns (1) and (4) report simple spatial lag models that control for conflicts in neighbouring regions. Columns (2) and (5) report Durbin spatial models that include as additional controls the vector of independent variables in neighbouring regions. Columns (3) and (6) report generalized spatial lag models that control both for conflicts in neighbouring regions and for the effect of the independent variables in nearby regions. In columns (1)-(3) we use a linear in Eucledian distance to the centroid of each country-ethnic region weighting matrix. In columns (4)-(6) we use a quadratic in Eucledian distance to the centroid of each country-ethnic region weighting matrix. The specifications in columns (1)-(4) are estimated in the full sample of country-ethnicity areas. The specification in column (5) excludes country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specification in column (6) excludes country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses standard errors accounting for heteroskedasticity and spatial correlation. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 10: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Accounting for Spillovers at the Country Level and at the Ethnic Family Level

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

The table reports OLS estimates, associating civil conflict incidents with ethnic partitioning at the country-ethnicity level, accounting for spatial spillovers. Panel A reports cross-sectional estimates. Panel B reports within country estimates (that include country constants). The dependent variable is the log of one plus the total number of civil conflict incidents in an ethnic region within a country over the period 1997-2010. The specifications in columns (4)-(6) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country to the national border; 62 kilometers). The specifications in columns (2) and (5) exclude country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (3) and (6) exclude country-ethnic homelands where capital cities fall. In all specifications we control for the log of one plus the total number of all conflict incidents in each ethnic family and the log of one plus the total number of all conflict incidents in each country minus conflicts in each country-ethnic area. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

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

at the country-ethnic homeland level. The dependent variable in columns (1) and (4) is the total number of conflict incidents; the dependent variable in columns (2) and (5) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3) and (6) is the number of violent events against civilian populations. In Panel A we estimate the specifications in the full sample. In columns (4)-(6) we restrict estimation on ethnicity-country areas close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland within a country to the national border; 62 kilometers). In Panel A we we control for the distance of each ethnic homeland to the centroid of the closest war during the period 1400-1700, usinig data from Besley and Reynal-Quaerol (2012). In Panel B we control an indicator variable that takes the value of one when the historical homeland falls within the boundaries of a large pre-colonial kingdom and empire using data from Besley and Reynal-Quaerol (2012). In Panel C we control for the log of one plus the number of slaves at the ethnicity level normalized by the surface area of each homeland, using data from Nunn (2008) and Nunn and Watchekon (2011). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 12: Ethnic Partitioning and Civil Conflict

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

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

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating various aspects of civil conflict with ethnic partitioning at the country-ethnic homeland level dropping from the estimation ethnic homelands that experienced a major war over the period 1400-1700. The dependent variable in columns (1) and (4) is the total number of conflict incidents; the dependent variable in columns (2) and (5) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3) and (6) is the number of violent events against civilian populations. In columns (1)-(3) we estimate the specifications in the full sample. In columns (4)-(6) we restrict estimation on ethnicity-country areas close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland within a country to the national border; 62 kilometers). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 13: Ethnic Partitioning and Civil Conflict

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

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

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating various aspects of civil conflict with ethnic partitioning at the country-ethnic homeland level dropping from the estimation ethnic homelands that were part of large pre-colonial states, kingdoms, and empires using the classification of O'Brian (1999) in Panel A and using the classification of Murdock (1967) in Panel B. The dependent variable in columns (1) and (4) is the total number of conflict incidents; the dependent variable in columns (2) and (5) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3) and (6) is the number of violent events against civilian populations. In columns (1)-(3) we estimate the specifications in the full sample. In columns (4)-(6) we restrict estimation on ethnicity-country areas close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland within a country to the national border; 62 kilometers). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 14: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Controlling for Population in 2000

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

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating various aspects of civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable in columns (1) and (4) is the total number of conflict incidents; the dependent variable in columns (2) and (5) is the total number of battles between government forces, rebel groups, and militias; the dependent variable in columns (3) and (6) is the number of violent events against civilian populations. In columns (1)-(3) we estimate the specifications in the full sample. In columns (4)-(6) we restrict estimation on ethnicity-country areas close to the national border (using as a cutoff the median distance from the centroid of each ethnic homeland within a country to the national border; 62 kilometers). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 2000. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 15: Ethnic Partitioning and Civil Conflict
Sensitivity Analysis: Results with Ethnologue's Mapping of Ethnic Groups

	All Language-Country Homelands						Language-Country Homelands Close to the National Border					
	All Observations				Excl. Outliers	Excl. Capitals	All Observations				Excl. Outliers	Excl. Capitals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: NB ML Estimates												
SPLIT (Partitioning)	0.7170***	0.4918**	0.4421***	0.3603***	0.3383***	0.3516***	0.5262***	0.3525***	0.3619***	0.3191***	0.2672***	0.3157***
Double-clustered s.e.	(0.1288)	(0.0968)	(0.0836)	(0.0899)	(0.0923)	(0.0905)	(0.1490)	(0.1230)	(0.0967)	(0.0948)	(0.1029)	(0.0961)
Log Likelihood	-4850.15	-4489.32	-4414.53	-4365.37	-4149.16	-4022.85	-2338.24	-2081.06	-2061.75	-2040.4	-1970.1	-1904.15
Panel B: LS Estimates												
SPLIT (Partitioning)	0.2467***	0.1496**	0.0747**	0.0903**	0.0803**	0.0856**	0.1729***	0.0843**	0.0794*	0.1097***	0.0933**	0.1148***
Double-clustered s.e.	(0.0586)	(0.0664)	(0.0377)	(0.0374)	(0.0334)	(0.0364)	(0.0172)	(0.0390)	(0.0440)	(0.0388)	(0.0397)	(0.0385)
Adjusted R-squared	0.381	0.546	0.597	0.609	0.569	0.54	0.283	0.537	0.587	0.602	0.579	0.542
Country FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	2405	2405	2405	2405	2379	2353	1202	1202	1202	1202	1194	1182

The table reports Negative Binomial Maximum Likelihood (ML) estimates in Panel A and OLS estimates (in Panel B), associating civil conflict incidents with ethnic partitioning at the country-ethnic homeland level. The dependent variable in Panel A is the total number of civil conflict incidents in an ethnic homeland within a country over the period 1997-2010. The dependent variable in Panel B is the log of one plus the total number of civil conflict incidents at each ethnic homeland in each country over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned groups as those that Ethnologue maps into more than one contemporary country. The specifications in columns (2)-(6) and (7)-(12) include country fixed effects (constants not reported). The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 2000. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 16: Ethnic Partitioning and Riots and Protests

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

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating riots and protests with ethnic partitioning at the country-ethnic homeland level. The dependent variable in all columns is the total number of riots and protests at the country-ethnic homeland level. In columns (1)-(4) we estimate the specifications in the full sample. The specification in column (5) excludes country-ethnic homelands where the dependent variable exceeds the 99th percentile. The specification in column (6) excludes country-ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 17: Ethnic Partitioning and Alternative Measures of Civil Conflict
Cross-Sectional Estimates

	All Ethnic Homelands						Ethnic Homelands Close to the National Border					
	All Observations			Excl. Outliers	Excl. Capitals		All Observations			Excl. Outliers	Excl. Capitals	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Battles												
SPLIT (Partitioning)	0.8545***	0.8543***	0.8536***	0.6579***	0.7706***	0.7288***	0.6687**	0.6884***	0.7843**	0.5764	0.6919*	0.6408
Double-clustered s.e.	(0.2970)	(0.2516)	(0.2305)	(0.2240)	(0.1991)	(0.2363)	(0.2788)	(0.1984)	(0.3383)	(0.3880)	(0.3456)	(0.4308)
Log Likelihood	-2483.17	-2452.05	-2418.24	-2406.04	-2349.76	-2171.71	-1213.03	-1193.44	-1176.18	-1165.08	-1113.3	-1060.28
Panel B: Violence against Civilians												
SPLIT (Partitioning)	0.6025***	0.5750***	0.5300**	0.5327**	0.6435***	0.6119**	0.6184**	0.6389***	0.9951***	0.9688***	1.1552***	0.9986***
Double-clustered s.e.	(0.2580)	(0.2489)	(0.2592)	(0.2507)	(0.2659)	(0.2549)	(0.2182)	(0.1913)	(0.2818)	(0.3062)	(0.2950)	(0.3186)
Log Likelihood	-2342.88	-2316.08	-2296.55	-2294.53	-2241.91	-2049.53	-1159.94	-1144.5	-1137.28	-1133.69	-1081.03	-1030.17
Panel C: Battles that Resulted in a Change of Territory												
SPLIT (Partitioning)	0.1037**	0.1165***	0.1181***	0.1191***	—	0.1364***	0.1156**	0.1272***		0.0905**	—	0.0989**
Double-clustered s.e.	(0.0490)	(0.0331)	(0.0414)	(0.0374)		(0.0395)	(0.0468)	(0.0390)	(0.0474)	(0.0365)		(0.0398)
Adjusted R-squared	0.091	0.225	0.257	0.278		0.281	0.1	0.199	0.235	0.283		0.282
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	826	826	826	826	821	779	413	413	413	413	408	393

The table reports Negative Binomial Maximum Likelihood (ML) estimates, associating civil conflict incidents with ethnic partitioning at the ethnic homeland level. The dependent variable in Panel A, B, C is the total number of battles, violence against civilians, battles resulting in territorial change at each ethnic homeland over the period 1997-2010, respectively. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. The specifications in columns (2)-(6) and (7)-(12) include region fixed effects (constants not reported). The specifications in columns (7)-(12) focus on country-ethnicity areas close to the national border (using as a cutoff the median distance from the centroid of each ethnicity-country homeland to the national border; 102 kilometers). The specifications in columns (5) and (11) exclude ethnic homelands where the dependent variable exceeds the 99th percentile. The specifications in columns (6) and (12) exclude ethnic homelands where capital cities fall. The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each ethnic homeland from the capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 18: Ethnic Partitioning and Civil Conflict.
Heterogeneous Effects A

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

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable is the total number of civil conflict incidents at each ethnic homeland within a country over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. All specifications include interactions terms of the SPLIT index with dummy variables identifying countries with high ethnic fractionalization (in (1)), linguistic fractionalization (in (2)), religious fractionalization (in (3)), cultural heterogeneity (in (4)), ethno-linguistic polarization (in (5), ethnic segregation (in (6)), large in terms of size, (in (7)), and landlocked countries (in (8)). With all measures we use the median value of the respective variable. All specifications include country fixed effects (constants not reported).

The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 19: Ethnic Partitioning and Civil Conflict. Heterogeneous Effects B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SPLIT (Partitioning)	0.8289***	0.9139***	0.9720***	0.9153***	1.1205***	1.2715***	0.8648***	1.4545***
	(0.2182)	(0.2503)	(0.2782)	(0.2295)	(0.2954)	(0.3326)	(0.2596)	(0.3684)
SPLIT (Partitioning) Interacted with								
High Rule of Law	-0.4397							
	(0.3677)							
Democracy Indicator		-0.8405***			-0.6958**			
		(0.3963)			(0.2727)			
High Democratic Capital			-0.7152***			-0.7032***		
			(0.2956)			(0.2106)		
High Income				-0.7244***	-0.6016**	-0.7195***		
				(0.3449)	(0.3358)	(0.2885)		
Executive Constraints at independence							-0.6346***	-0.6467***
							(0.2916)	(0.2679)
High Income at Independence								-1.0945***
								(0.3795)
Log Likelihood	-3601.185	-3596.972	-3597.876	-3597.585	-3593.267	-3592.267	-3599.011	-2914.9
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simple Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location & Geo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1182	1182	1182	1182	1182	1182	1182	1023

The table reports Negative Binomial Maximum Likelihood (ML) estimates associating civil conflict with ethnic partitioning at the country-ethnic homeland level. The dependent variable is the total number of civil conflict incidents at each ethnic homeland within a country over the period 1997-2010. SPLIT is an indicator variable that identifies partitioned ethnicities as those with at least 10% of the historical homeland falling into more than one contemporary country. All specifications include interactions terms of the SPLIT index with dummy variables identifying countries scoring high in a composite rule of law index (in (1)), classified as democracies (in (2) and (5)), scoring high in a democratic capital variable (in (3) and (6)), with high income (in (4)-(6)), with high executive constraints at independence (in (7)-(8)), and with high income in 1961 (in (8)). With all measures we use the median value of the respective variable. All specifications include country fixed effects (constants not reported).
The set of simple controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), and the log of population in 1960. The set of location controls includes the distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, and an indicator that takes on the value of one if a capital city falls in the homeland of an ethnic group within a country. The set of geographic controls includes an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 20: Summary Statistics for Data from the Demographic and Health Surveys (DHS)

variable	Obs.	mean	st. dev.	p25	median	p75	min	max
Composite Wealth Index	88124	3.162	1.433	2	3	4	1	5
Education	88123	1.732	1.554	0	1	3	0	5
Literacy Indicator	88124	0.491	0.500	0	0	1	0	1
Baseline Ethnic Partitioning Index (origin-based)	88124	0.429	0.495	0	0	1	0	1
Location-based Ethnic Partitioning Index	88124	0.445	0.497	0	0	1	0	1
Distance to the National Border	88124	0.103	0.093	0.033	0.079	0.143	0.000	0.564
Distance to the Sea	88124	0.466	0.338	0.162	0.432	0.751	0.000	1.716
Distance to the Capital	88124	0.306	0.252	0.127	0.271	0.428	0.000	1.902
Capital City Indicator	88124	0.257	0.437	0	0	1	0	1

The table reports descriptive statistics for all variables employed in the empirical analysisusing data from the Demograpic and Health Surveys (DHS). The data cover 20 countries. The Data Appendix gives detailed variable definitions and data sources.

Appenix Table 21: The Long-Run Effects of Ethnic Partitioning. Channels; Location and Identity
Sensitivity Analysis: Results with Literacy

	All Observations (households)						Observations (households) close to the National Border					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Partitioning - Idenity	-0.0847**	-0.0586**	-0.0373*	-0.0747**	-0.0522**	-0.0423*	-0.0533*	-0.0447*	-0.0379*	-0.0548*	-0.0463*	-0.0413**
Double-clustered s.e.	(0.0331)	(0.0251)	(0.0222)	(0.0341)	(0.0264)	(0.0231)	(0.0276)	(0.0243)	(0.0200)	(0.0285)	(0.0250)	(0.0204)
Partitioning - Location				-0.0308	-0.0199	0.0197				0.0062	0.0069	0.0154
Double-clustered s.e.				(0.0193)	(0.0155)	(0.0155)				(0.0257)	(0.0233)	(0.0187)
Adjusted R-squared	0.1970	0.2414	0.2610	0.1976	0.2417	0.2612	0.2433	0.2738	0.2881	0.2433	0.2738	0.2882
Observations	88124	86574	86574	88124	86574	86574	44525	43503	43503	44525	43503	43503
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 (linear probability model) estimates associating literacy with ethnic partitioning at the individual level. The dependent variable equals one when a respondent can read a sentence that has been presented to him by the DHS team in the respondent's language; the index equals zero when the respondent can neither read at all or can only read parts of the sentence. The ethnic partitioning index takes on the value of one for individuals that identify with a partitioned ethnicity and zero otherwise. The location based ethnic partitioning index takes on the value of one for individuals that currently reside in ethnic homelands that have been partitioned by the national border and zero otherwise.

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 (7)-(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.

Appendix Table 22: The Long-Run Effects of Ethnic Partitioning. Further Evidence on Channels
Sensitivity Analysis. Looking within Partitioned Ethnic Homelands; Ethnic-Homeland Fixed Effects Estimates

	Composite Wealth Index			Education			Literacy		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnic Partitioning	-0.2237*	-0.2119*	-0.2183**	-0.2291***	-0.2184***	-0.1949**	-0.0772***	-0.0742***	-0.0636***
Double-clustered s.e.	(0.1187)	(0.1202)	(0.1067)	(0.0809)	(0.0801)	(0.0791)	(0.0225)	(0.0219)	(0.0225)
Adjusted R-squared	0.182	0.198	0.204	0.204	0.252	0.255	0.270	0.289	0.294
Observations	39215	38371	38371	39215	38371	38371	39215	38371	38371
Country-Ethnic Homeland Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Location Controls	No	No	Yes	No	No	Yes	No	No	Yes

The table reports OLS estimates associating a composite wealth index (in columns (1)-(3)), education (in columns (4)-(6)), and literacy (in columns (7)-(9)) with ethnic partitioning at the individual level. All specifications are estimated in the sample of individuals residing in homelands of ethnicities that have been systematically partitioned by the national border.

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., 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 literacy indicator takes the value of one when a respondent can read a sentence that has been presented to him by the DHS team in the respondent's language; the index equals zero when the respondent can neither read at all or can only read parts of the sentence.

The ethnic partitioning index takes on the value of one for individuals that identify with a partitioned ethnicity and zero otherwise. All specifications include a vector of country-ethnic-homeland 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 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.