Appendix A

Description and Coding of African Territorial Disputes

The data set consists of territorial disputes between Africa states. A territorial dispute exists when states have incompatible territorial claims and these claims are pursued by central state governments using diplomatic and/or military means. In some cases, disputes were inherited from the colonial era, but to enter the data set, they had to be reiterated by the states after independence. Moreover, we sought to identify disputes that represented genuine incompatibilities, rather than technical issues arising from poorly delimited or demarcated borders. There are a number of instances in which states sought to clarify their borders, usually through the appointment of mixed commissions. We did not code such cases as disputes unless and until there was some diplomatic or military act signifying an incompatibility being pursued at the political, rather simply technical, level. Because of limitations on data dealing with militarized actions, we restricted the collection to disputes arising no later than 2001. This means that the conflict between Sudan and South Sudan, starting with the latter's independence in 2011, is not reflected in our data set, nor is a relatively recent dispute between Kenya and Uganda over islands in Lake Victoria.

For each case, we determine both the region of the dispute and the claimant(s). The claimant in a dispute is that state or states making a claim beyond the status quo at time of independence. The coding is thus unproblematic when the status quo was clear and corresponded to the division of effective administration at the time of independence. Two complications arise. First, the status quo may be unclear because of uncertainty over the colonial-era border. In that case, both states were considered claimants. Second, there

are a small number of cases in which administration at independence differed from the international border—e.g., Sudan's control of the Halaib triangle and Kenya's control over the Ilemi triangle. In these cases, where one state asserted administrative control over territory beyond the de jure border, we coded those states as the claimants. In the table below, the claimant state(s) are italicized.

Dyad	Years	Description of Dispute	Map Data	Source
Algeria-*Tunisia*	1962-1970	Tunisia claimed a triangle of territory east of the line running south from the boundary marker at Bir-Romane, Tunisia, north of the line running west from boundary maker 233 at Garet el-Hamel, Libya, and west of the provisional boundary line inherited from the French.	Straight line segments drawn from coordinates of boundary markers	Touval (1972, 251-55); Keesing's (1961, 18341-48).
Benin-Niger	1960-2005	Benin and Niger had a disputed possession of Lété island, which sits in the Niger river.	Claim lines surround Lété island.	Touval (1972, 281-2); Google Earth
Burkina Faso-*Ghana*	1963-1966	A dispute arose over the alignment of the far eastern section of the border, as Ghana encroached on territory thought to belong to Upper Volta (now Burkina Faso).	Based on fragmentary evidence, the claim line was drawn 5km deep into Burkina Faso's territory along an approximately 50 mile segment of the border in the north east Ghana.	Saffu (1970, 204); Waters (1969, 183); OAU (1964, 217).
Cameroon-*Nigeria*	1987-2002	Nigeria contested the delimitation in the Lake Chad region and claimed sovereignty over a number of villages on the now-dried up lake bed. Since the claims was based on de facto control over villages in question, there is no exact claim line.	A line encompassing the villages claimed by Nigeria.	ICJ (1999, 336-7)
Cameroon-*Nigeria*	1965-2006	Nigeria claimed the Bakassi Peninsula.	Claim line follows the physical contours of the Bakassi peninsula and coordinates of proposed border provided by Nigeria.	ICJ (1999, 409)
Cote d'Ivoire-*Ghana*	1960-1966	Ghana claimed the former Sanwi Kingdom in southeastern Ivory Coast.	Line following approximate border of former Sanwi Kingdom.	Handloff (1990, 76)[1]
Chad-*Libya*	1973-1994	Libya claimed a region known as the Aouzou strip on the basis of the 1935 Laval-Mussolini agreement.	Line following the boundary as set out in in the 1935 Laval-Mussolini agreement.	Brownlie (1972, 122)

Djibouti-*Eritrea*	1995-	Eritrea claimed a portion of northern Djibouti on the basis of the 1935 Laval-Mussolini agreement.	A straight line from Daadato to Der Eloua.	Google Earth; Mesfin (2008, 7)
Zaire-Zambia	1980-	A dispute arose over the location of the tripoint with Tanzania in Lake Tanganyika, leading to two versions of the straight line segment from Tanzania to Lake Mweru.	Straight lines drawn from the two versions of the Lake Tanganyika tripoint to the agreed point on Lake Mweru.	IBS, no. 51, p. 4.
Ethiopia-Eritrea	1998-	A dispute arose over different interpretations of colonial era treaties.	Claims lines based on maps generated by the United Nations' Eritrea-Ethiopia Boundary Commission.	Eritrea-Ethiopia Boundary Commission (2002, 14-16)
Ethiopia-Sudan	1966-2002	Ethiopia claimed two regions along the border known as the Fashqa and Umbrega triangles.	The Fasqa triangle is bounded by the Setit and Atbara rivers. The Umbrega triangle is bounded by the Setit and a straight line segment that heads towards Tod luk.	Taha (1975); river data from Digital Charts of the World.
Ghana-Togo	1960-	Ghana claimed all of Togo. Togo, in response, called for the unification of the territories of former German Togoland. In practice, this entailed a claim to parts of Ghana that had been in the British mandate following WWI.	The Ghananian claim follows boundary of Togo. The Togolese claim follows the historic boundary of German Togoland.	Paul Sprigade and Moisel (1920).[2]
Kenya-*Somalia*	1963-1981	Prior to Kenya's independence, Somalia claimed Kenya's Northern Frontier District. Upon independence, this region was reorganized and the Somali inhabited region became the North Eastern Province.	Claim line follows the internal administrative boundary of Kenya's North Eastern Province.	Global Administrative Areas (GADM) database

Kenya-Sudan	1963-	Kenya claims the Ilemi triangle, a region north of the straight line border drawn in 1914. While several alternative borders have been proposed, the Kenyan claim, and de facto control, typically extends to the red line boundary demarcated in 1938.	The claim line follows the red line boundary of 1938.	Google Earth
Mauritania -Morocco	1960-1979	Mauritania and Morocco both claimed Western Sahara.	Claim follows the borders and physical outline of Western Sahara.	Natural Earth base map
Mauritania-*Morocco*	1960-1970	Morocco claimed all of Mauritania	Claims line follows the boundaries of Mauritania	Natural Earth base map
Mali-Burkina Faso	1974-1987	Mali and Burkina Faso had conflicting claims to a region known as the Agacher Strip. Both countries are coded as claimants because, in its exhaustive review of the evidence, the ICJ could not determine a clear boundary upon independence.	The claim lines follow the lines each side presented to the ICJ.	International Court of Justice, Frontier Dispute Between Burkina Faso and Mali: Summaries of Judgments and Orders, map 1
Mali-Mauritania	1960-1963	Mali had claims in two places, both based on rejection of changes France made to the colonial border in 1944: (a) Mali claimed a portion of the Eastern Hodh desert. The north-south border in this region had never been well-defined, but the French moved it east (at Mali's expense) in 1944. The prior border was poorly defined, but the border agreed to in 1963 appears to closely approximate it. (b) Mali claimed that the border in western Hodh followed the bend of the Ouadou River between Djel Mael to Gueneibe. This had been the traditional border until a French order of 1944 shifted it south to the present location.	(a) Claim line follows the existing north-south international border, which was established by the 1963 Treaty of Kayes.[3] (b) Claim line follows the bend of the Ouadou River.	IBS, no 23, map 1.[4]

Morocco-Algeria	1962-1972	The border south of Figuig had not been fully delimited during the colonial period. Upon independence, Morocco claimed large swaths of southwest Algeria, including Tindouf and Colomb Bechar. These claims were renewed as soon as Algeria became independent in 1962.	Morocco's claim line encompasses the general contours of "Greater Morocco" from Figuig to the Algeria-Mali border.	Pennell (2000, 341)
Namibia-Botswana	1990-1999	A dispute arose over which channel of the Chobe river forms the boundary around Kasaikili/Sendudu island. Namibia claimed that the border was the southern channel, while Botswana claimed the northern channel.	The Namibia claim follows the southern channel of the river around the island.	Google Earth
Nigeria-Chad	1983-	A dispute arose over the ownership of islands in Lake Cad that were exposed when the lake level dropped. Although the primary problem appears to be one of demarcation, efforts to resolve the issue foundered, in part, on Nigeria's assertion that the boundary was never definitively delimited. Nigeria's claims were not clearly articulated, but in its dispute with Cameroon, it pointed to ambiguity in the location of the tripoint that serves as the terminus of this border.	The claim line is drawn to reflect a possible interpretation of the Nigeria-Chad-Cameroon tripoint that would be advantageous to Nigeria.	ICJ (1999, 341-410)
South Africa-*Lesotho*	1966-	Lesotho claimed "conquered territories" that had been taken from it in the 19th century. The claim was not well-defined but was understood to encompass parts of the former Orange Free State, Natal, and Eastern Cape Province. It included several named municipalities including Herchel and Matatiele.	The claim line in the north and west follows the approximate 1858 boundary between Lesotho and the Orange Free State. In the south, the claim line encompasses two regions of Eastern Cape and Natal Provinces identified as part of the dispute.	Brownlie (1972, 1108)
South Africa-*Swaziland*	1982-	Swaziland claimed the KaNgwane homeland and Ingwavuma district of KwaZulu homeland.	Claim line follows borders of these districts as they existed at the time of the claim (e.g. early 1980s).	King (2007, 16); Griffiths and Funnell (1991, 53)

Somalia-Ethiopia	1960-	Somalia claimed the Ogaden region of Ethiopia, a region inhabited by ethnic Somalis. The region did not have a precise geographic definition, but was instead defined ethnically. Today, it corresponds closely with Ethiopia's Somali Region.	Claim line follows the border of the Somali Region in Ethiopia.	GADM database
Sudan-*Egypt*	1958-	Egypt asserted that the border follows the 22nd parallel, giving it the Wadi Halfa salient and the Sudanese-administered Halaib triangle. Sudan claimed the Halaib triangle on the basis of colonial-era administration.	The claim line in the Halaib triangle follows the administrative border. In the Wadi Halfa salient, it follows the 22nd parallel.	IBS no. 18, map 1.[5]
Tanzania-Malawi	1967-	Tanzania claimed that the border follows the median line of Lake Nyasa from the point where the River Songwe meets the lake.	Median line of Lake Nyasa.	Mayall (1973, 613)
Uganda-Kenya	1976	Uganda asserted a claim to parts of western Kenya that had been transferred from the Uganda Protectorate by the British in 1902 and 1926.	Claim line follows the former border between the Uganda and East Africa Protectorates.	Huxley (1935, map facing p. 110).
Uganda-Tanzania	1974-1979	Uganda claimed the Kagera Salient, a patch of land south of its border and north of the Kagera River	Claim line follows the Kagera River in Tanzania.	Natural earth 10m Rivers and Lakes layer
Zambia-*Malawi*	1968-1986	Malawi claimed that its border with Zambia should be the Luanga River. Additional border flare ups have happened in this region, though it is not clear what their relation is to the larger claim.	Claim line follows the Luanga River in Zambia.	Natural earth 10m Rivers and Lakes layer

In addition to these disputes, we investigated several cases that are occasionally raised in the literature on African boundary problems. In each case, we determined that the case did not meet our criteria for inclusion.

Dyad	Dispute description and source	Reason for rejection	Source
Liberia-Guinea	Mount Nimba region (Waters 1969)	Upon Guinea's independence in 1958, Liberia renounced all claims to territory in Guinea, and in 1960, the two countries recognized their border. There may have been a demarcation issue, given that the mountain serves as the tripoint with Cote d'Ivoire.	Brownlie (1972, 305-9); Zartman (1969, 85)
Liberia-Cote d'Ivoire	Area between Cess and Cavally Rivers (Waters 1969)	There was no evidence of a dispute, and both countries recognized the border at independence in 1961.	Brownlie (1972, 359-69) ; Zartman (1969, 85)
Cote d'Ivoire-Guinea	The MID data set records a few incidents that took place in the Ivorian department of Biankouma in 1996.	After researching these incidents we were unable to identify claims made on the affected villages. These events appear to be isolated and the result of a poorly demarcated border.	African Research Bulletin (1995, 1996)
Kenya-Ethiopia	Gadaduma Wells (Waters 1969; Huth and Allee 2002)	This issue was resolved in principle in 1963, upon Kenya's independence. It was incorporated into a definitive border treaty in 1970, but there is no evidence of an active dispute in the interim.	Brownlie (1972, 824); IBS no. 152, 1975, pp. 3-5
Mozambique-Malawi	Eastern shore of Lake Shirwa (Waters 1969)	There is no evidence of a dispute, and Mozambique was not independent at the time it was alleged to have happened (1963). A redemarcation happened in that year, but it appears to have been undisputed.	Brownlie (1972, 1198-1212); IBS no. 112, 1971,p. 9.

Congo (Brazzavillle) -Gabon	Haut-Ogooué (Touval 1972; Waters 1969)	Although there were people in Haut-Ogooué and in the Congo who openly aspired for the region's return to Congo, we found no evidence that the claim was pursued by the Congolese government after independence. When the issue flared in 1962, due to soccer riots, Gabon alleged that Congo sought to annex the region, but Congo did not openly make a claim.	Touval (1972, 196-8); Weinstein (1966, 220-24); Brownlie (1972, 657)
Cameroon-Gabon	Rio Muni (Waters 1969)	Cameroon and Gabon both expressed interest in the Rio Muni region, which later became the continental portion of Equatorial Guinea. Although some Cameroonians considered the region to be part of their country, there is no evidence that the claim was officially advanced by state leaders.	Weinstein (1966, 230).
Sudan-Chad	845-mile border region (Waters 1969)	There appears to be no official dispute between these countries, although local issues have at times cropped up, requiring more precise demarcation.	Brownlie (1972, 638-9)
Nigeria-Benin	Yoruba area of Benin (Waters 1969)	There is no evidence of a dispute. Brownlie (1972, 188) explicitly mentions that the splitting of the Yoruba did not create problems because people were able to move freely across the border.	Brownlie (1972, 188)

[1] Map available online at http://www.cartoko.com/content/wp-content/uploads/2010/05/CotedIvoire_1990_p006.jpg. Accessed 1/25/2013.
[2] Available on line at http://en.wikipedia.org/wiki/File:Togo_Deutsches_Koloniallexikon,_Verlag_von_Quelle_%26_Meyer_Leipzig.jpg. Accessed 1/25/2013.
[3] This is the one case in which the existing border departs substantially from the border inherited upon independence. Hence, the currently recognized border is the claim line, and the inherited border is no longer in existence.
[4] Available at http://www.law.fsu.edu/library/collection/LimitsinSeas/maps/bs23asmall.html. Accessed 1/25/2013.
[5] Available online at http://www.law.fsu.edu/library/collection/limitsinseas/maps/bs18.html. Accessed 2/5/13.

Works Cited

Brownlie, Ian. 1979. *African Boundaries: A Legal and Diplomatic Encyclopaedia*. London: C. Hurst & Company.

Griffiths, Ieuan Ll, and D. C. Funnell. 1991. "The Abortive Swazi Land Deal." *African Affairs* 90 (358): 51-64.

Handloff, Robert E., ed. 1990. *Cote d'Ivoire: A Country Study*. Washington, DC: Library of Congress.

Huth, Paul K., and Todd L. Allee. 2002. *The Democratic Peace and Territorial Conflict in the Twentieth Century*. Cambridge: Cambridge University Press.

Huxley, Elspeth. 1935. *White Man's Country: Lord Delamere and the Making of Kenya*. Vol. 1. London: MacMillan and Co., Limited.

International Court of Justice. 2002. Case concerning the land and maritime boundary between Cameroon and Nigeria: Judgment of 10 October 2002. *ICJ Reports.*

International Court of Justice. 1999. "Case Concerning the Land and Maritime Boundary between Cameroon and Nigeria: Counter-Memorial of the Federal Republic of Nigeria." , http://www.icj-cij.org/docket/files/94/8602.pdf.

Keesing's. 1961. *Keesing's Record of World Events* 7 (September): 18341-8.

King, Brian H. 2007. "Developing KaNgwane: Geographies of Segregation and Integration in the New South Africa." *Geographical Journal* 173 (1): 13-25.

Mayall, James. 1973. "The Malawi-Tanzania Boundary Dispute." *The Journal of Modern African Studies* 11 (4): 611-28.

Mesfin, Berouk. 2008. The Eritrea-Djibouti border dispute. *Institute for Strategic Studies Situation Reports* (15 September 2008), http://www.issafrica.org/uploads/SITREP150908.PDF.

Organization of African Unity. 1964. *Assembly of the Heads of State and Government of the O.A.U: Addresses Delivered at the Assembly*. Cairo: Information Department.

Paul Sprigade, Bearbeitet, and Max Moisel. 1920. *Togo*. Deutsches Koloniallexikon Von 1920. Leipzig: Verlag von Quelle and Meyer.

Pennell, C. R. 2000. *Morocco since 1830: A History*. London: Hurst & Co.

Saffu, E. O. 1970. "The Bases of Ghana-Upper Volta Relations during the Nkrumah Regime." *Canadian Journal of African Studies* 4 (2): 195-206.

Taha, Faisal Abdel Rahman Ali. 1975. *The Settlement of the Sudan-Ethiopia Boundary Dispute*. Khartoum: Khartoum University Press.

Touval, Saadia. 1972. *The Boundary Politics of Independent Africa*. Cambridge, Mass.: Harvard University Press.

Waters, Robert. 1969. "Appendix 1: Inter-African Boundary Disputes." In *African Boundary Problems*, ed. Carl Gosta Widstrand, pp. 183-4. Uppsala: Scandinavian Institute of African Studies.

Weinstein, Brian. 1966. *Gabon: Nation-Building on the Ogooué*. Cambridge, MA: M.I.T. Press.

Wild, Patricia Berko. 1966. “The Organization of African Unity and the Algerian-Moroccan Border Conflict: A Study of New Machinery for Peacekeeping and for the Peaceful Settlement of Disputes among African States.” *International Organization* 20 (1): 18-36.

Zartman, I. William. 1969. “The Foreign and Military Politics of African Boundary Problems.” In *African Boundary Problems*, ed. Carl Gösta Widstrand, pp. 79-100. Uppsala: The Scandinavian Institute of African Studies.

Appendix B

Sources, Coding, Summary Statistics, and Bivariate Relationships

The base map used to identify border segments is the 10m boundary line map produced by Natural Earth.[1] All calculations involving lengths or distances used an equidistant conic projection defined over the African continent; calculations involving area used an Albers equal area conic projection, also defined over Africa.

1. Ethnic Variables

The ethnic variables used in the main text were derived from two different sources. The first is the Soviet atlas *Narodov Mira* (Bruk and Apenchenko 1964), which was digitized by Weidmann, Rød, and Cederman (2010) into the data set Georeferencing of Ethnic Groups (GREG). GREG differentiates Arabs on the basis of country, creating separate polygons for Sudanese, Tunisian, Egyptian, Moroccan, Algerian, Libyan, and Western Saharan Arabs. In the tests reported in the paper, we aggregated these into a single Arab group. In appendix C, we report results without this change. We also report results using an alternative ethnic map produced by Murdock (1959).

[1] All maps from Natural Earth are in the public domain and available at www.naturalearthdata.com. The Kenya-Sudan border was changed to follow the straight east-west line established as the provisional boundary in 1914 (see Brownlie 1979, pp. 917-21). Border lines were split into 1km segments using a tool in ET GeoTools (http://www.ian-ko.com/), an add-on to ArcMap.

The second source is the Ethnic Power Relations family of datasets (Vogt et al. forthcoming), particularly the geo-located data in GeoEPR (Wuchperfenning et al. 2011) and the identification of transborder ethnic kin (TEK) in neighboring states (Cederman et al. 2013).[2] Two sets of changes were needed to make these data appropriate for our study:

1. Four countries—Somalia, Lesotho, Swaziland, and Burkina Faso—were coded as having no politically relevant ethnic groups. In these countries, the EPR identify a single group that accounts for either all (Somalis, Sotho, Swazi) or most (Gur in Burkina Faso) of the population. They are nonetheless coded as politically "irrelevant" presumably because, given the hegemonic position of these groups, these ethnic identities do not comprise politically salient cleavages domestically. They are, however, politically salient internationally. For this reason, we recode these groups as politically relevant, and code their power status as "hegemonic." For the purposes of geo-locating these groups (which are omitted from GeoEPR), the Somali, Sotho, and Swazi were mapped as encompassing the entire territory of their respective countries. For the Gur in Burkina Faso, we used *Ethnologue* and the associated *World Language Mapping System*, to identify regions inhabited by groups in the Gur language family.

2. In three countries—Liberia, Zimbabwe, and South Africa—EPR aggregates all of the African ethnic groups at the time of independence into undifferentiated umbrella categories: Indigenous Peoples, Africans, and Blacks. This coding reflects the nature of the political cleavages in those countries at the time, and it is only later that EPR identifies politically salient cleavages corresponding to the ethnic groups within these larger aggregations. This coding rule makes sense given the primary purpose of the EPR data, but it is less appropriate when thinking

[2] The 2014 versions of these data sets were accessed using Luc Girardin, Philipp Hunziker, Lars-Erik Cederman, Nils-Christian Bormann, and Manuel Vogt. 2015. *GROW*^{up} - Geographical Research On War, Unified Platform. ETH Zurich. http://growup.ethz.ch/.

about cross-border ethnic attachments, as ethnic groups in neighboring states can presumably identify the areas inhabited by their kin within these umbrella groups. For example, the Sotho in Lesotho did not see Blacks in South Africa as an undifferentiated mass of ethnic kin, as the TEK data imply; presumably, they could identify the regions where fellow Sotho lived. Thus, for the purposes of locating border segments that partition ethnic kin, we break these umbrella groups into their constituent parts, relying on polygons corresponding to later time periods, when these groups are coded as distinct. Each constituent group still retains the power status of its umbrella group at the time of independence (i.e., discriminated).

From these two data sets, several variables were created. When a variable was unique to one of the sources, we indicate that below in parentheses.

Partition: This variable indicates whether or not the same ethnic group was located on both sides of the border segment. For the GREG data, this variable captures segments that have the same group on both sides of the border.[3] For the EPR data, this variable captures whether the border segment separated a politically relevant ethnic group from transborder kin, based on the GeoEPR map. Since both GREG and EPR use the national borders from a different base map—the now-discontinued Digital Charts of the World (DCW)—we first identified partitioning border segments on that map and then joined them to the segments in our base map on the basis of proximity.

[3] Note that the GREG map identifies some mixed regions, containing two ethnic groups, so a small number of border segments partition two groups.

In the main text, we do not restrict the extent of the ethnic overhang required to consider a group partitioned; in Appendix C, we report tests with minimum thresholds for how much a group's area has to be on each side of the border for the group to be considered partitioned.

Partition and 1st Leader (GREG, Murdock): This variable indicates whether a segment partitions the ethnic group that produced the head of state of the claimant country at the time the dyad became independent. Leaders' ethnicity was identified using data from Fearon, Kasara, and Laitin (2007). These ethnicities were then matched to GREG (and Murdock) groups through additional research by the authors. In some instances, leaders were attached to GREG groups based on the location of their birthplace or other circumstantial evidence. In a few cases, a match could not be found, but we could rule out the possibility that the leader's group was partitioned by a border.

Partition and national capital (GREG, Murdock): This variable indicates whether a segment partitions the ethnic group whose homeland includes the capital of the potential claimant state in the directed dyad.

Partition and In Control (EPR): This variable indicates whether a segment partitions an ethnic group whose power status in the claimant state was hegemonic, dominant, or monopoly at the time of the dyad's birth.

Partition and In Coalition (EPR): This variable indicates whether a segment partitions an ethnic group whose power status in the claimant state was senior or junior partner at the time of the dyad's birth.

Partition and Powerless in Target (EPR): This variable indicates whether a segment partitions an ethnic group whose power status in the target state was powerless, discriminated, or irrelevant at the time of the dyad's birth.

Partitioned group population share: For segments that partition a group, two variables measure the population of the group as a percentage of total population in the claimant and target states, respectively. For the GREG (and Murdock) data, group and national populations were estimated using spatially coded 1960 population data from provided by the United Nations Environment Programme/Global Resource Information Database (UNEP/GRID), Sioux Falls.[4] For each GREG (and Murdock) group in each country, the group's population share was assumed to equal the share of the country's population residing in the group's geographic homeland. Population shares for EPR groups are reported in that data set.

Border fractionalization of claimant: This variable counts the number of ethnic groups partitioned by all of the borders of the potential claimant state. Versions of this variable were created for the GREG, Murdock, and EPR groups.

[4] These data are available at http://na.unep.net/siouxfalls/datasets/datalist.php.

Ethno-linguistic fractionalization: This variable measures the ethnic fractionalization of the potential claimant state. It runs from zero to one, with higher scores corresponding to more diverse states. The data are from Fearon and Laitin (2003).

2. Realist Variables

Capability share of claimant: This variable measures the material capabilities of the claimant state as a fraction of total dyadic capabilities. Data are from the Correlates of War National Material Capabilities data set (Singer, Bremer, and Stuckey 1972).[5] The capability scores are measured in the year of the dyad's independence (e.g., the later of the two independence dates). Additional measures were taken at 1, 5, and 10 years after independence.

Area share of claimant: This variable measures the share of total dyadic land area that belongs to the claimant state. Country areas were determined from the base map,

Distance from capital of claimant: The variables measures the distance in kilometers from the claimant's capital city to the border segment (logged). An analogous variable was created for the target.

Oil in target: There are two variables. The first is a dummy variable indicating whether there are any oil deposits located within 1000km of the border segment and on the side of the potential target state in the directed dyad. The second indicates the distance from the border segment to

[5] Available at http://correlatesofwar.org

the closest oil deposit in the target, if one exists. Data are from Lujala, Rod, and Thieme (2007). Note that this coding does not take into account the year that the oil deposit was discovered, as this is not known for all observed deposits. Hence, the variable indicates whether there was ever an oil deposit near the border.[6]

Minerals in target: As with oil, there are two variables, one indicating whether a mineral deposit exists in the target within 1000km of the border segment, the second indicating the distance to the closest such deposit. Data are from the Mineral Resource Data System (MRDS), a coding of metallic and nonmetallic mineral resources produced by the U.S. Geological Survey (USGS). [7] As above, this coding does not take into account the timing of the discovery.

Area of partitioned group in target (claimant): This variable indicates, for each segment that partitions a group, the percentage of the target (claimant) country's land area that is covered by the partitioned group. Areas determined from GREG, GeoEPR, and the base map.

Minerals in partitioned group in target: This is a dummy variable indicating, for each segment that partitions a group, whether the homeland of the group in the target state includes at least one mineral deposit.

3. Institutional Variables

[6] This and all other distance measures were obtained by creating a "near table" in ArcMap, which reports the distance of the closest feature(s) in one layer (in this case, oil deposits) to each of the features in the target layer (the 1km segments).
[7] Available at http://tin.er.usgs.gov/mrds/.

Rivers: Segments that follow rivers were identified using the Natural Earth 1:10m rivers layer, information from the International River Boundaries Database at Durham University,[8] and maps and descriptions in Brownlie (1979). A few borders follow lakes. The results do not change if these are coded as river boundaries or if a separate indicator for lake segments is included in the models.

Watersheds: Segments that follow watershed boundaries were identified using HydroSHEDS (Lehner et. al 2006) and descriptions in Brownlie (1979).

Straight: Straight line segments were identified by eye and by consulting sources on the boundary alignment, including Brownlie (1979) and the International Boundary Study series. A complication arises because almost every piece of a border is straight when viewed over a short enough distance. Moreover, linking together a number of very short straight lines can yield a border that is not straight overall. To address this, a subset of the border was considered to be straight if it followed a straight line for at least 50km, though in a few cases a smaller threshold was used. A 1km segment was coded as straight if it was part of a straight subset.

Demarcated: A dummy variable indicates, for border segments that did not follow a river or lake, whether they were physically demarcated at the time of independence. This determination was based on information in Brownlie (1979).

[8] Available at https://www.dur.ac.uk/ibru/resources/irbd/.

Border age: For each border segment, we coded the year in which the line inherited on independence was defined, based on information in Brownlie (1979) and the US State Department's *International Boundary Study* series. In most cases, the creation of a boundary happens in stages, which might include a rough statement of spheres of influence, followed by a more precise definition of a boundary, followed by successive stages of clarification, rectification, and demarcation. For each case, we selected the year when the inherited boundary was largely defined and recognizable in an official agreement or decree. Neither rough versions done in the course of separating spheres of influence, nor minor rectifications or clarifications were considered as determining the age of the boundary. Since there is some subjectivity in determining the border age, we distinguish only between borders that were essentially defined prior to 1920 and those that were defined thereafter. As noted in the text, there are a few dyads in which all or part of the border was undefined or only provisionally defined at the time of independence. These are indicated separately.

Border change in colonial period: This is a dummy variable that indicates whether a border segment underwent a significant change during the colonial period. Minor rectifications (e.g., arising from errors or efforts to more closely align the border with the physical geography) or more precise definitions of rough lines that had originally separated spheres of influence were not counted as changes. Rather, this variable identifies cases in which a line defined in an earlier agreement or decree was subsequently changed in a later agreement or decree, in a way that transferred an appreciable piece of territory. For each case, we also code which state lost territory due to the change. In a few cases, mostly in French West Africa, we only know that the

border shifted a number of times before finally settling in the inherited location, in which case both states are coded as losing territory in a change.

4. Summary Statistics

Summary statistics for all variables in the directed-dyadic data set are as follows:

Table B1: Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Claim on segment	153370	0.11	0.32	0	1
Partition (GREG)	153370	0.74	0.44	0	1
Partition (EPR)	153370	0.31	0.46	0	1
Partition and 1st leader (GREG)	153370	0.19	0.39	0	1
Partitioned group pop. share in claimant (GREG)	153370	0.2	0.32	0	0.99
Partitioned group pop. share in target (GREG)	153370	0.2	0.32	0	0.99
Partition and in control (EPR)	153370	0.089	0.29	0	1
Partition and in coalition (EPR)	153370	0.072	0.26	0	1
Partition and powerless in Target (EPR)	153370	0.14	0.35	0	1
Border fractionalization of claimant (GREG)	153370	9.57	4.86	1	19
Border fractionalization of claimant (EPR)	153370	3.17	1.12	1	6
Ethnic fractionalization of claimant	152904	0.67	0.22	0.036	0.93
Capability share of claimant	153370	0.5	0.31	0.0048	1
Area share of claimant	153370	0.5	0.25	0.011	0.99
Distance to capital of claimant (logged)	153370	6.27	0.77	-0.68	7.59
Oil on target side	153370	0.43	0.5	0	1
Distance to oil on target side	153370	166.7	251.2	0	1000
Minerals on target side	153370	0.92	0.27	0	1
Distance to minerals on target side	153370	157.3	164.3	0	990
Minerals in partitioned group in target (GREG)	153370	0.51	0.5	0	1
Partitioned group area share in target (GREG)	153370	0.2	0.27	0	1
Partitioned group area share in claimant (GREG)	153370	0.2	0.27	0	1
Minerals in partitioned group in target (EPR)	153370	0.25	0.43	0	1
Partitioned group area share in target (EPR)	153370	0.12	0.24	0	1
Partitioned group area share in claimant (EPR)	153370	0.12	0.24	0	1
River border	153370	0.29	0.46	0	1
Watershed border	153370	0.069	0.25	0	1
Straight line segment	153370	0.31	0.46	0	1
Demarcated prior to independence	153370	0.35	0.48	0	1
New border created after 1920	153370	0.15	0.36	0	1
Border undefined or provisional	153370	0.033	0.18	0	1
Border changed in colonial period	153370	0.17	0.38	0	1
Claimant lost territory in colonial period	153370	0.11	0.31	0	1
Border length (logged)	153370	6.85	0.61	4.14	7.71

5. Bivariate Relationships

Table B2 presents the frequency of claims as a function of the main dichotomous variables in the data set using directed dyadic data. In **Table B2** we see that segments that partition a GREG- or an EPR-, but not a Murdock-group appear slightly more likely to be the subject of a claim than those that do not, but any effect is very small.[9] A much stronger relationship emerges when we condition on whether the segment partitions a group that produced the first leader of a GREG-group or is an EPR-group that is politically in control in the claimant. Segments that have both of these qualities experienced a claim between three and four times the rate of segments that had neither. In the EPR data, segments that partition groups that were part of a ruling coalition were much less likely to be subject of a claim although that pattern does not survive in the multivariate analysis. Segments that partition EPR groups that are powerless in the target, in contrast, were much more likely to be the subject of a claim. The presence of minerals on the target side in the area of a partitioned ethnic group is correlated with a higher risk of a claim; however, this result does not survive the multivariate analysis. Oil is relatively rare – only nine dyads have any border segments that are close to oil deposits – and its presence is associated with slightly fewer disputed segments. Straight line segments are contested at 2.5 times the rate of non-straight segments, while segments that follow rivers or watersheds are much less likely to be contested. A border change in the colonial period is associated with a higher frequency of claims, an effect

[9] Since the number of 1km segments is large (over 76,000), even small differences are statistically significant. But since these comparisons do not take into account the interdependence among segments within dyads, this table is mainly useful for assessing the magnitude and direction of the reported effects.

driven entirely by changes that created a loss for the potential claimant in the directed dyad. Another striking relationship deals with the status of the border at independence: a claim is five times more likely if the border was undefined or provisional at the time of independence. The directed-dyadic setup actually understates this effect, since a claim may be made by one state but not the other. If we turn to the dyadic level and ask whether either state made a claim on the segment, we find that almost all (roughly 95%) of all border segments that the parties had not agreed upon before the birth of the dyad were also disputed afterwards. As we suggest in the text, this suggests that there is some endogeneity in this relationship.

Table B2: Summary of Bivariate Relationships

Variable	Variable=1 % in data	Variable=0 % disputed	Variable=1 % disputed
GREG partition	73.8	9.6	12
GREG partition and 1st leader	18.5	6.6	32.5
EPR partition	30.6	10.3	13.9
EPR partition & in control in Claimant	8.9	9.2	33.3
EPR partition & in coalition in Claimant	7.2	12.2	1.5
EPR partition & powerless in Claimant	14.4	11.9	8.1
EPR partition & powerless in Target	14.4	10	19.6
Murdock partition	93.2	13.0	11.3
Murdock partition and 1st leader	6.5	10.8	17.8
Oil in Target	43.1	13.4	8.7
Minerals in Target	91.9	9.7	11.5
Minerals in EPR partitioned group in target	25.1	11.2	11.9
Watershed border	6.9	12.1	1.3
River border	29.5	13.9	5.4
Straight line border	30.7	7.9	19.2
Demarcated at independence	34.8	11	12.2
Border undefined or provisional	3.3	10.2	47.1
Claimant lost territory in colonial period	10.5	10.6	18.4
Same colonial parent	44.9	8.5	14.9

Additional Works Cited

Brownlie, Ian. 1979. *African Boundaries: A Legal and Diplomatic Encyclopaedia*. London: C. Hurst & Company.

Bruk, Solomon Ilich, and V. S. Apenchenko, eds. 1964. *Atlas Narodov Mira*. Moscow: Glavnoe upravlenie geodezii i kartografii.

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

Fearon, James D., Kimuli Kasara, and David D. Laitin. 2007. "Ethnic Minority Rule and Civil War Onset." *American Political Science Review* 101 (01): 187-93.

Lehner, B., K. Verdin, K., and A. Jarvis. 2006. *HydroSHEDS Technical Documentation*. Washington, DC: World Wildlife Fund US. Available at http://hydrosheds.cr.usgs.gov.

Lujala, Paivi, Jan Ketil Rod, and Nadja Thieme. 2007. "Fighting Over Oil: Introducing a New Dataset." *Conflict Management and Peace Science* 24 (3): 239-56.

Singer, J. David, Suart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." In *Peace, War, and Numbers*, ed. Bruce Russett, pp. 19-48. Beverly Hills, CA: Sage.

Appendix C

Additional Tests

This appendix performs three tasks. First, we delve deeper into some of the results of the main text, emphasizing a more thorough analysis of the effects implied by the interaction terms. Second, we present a set of models that operationalize our key independent variables in different ways. Third, we analyze the main models from the article with two sets of alternative estimators – spatial lag models and fixed effects models – to subject our hypotheses to the most exacting tests. We present the results from these models here rather than in the main text because each suffers from some significant shortcomings of their own, as we discuss in more depth below. For convenience, we present a brief table of contents:

1. Additional Analysis of Ethnic Models 1 & 5 in the main text

To assess the explanatory power of the ethnic model we can establish different baselines. First, we can compare segments that partition ethnic groups with segments that do not partition groups. Second, we can compare segments that partition ethnic groups by the characteristics of those sections, such as the size of the partitioned group in either the claimant or the target state, and whether the partitioned group produced the country's first leader. While we by and large focus on the second sets of comparisons in the main text, here we delve deeper into the first basic comparison. We do so by examining the linear combinations of the relevant variables.

a) Main text – Table 1

For ease of comparison we reproduce Table 1 from the main text below as **Table C1a**, and complement that with **Table C1b** which reports the respective linear combinations.

Table C1a. Ethnic Politics Models – reproduced from the main text

VARIABLES	(1) GREG	(2) GREG	(3) GREG	(4) EPR	(5) EPR	(6) EPR
1.Partition	0.23	-0.96*	-0.08	0.38	-1.23	0.46
	(0.44)	(0.52)	(0.45)	(0.44)	(0.96)	(0.89)
Partitioned group's status in claimant						
2.First leader		1.35***	-1.52			
		(0.41)	(0.95)			
3.Population share		1.59**	-4.90***			
		(0.80)	(1.87)			
4.First leader * population share			8.69***			
			(2.17)			
5.In control					2.10**	2.13**
					(0.89)	(0.84)
6.In coalition					-1.58	-1.67
					(1.32)	(1.32)
Partitioned group's status in the target						
7.Powerless in target					1.94**	0.53
					(0.85)	(0.84)
8.Population share		-0.10	-0.45			-2.97***
		(0.84)	(0.88)			(0.93)
Border fractionalization of claimant			-0.11**		-0.73***	-0.75***
			(0.05)		(0.23)	(0.23)
Border length (logged)	0.77	0.62*	0.69*	0.79	1.05*	1.12*
	(0.52)	(0.37)	(0.38)	(0.54)	(0.63)	(0.64)
Constant	-7.61**	-6.55**	-6.03**	-7.66**	-7.49*	-7.95*
	(3.70)	(2.64)	(2.64)	(3.59)	(4.11)	(4.18)
Observations	153,370	153,370	153,370	153,370	153,370	153,370
χ^2 statistic	2.20	14.68	30.60	4.43	50.05	59.24
Pseudo R-squared	0.03	0.16	0.23	0.03	0.18	0.19

Standard Errors, corrected for clustering on dyads, are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table C1b. Ethnic Politics Models - Linear Combinations

Linear Combinations	**(2) GREG**	**(3) GREG**	**(5) EPR**	**(6) EPR**
Linear Combination 1 – 3 |	1.58**			
Population share = .75	(0.66)			
Linear Combination 1 – 4 |		1.25*		
Population share = .75		(0.68)		
Linear Combination 1 & 5			0.86*	2.60***
In control in claim, not powerless in target			(0.45)	(0.53)
Linear Combination 1 & 6			-2.81**	-1.21
In coal. in claim, not powerless in target			(1.41)	(1.31)
Linear Combination 1 & 7			0.71	0.99*
Powerless in claimant & target			(0.61)	(0.59)
Linear Combination 1, 5 & 7			2.80***	3.12***
In control in claim, powerless in target			(0.81)	(0.76)
Linear Combination 1, 6 & 7			-0.87	-0.68
In coalition in claim, powerless in target			(1.14)	(1.18)
Linear Combination 1 & 8	-1.03	-0.39		-1.61*
Pop. share in target = .7	(0.64)	(0.57)		(0.92)

Note: Standard errors, corrected for clustering on dyads, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.10

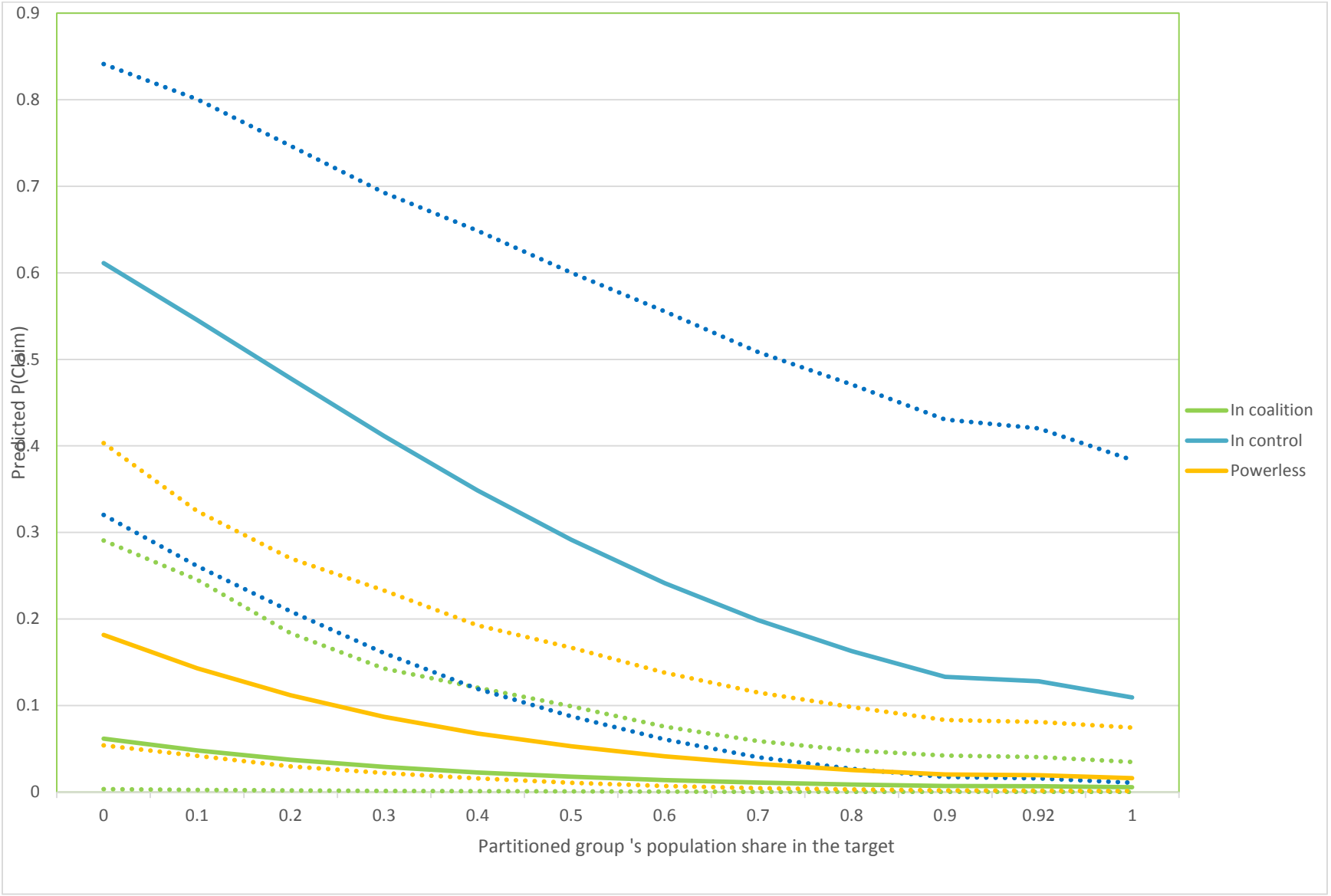
When we focus on the linear combinations of the variables that measure partition and the relevant (partitioned) ethnic group variable(s), the coefficients measure the combination's effect as compared to segments that do not partition a group. There is consistent evidence that segments that partition ethnic groups that are large and powerful in the claimant are significantly more likely to be disputed. (We should note that in Model (3) for partitioned groups that c *less* than 70% of the population in the claimant, the linear combination of Partition, First leader, Population share and First leader * Population share fails to reach significance at the 10% level; p < .107.) The effect is particular strong if the segment partitions an ethnic group that is also powerless in the target. In spite of

this result, the evidence for Hypothesis 1c which conditions the likelihood of a claim on the status of the partitioned group in the target is mixed.[1] In model 2 the linear combination of Partition and Partitioned group's share of the population in the target becomes negative and significant at the 0.1 level for groups that constitute less than 60% of the population in the target. In the EPR models, once the group is larger than 70% in the target (model 6) the linear combination becomes significant, which holds in less than 15% of the cases where the segment partitions a group. However, the other linear combinations that directly tap into the political status of the group in the target (Partitioned and powerless) fail to reach significance.

To capture the effect of the partitioned group's share of the population in the target – another potential way to capture political power and influence (Cederman et al. 2013) – we plot the group's population share in the target against the predicted probability of a claim for all three potential scenarios, where the partitioned group is 1) in control in the claimant (in blue), 2) in coalition in the claimant (in green), and 3) powerless in the claimant (in orange). The color-matched dotted lines represent the 95% confidence interval for each scenario. We see that for every level of the partitioned group's population share, the predicted probability of a claim differs only between partitioned groups that were in control and partitioned groups that were in coalition. Note that the predicted probability of a claim when the segment does *not* partition a group is 0.07 with a confidence interval [0.04, 0.12].

[1] We do find that the partitioned group's status in the target has the expected negative effect on the likelihood a segment is contested if we set a high threshold – a minimum of 30 percent of the group's area is on either side of the border – for considering a group to be partitioned.

Figure C1: EPR data: Predicted probability of a claim, given population share in the target and political status in the claimant



b) Main text – Tables 5a and 5b

We next switch to a re-examination of the results of **Tables 5a** and **5b** in the main text. For ease of interpretation, we split these into **Table C2a**, which reproduces the GREG results from **Table 5a** and adds the linear combination. **Table C2b** reproduces the results from the EPR models and **Table C2c** presents the linear combinations of the EPR models. In **Table C2a** we see that the linear combinations of the variables that measure partition and the relevant (partitioned) ethnic group variable(s) indicate that, as expected, the dynamic of partition, access to power through the first leader and partitioned group's share of the population in the claimant significantly increases the likelihood a segment will be claimed early on in the life of the dyad, compared to a segment that does not partition. However, when no claim is made early on, this constellation of factors significantly *decreases* the likelihood a partitioned segment will be disputed later on. In **Table C2c** with the EPR data we see a similarly expected pattern, the linear combinations show that when a segment partitions an ethnic group that is politically in control in the claimant state, that segment is significantly more likely to be claimed, both in the Full and the Early dispute samples, whether or not the partitioned group is powerless in the target (linear combinations 1 & 2 for the not powerless in the target, linear combinations 1, 2 & 4 for powerless groups in the target). Evidence for Hypothesis 1c is mixed since partitioning a powerless group in the target is significant only for groups that are powerless or in control, but not when it is in coalition in the claimant.

Table C2a: Combined GREG Models (Table 5a in main text)

VARIABLES	(1) GREG-Full	(2) GREG-Early	(3) GREG-Late
Partition	0.08	0.68	-0.67
	(0.47)	(0.57)	(0.49)
Partition and 1st leader	-2.38**	-2.58*	-16.19**
	(1.17)	(1.32)	(7.29)
Partitioned group population share in claimant	-5.87**	-10.19***	-2.19*
	(2.63)	(3.26)	(1.26)
Partition and 1st leader * population share	11.38***	16.52***	18.38**
	(3.09)	(3.78)	(9.33)
Linear Combination 1-4 |	1.83***	2.86***	-4.72***
Population share = .75	(0.60)	(0.70)	(1.16)
Partitioned group population share in target	-1.68	-0.69	-2.97
	(1.10)	(1.34)	(2.61)
Border fractionalization of claimant	-0.13*	-0.11*	-0.17
	(0.07)	(0.06)	(0.12)
Area share of claimant	0.34	1.33	-2.40*
	(0.83)	(0.91)	(1.39)
Oil on target side	-4.14***	-4.77***	-3.08**
	(1.09)	(1.45)	(1.38)
Distance to oil on target side	0.01***	0.01***	0.00
	(0.00)	(0.00)	(0.00)
Minerals on target side	-0.32	3.11**	-2.37**
	(1.22)	(1.43)	(1.04)
Distance to minerals on target side	-0.00***	-0.00**	-0.01
	(0.00)	(0.00)	(0.00)
Minerals in partitioned group in target	0.47	0.20	1.48***
	(0.42)	(0.46)	(0.37)
Partitioned group area share in target	1.02	0.10	1.11
	(0.83)	(0.83)	(1.47)
River border	-1.09**	-0.68	-2.47***
	(0.49)	(0.54)	(0.81)
Watershed border	-2.70***	-2.23***	
	(0.69)	(0.80)	
Straight line segment	0.06	-0.36	0.46
	(0.39)	(0.42)	(0.76)
Border at independence undefined or provisional	4.39***	4.91***	
	(0.75)	(0.82)	
Claimant lost territory in colonial period	2.17***	2.67***	1.20
	(0.73)	(0.83)	(1.37)
Observations	153,370	153,370	120,962
Chi2 statistic	138.02	166.24	164.57
Pseudo R-squared	0.40	0.46	0.38

Standard errors, corrected for clustering on dyads, are reported in parentheses. Constant and variable for

border length included but not reported. *** p<0.01, ** p<0.05, * p<0.1

Table C2b: Combined EPR Models (Table 5b in main text)

Variables	(1) All Disputes	(2) Early Disputes	(3) Late Disputes
1. Partition	0.63	0.94	-18.15***
	(1.12)	(1.11)	(2.52)
2. Partition and in control in claimant	1.76*	1.80*	1.02
	(0.97)	(1.01)	(1.79)
3. Partition and in coalition in claimant	-2.11	-2.06	
	(1.43)	(1.74)	
4. Partition and powerless in target	1.96**	1.38	18.12***
	(0.97)	(0.93)	(1.56)
Border fractionalization of claimant	-0.83**	-1.43***	0.34
	(0.42)	(0.54)	(0.23)
Area share of claimant	-0.58	1.38	-4.06***
	(1.50)	(1.87)	(1.36)
Oil on target side	-3.71***	-5.11***	-2.17*
	(1.07)	(1.60)	(1.17)
Distance to oil on target side	0.00***	0.01***	0.00
	(0.00)	(0.00)	(0.00)
Minerals on target side	0.45	3.25**	-0.19
	(1.06)	(1.39)	(1.35)
Distance to minerals on target side	-0.00**	-0.00*	-0.01*
	(0.00)	(0.00)	(0.00)
Minerals in partitioned group in target	-1.54*	0.01	-2.79***
	(0.79)	(0.89)	(1.08)
Partitioned group area share in target	-0.53	-3.10	6.23***
	(1.62)	(2.31)	(1.60)
River border	-1.05**	-0.56	-2.99**
	(0.48)	(0.47)	(1.16)
Watershed border	-2.84***	-2.63***	
	(0.83)	(0.87)	
Straight line segment	0.64	0.58	0.55
	(0.40)	(0.52)	(0.73)
Border undefined or provisional	3.46***	4.40***	
	(0.69)	(0.79)	
Claimant lost territory in colonial period	1.44	1.87*	1.34
	(0.89)	(1.09)	(1.45)
Observations	153,370	153,370	111,485
χ^2 statistic	164.35	172.17	
Pseudo R^2	0.36	0.42	0.34

Note: Standard errors, corrected for clustering on dyads, are reported in parentheses. Constant and variable for border length included but not reported. *** p<0.01, ** p<0.05, * p<0.1

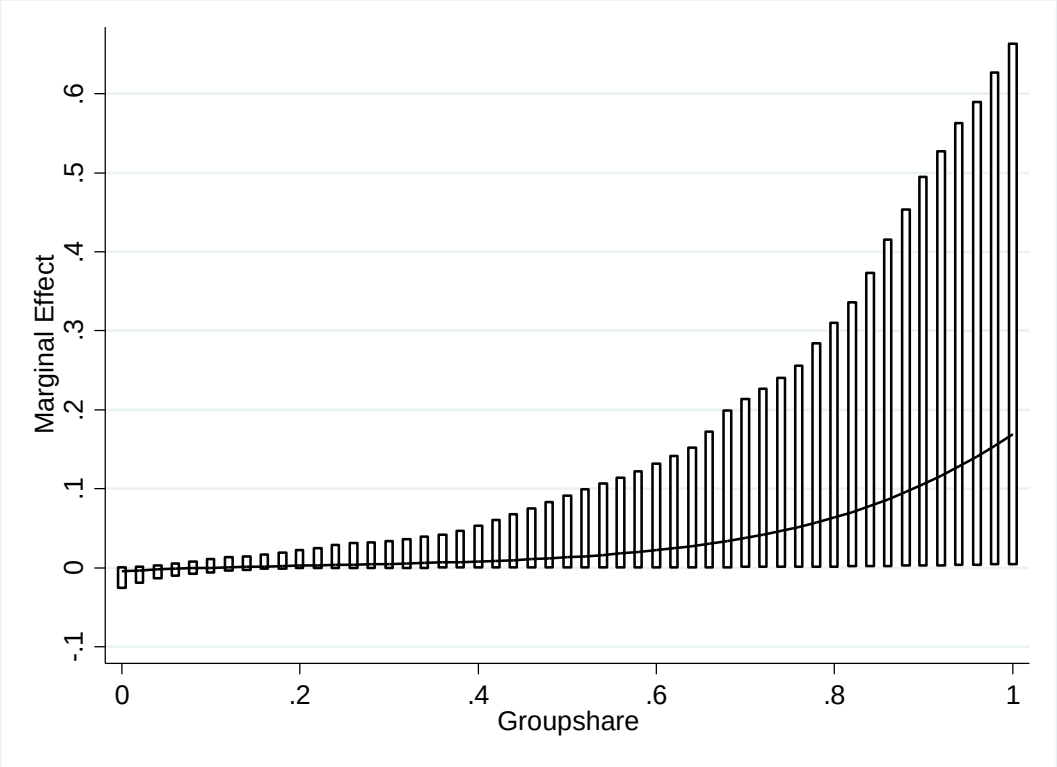
Table C2c: Combined EPR models – linear combinations

Variables	(1) EPR-Full	(2) EPR-Early	(3) EPR-Late
Linear Combination 1 & 2	2.39***	2.74***	-17.13***
	(0.81)	(0.71)	(1.42)
Linear Combination 1 & 3	-1.48	-1.12	--
	(1.79)	(1.96)	
Linear Combination 1 & 4	2.59***	2.32**	-0.03
	(0.67)	(0.92)	(1.31)
Linear Combination 1, 2 & 4	4.35***	4.12***	0.99
	(1.02)	(1.00)	(1.01)
Linear Combination 1, 3 & 4	0.47	0.25	--
	(1.38)	(1.66)	

Note: Standard errors, corrected for clustering on dyads, are reported in parentheses. Constant and variable for border length included but not reported. *** p<0.01, ** p<0.05, * p<0.1

In **Figure C2** we present the marginal effect of whether the segment partitions a group that produced the country's first leader, conditional on the partitioned group's share of the claimant's population from model 2 (the Early disputes) in **Table C2a**. The results are again very similar to the results of **Figure 3** in the main text. Segments that partition a group that produced the first leader are more likely to be contested (relative to segments that partition groups that did not produce the first leader) as the group's population share increases. Specifically, the marginal effect of partitioning the first leader's group is significantly different from zero once the group exceeds one-half of the country's population.

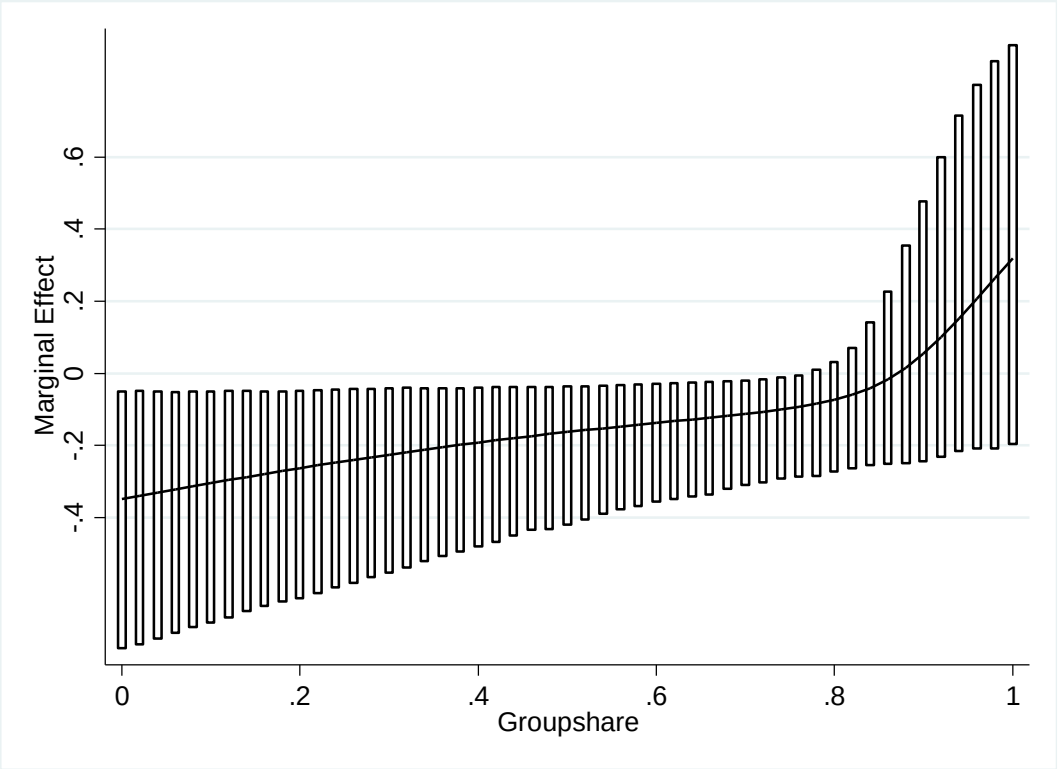
Figure C2 – *Early* Disputes: The marginal effect of 1st leader and group share



In **Figure C3** we present the effect of 1st leader conditional on the partitioned group's share of the population in the claimant state from model 3 in **Table C2a** (the Late disputes). The marginal effect of

producing the 1st leader is *negative* and significant as long as the group constitutes *less* than roughly 80% of the population in the claimant.

Figure C3: ***Late*** **disputes, the marginal effect of 1st leader and group share**



2. Alternative specifications of the key explanatory variables

a) Alternative indicator for political access

In **Table C3,** we offer an alternative operationalization of an ethnic group's political power in the GREG data, replacing the *1st leader* variable with an indicator for whether the ethnic group's homeland included the capital of the country. Note that that these two indicators are very highly correlated (.798), so we do not include both variables in the same regressions. The table replicates the model specifications in Table

1 using this indicator, but we also include one model that substitutes the ELF fractionalization measure for our border fractionalization measure. To capture the overall effects of the partitioned group's associated variable, we again need the linear combination of *Partition* and the other Partition variables; we set the *Partitioned group's population share* in the claimant at .93, which holds in over 25% of the cases where a partitioned group controlled the capital. At lower levels of population share, the linear combination is not significant. As expected, the results are similar, specifically, the linear combinations are significant at the same level as before and of similar magnitude.

Table C3: Ethnic politics, with capital group as an indicator of political access

VARIABLES	(1) GREG	(2) GREG	(3) GREG	(4) GREG	(5) GREG
1.Partition	-0.60	-0.46	-0.44	-0.11	-0.02
	(0.40)	(0.41)	(0.39)	(0.41)	(0.45)
2.Partition and in capital	1.82***	1.27**	1.12**	-1.00	-1.19
	(0.56)	(0.54)	(0.44)	(1.03)	(1.13)
3.Part. group pop. share in claimant				-4.95**	-4.92**
				(2.24)	(2.14)
4.Partition and capital * pop. share				7.29***	7.67***
				(2.54)	(2.54)
Linear Combinations |	1.22**	0.80	0.68	0.99*	1.27*
Models 4 & 5: Pop. share=.9	(0.54)	(0.53)	(0.50)	(0.59)	(0.75)
Border fractionalization of claimant			-0.15***	-0.12**	-0.12**
			(0.05)	(0.05)	(0.05)
ELF fractionalization of claimant		-2.13**			
		(1.07)			
Border length (logged)	0.71	0.73*	0.80**	0.78*	0.82**
	(0.45)	(0.44)	(0.40)	(0.40)	(0.40)
Partitioned group pop. share in target					-0.48
					(0.85)
Constant	-7.17**	-5.94*	-6.44**	-6.51**	-6.85**
	(3.16)	(3.14)	(2.77)	(2.80)	(2.78)
Observations	153,370	152,904	153,370	153,370	153,370
Chi2 statistic	10.57	12.08	12.34	20.09	21.70
Pseudo R-squared	0.11	0.14	0.16	0.17	0.18

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

b) Disaggregating Arabs

As noted in the text, the GREG data set breaks the Arabs of northern Africa into separate groups depending on their country, leading to separate polygons for Egyptian, Sudanese, Libyan, Algerian, Moroccan, and Western Sahara Arabs. In our main tests, we merged all of these groups into a single Arab group, which was consequently treated as being partitioned by the borders between these countries. Here, we use the original GREG coding, which treats these groups, for the most part, as not partitioned. To calculate the linear combination we again set the level of the *Partitioned group's population share* in the claimant at .7, which holds for roughly a third of the observations where the partitioned group produced the first leader. **Table C4** re-estimates the specifications from Table 1, column (3) and Table 5, column (1), with these new variables. We drop the consistently insignificant "group population share in the target" variable. The effects of the ethnic variables are if anything more significant.

Table C4: Results with Arab Groups Disaggregated

VARIABLES	(1) Basic Ethnic Model	(2) Full Model
1.Partition	0.25	0.13
	(0.48)	(0.40)
2.Partition and 1st leader	-2.95**	-3.71***
	(1.43)	(1.33)
3.Partitioned group population share in claimant	-7.84*	-7.33*
	(4.23)	(3.75)
4.Partition and 1st leader * population share	13.11***	14.41***
	(4.76)	(4.19)
Linear Combination 1-4 |	0.99*	1.37***
Pop. share=.7	(0.53)	(0.47)
Border fractionalization of claimant	-0.14**	-0.11
	(0.06)	(0.07)
Area share of claimant		-1.55
		(1.25)
Oil on target side		-3.95***
		(0.81)
Distance to oil on target side		0.00***
		(0.00)
Minerals on target side		-0.47
		(1.29)
Distance to minerals on target side		-0.00***
		(0.00)
Minerals in partitioned group in target		0.35
		(0.38)
Partitioned group area share in target		0.96
		(0.91)
River border		-1.18**
		(0.50)
Watershed border		-3.14***
		(0.68)
Straight line segment		0.23
		(0.36)
Border undefined or provisional		3.51***
		(0.61)
Claimant lost territory in colonial period		1.85***
		(0.69)
Constant	-6.52**	-6.45**
	(3.08)	(2.80)
Observations	153,370	153,370
χ^2 statistic	26.37	178.43
Wald test	0.00	0.00
Pseudo R^2	0.21	0.38

Note: Standard errors, corrected for clustering on dyads, are reported in parentheses. Variable for border length included but not reported. *** p<0.01, ** p<0.05, * p<0.1

Although this does not constitute an alternative coding of the explanatory variables, but rather a different sample, to further check for the potential effects of regional variation and the potential confounding effects of the widespread Arab population in North Africa, we replicate the analysis in **Tables 5a** and **5b** on a sample from Sub-Saharan countries. The first two columns in **Table C5a** report the results based on GREG codings, while the second two columns use ethnic variables derived from EPR. For easier comparison we report the results from **Table 5** in the main text in models 1 and 3. To save space, some variables that are consistently insignificant have been dropped. In both samples, the results remain broadly as before. One notable change is that minerals have a more pronounced effect in the sub-Saharan Africa sample, with a very clear jump in the probability of a claim when there are minerals in the target and the probability dropping as those minerals are farther from the border. The other results, including the linear combinations remain similar. In the GREG sub-Saharan model (2) the linear combination of partition, First leader, Population share, and First leader * Population share is significant at the $p < 0.005$ level when the partitioned group makes up more than 65% of the overall population in the claimant; in the Full model (1), at this level of population share the linear combination is significant at $p < .014$. We report the linear combinations of the sub-Saharan EPR models in **Table C5b** below. In contrast to the Full sample, the linear combination of partition and in control no longer is significant at conventional levels. The other linear combinations remain similar in their sign and significance.

Table C5a: Combined Models—Sub-Sahara

Variables	(1) GREG-Full	(2) GREG-SSA	(3) EPR-Full	(4) EPR-SSA	Variables
Partition	0.08	0.07	0.63	-0.80	Partition
	(0.47)	(0.68)	(1.12)	(1.23)	
First leader	-2.38**	-3.99***	1.76*	2.18**	In control
	(1.17)	(1.34)	(0.97)	(0.98)	
Partitioned group pop. share	-5.87**	-3.66	-2.11	-2.29	In coalition
	(2.63)	(3.08)	(1.43)	(1.67)	
First leader * pop. share	11.38***	12.58***			
	(3.09)	(3.54)			
Group pop. share in target	-1.68	-0.64	1.96**	2.72**	Powerless in target
	(1.10)	(1.26)	(0.97)	(1.06)	
Border fractionalization	-0.13*	-0.01	-0.83**	-1.44**	
	(0.07)	(0.06)	(0.42)	(0.59)	
Area share of claimant	0.34	0.30	-0.58	2.27	
	(0.83)	(1.06)	(1.50)	(1.91)	
Oil on target side	-4.14***	-3.87***	-3.71***	-5.83**	
	(1.09)	(1.29)	(1.07)	(2.37)	
Distance to oil on target side	0.01***	0.00*	0.00***	0.01*	
	(0.00)	(0.00)	(0.00)	(0.00)	
Minerals on target side	-0.32	2.73*	0.45	3.39**	
	(1.22)	(1.53)	(1.06)	(1.49)	
Distance to minerals in target	-0.00***	-0.00*	-0.00**	-0.00	
	(0.00)	(0.00)	(0.00)	(0.00)	
Minerals in partitioned group	0.47	0.49	-1.54*	1.03	
	(0.42)	(0.35)	(0.79)	(0.90)	
Group area share in target	1.02	-0.60	-0.53	-2.08	
	(0.83)	(0.74)	(1.62)	(2.12)	
River border	-1.09**	-1.34***	-1.05**	-1.03**	
	(0.49)	(0.49)	(0.48)	(0.43)	
Watershed border	-2.70***	-3.19***	-2.84***	-3.26***	
	(0.69)	(0.78)	(0.83)	(0.80)	
Straight line segment	0.06	-0.49	0.64	-0.02	
	(0.39)	(0.51)	(0.40)	(0.47)	
Border undefined or provisional	4.39***	2.53***	3.46***	2.64**	
	(0.75)	(0.83)	(0.69)	(1.19)	
Claimant lost territory	2.17***	2.96***	1.44	2.83***	
	(0.73)	(0.87)	(0.89)	(0.99)	
Observations	153,370	114,748	153,370	114,748	
χ^2 statistic	138.02	93.69	164.35	97.94	

Wald test	0.00	0.00	0.00	0.00	
Pseudo R^2	0.40	0.38	0.36	0.42	

Note: Standard errors, corrected for clustering on dyads, are reported in parentheses. Constant and variable for border length included but not reported. *** p<0.01, ** p<0.05, * p<0.1

Table C5b: EPR models (3 & 4) – linear combinations

Variables	(1) EPR-Full	(2) EPR-SSA
Linear Combination 1 & 2	2.39***	1.38
	(0.81)	(0.85)
Linear Combination 1 & 3	-1.48	-3.09
	(1.79)	(2.02)
Linear Combination 1 & 4	2.59***	1.92**
	(0.67)	(0.77)
Linear Combination 1, 2 & 4	4.35***	4.10***
	(1.02)	(0.97)
Linear Combination 1, 3 & 4	0.47	-0.36
	(1.38)	(1.69)

Note: Standard errors, corrected for clustering on dyads, are reported in parentheses. Constant and variable for border length included but not reported. *** p<0.01, ** p<0.05, * p<0.10

c) More restrictive codings of partition

In our base model, a segment was coded as partitioning a group if it cut through the polygon corresponding to the group's homeland. This rule does not distinguish groups that have substantial overhang on each side of the border from those that are merely "shaved" by the border. Given the uncertainty associated with mapping group areas, we might want to insist that a sizable portion of the group extend into another state in order to consider it partitioned. To do this, we calculate, for each partitioned group, what percentage of the group's area was on either side of the border. A group that is perfectly split in half by the border has 50 percent of its area in each state, while a group that is merely shaved by the border will have almost all of its area in one state and very little in the other. We can then impose more restrictive codings of partition by requiring that the percentage of the group's area that falls in each state exceed some minimum threshold. **Table C6** replicates the model specification in Table 1, column (3), after imposing increasingly strict criteria for a group to be considered partitioned. Minimum thresholds of 1, 5, 10, 20, and 30 percent were used, as indicated. The table also indicates, in each column, the percentage of segments that are coded as partitioning a group given the threshold. It is clear from the table that, with one exception, none of the main results depend on how strict a coding of partition we use, although the magnitude and significance of the main effect—the linear combination of *Partition*, *First leader*, *population share*, and the *interaction*—increases as the partition criterion becomes more stringent. Note that the border fractionalization score, which counts the number of groups partitioned by the claimant country's borders, also changes with the minimum threshold, and its effects remain largely unchanged as well, except in model (3) where the variable obtains a p-value of .102. The table yields one strikingly different result in model (5): when we only consider groups to be partitioned when their area constitutes more than 30

percent in either state, the larger the partitioned group's population share in the target – and thus presumably the more powerful in the target – the less likely the segments that partition that group will be disputed. This is an intriguing and suggestive finding, which confirms Hypothesis 1c, and fits our findings with the EPR-data.

Table C6: The Effect of Changing the Criterion for Group Partition

	(1)	(2)	(3)	(4)	(5)
VARIABLES	1 percent	5 percent	10 percent	20 percent	30 percent
1.Partition	-0.06	0.24	0.56	0.44	0.49
	(0.43)	(0.45)	(0.43)	(0.55)	(0.74)
2.Partition & 1st leader	-1.58	-1.52	-1.48	-1.36	-0.49
	(0.97)	(0.95)	(0.92)	(1.09)	(1.68)
3.Part. group pop. share in claim	-4.56**	-4.49**	-4.25**	-3.50*	-3.09*
	(1.85)	(1.82)	(1.73)	(1.84)	(1.60)
4.Part. & 1st leader * pop. share	8.68***	8.35***	8.07***	7.57***	6.12**
	(2.13)	(2.12)	(2.05)	(2.32)	(2.54)
Linear Combination 1-4 |	1.25*	1.42**	1.76**	1.92***	2.12***
Pop. Share =.7	(0.68)	(0.70)	(0.72)	(0.64)	(0.76)
Border fractionalization claimant	-0.10**	-0.13**	-0.13	-0.16*	-0.41**
	(0.05)	(0.07)	(0.08)	(0.09)	(0.17)
Part. Group pop. share in target	-0.57	-0.60	-0.71	-0.87	-3.57***
	(0.94)	(0.96)	(0.97)	(1.07)	(1.23)
Border length (logged)	0.65*	0.67*	0.67*	0.63*	1.01
	(0.37)	(0.37)	(0.38)	(0.34)	(0.66)
Constant	-5.86**	-6.08**	-6.33**	-5.89**	-7.75*
	(2.60)	(2.63)	(2.66)	(2.33)	(4.15)
Percent of segments that partition	70.8	64.6	57.6	38.3	21.8
Observations	153,370	153,370	153,370	153,370	153,370
Chi2 statistic	32.50	33.59	34.35	28.77	37.09
Pseudo R-squared	0.23	0.23	0.23	0.18	0.17

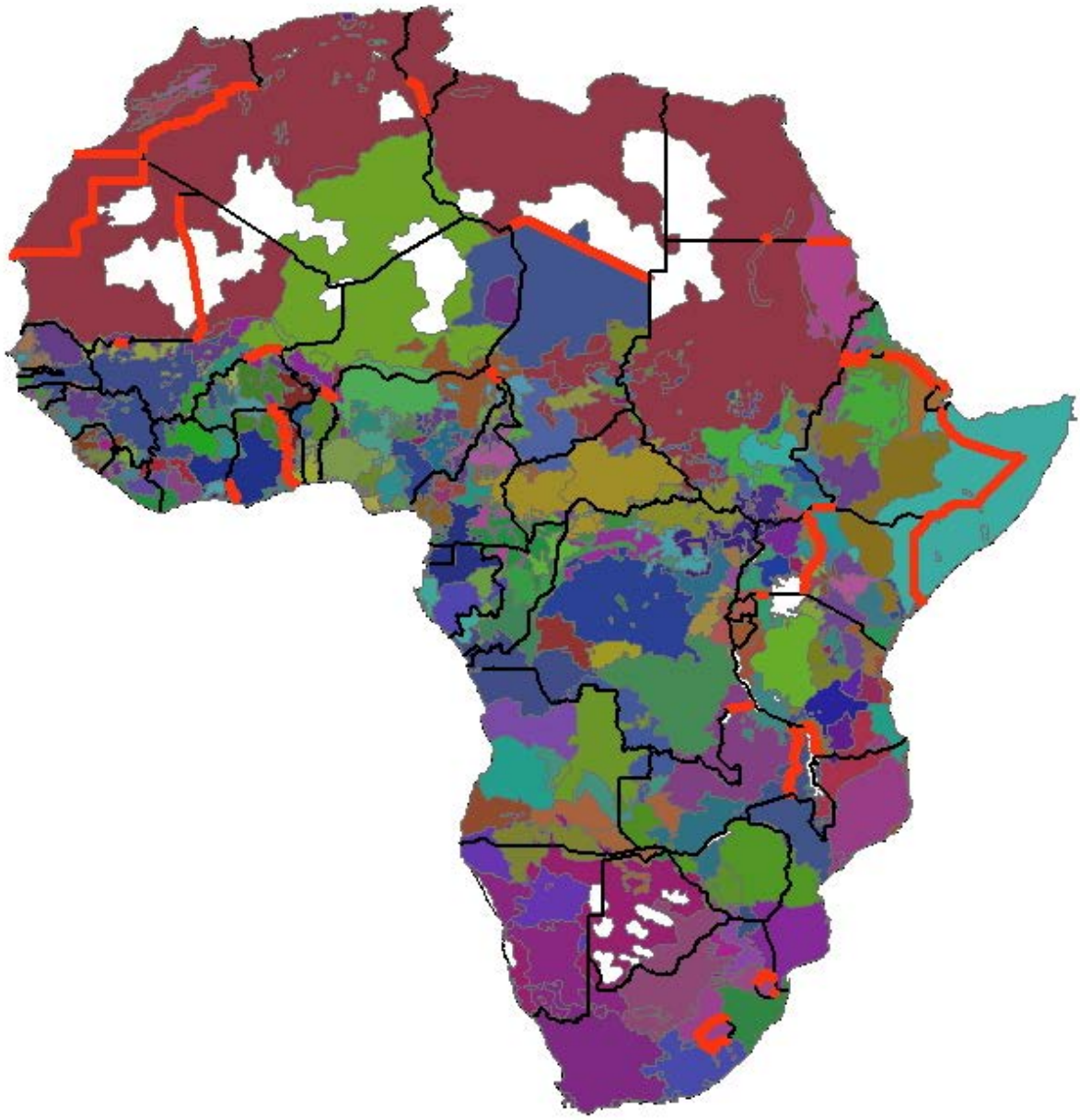
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

d) Alternative map of ethnic groups

We next replicate the analyses of the main text in the paper replacing the Soviet Atlas *Narodov Mira* geo-referenced as GREG – presented in **Figure C4** – with a map of Africa tribes produced by Murdock's (1959). Murdock's map, digitized by Nathan Nunn, is shown **Figure C5**. A quick comparison with the GREG map in **Figure C4** shows that the Murdock alternative has more, smaller groups, so some of the large groupings that appear in GREG (e.g., Arabs, Somalis) are broken into smaller units. As a result, comparably fewer countries appear to be ethnically homogeneous. The Murdock map was also not snapped to country borders in the same way that GREG map was, leading to more small overhangs. As a results of these two differences, using the Murdock map leads to comparably more segments that partition a group (93 percent vs 74 percent) and fewer segments that partition a group that produced the first leader (5.6 percent vs. 19.8 percent). To deal with the possibility that map inaccuracies overstate the incidence of partition, we estimated the models with these data using several different criteria for partition, as described above. To calculate the linear combination, we set Partitioned group's population share in the claimant at .4, which holds for over a quarter of the cases where a partitioned group in the Murdock data produced the first leader. **Table C7** estimates the model specification in Table 1, column (4), using variables created using the Murdock map. The first column imposes no threshold on the size of the overhang required for a segment to be coded as partitioning a group; the next four columns impose increasingly strict minimum thresholds, as indicated. Overall, the results obtained using this map are similar to those obtained with the GREG data, particularly when map inaccuracies are reduced by eliminating tiny overhangs. The linear combination of *Partition, First leader, Partitioned group population share* and the *triple interaction* is consistently positive, and it is significant at increasing levels in models (3) - (5), that is, as we set

more stringent partitioning requirements the linear combination becomes significant at the 0.10, 0.05 and 0.01 level. In addition, border fractionalization, now calculated with the new map, continues to have a significant negative effect on the likelihood of a claim. When we reproduce **Figure 3** from the main text, the figure obtained with the Murdock map looks very similar with similarly sloping predicted probabilities of making a claim. However, the confidence intervals of *Partition* and *1st leader* consistently overlap with those of *Partition* and **not** *1st leader*. The reason is that there are very few partitioned 1st leaders identified with the Murdock map.

Figure C4: GREG – Soviet Atlas *Narodov Mira*



Note: In this map, the Arab groups in north Africa have been combined into a single group.

Figure C5: Murdock's (1959) map

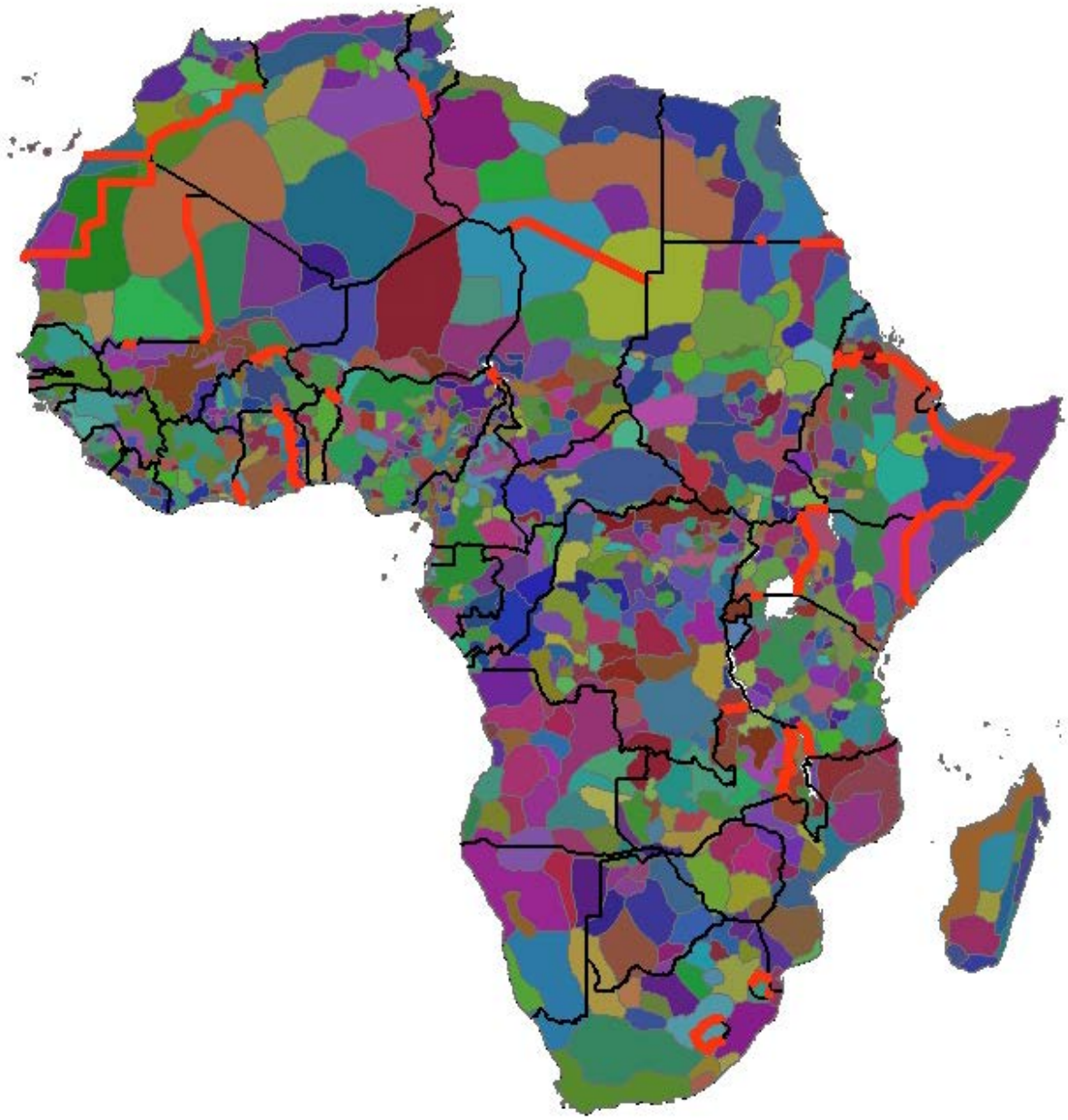


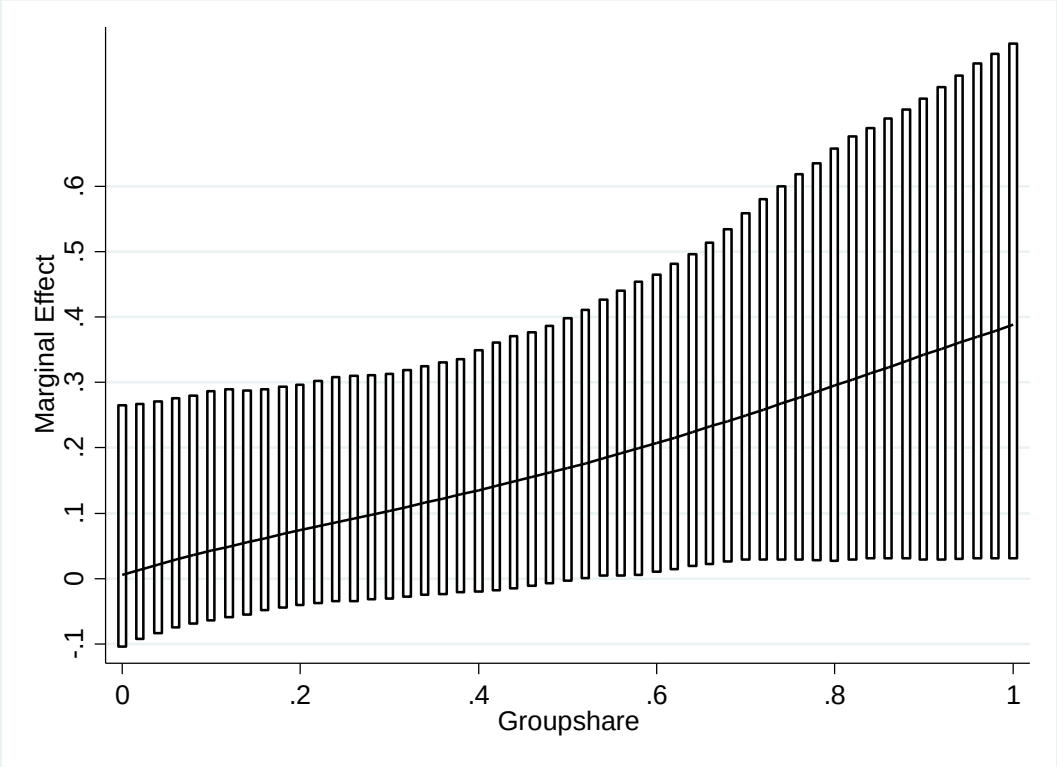
Table C7: The Ethnic Model using the Murdock Map

VARIABLES	(1) 0 percent	(2) 1 percent	(3) 5 percent	(4) 10 percent	(5) 20 percent
Partition	0.05	0.33	0.77**	0.60	0.97**
	(0.52)	(0.46)	(0.38)	(0.39)	(0.42)
Partition & 1st leader	-0.79	-0.79	-0.48	-0.26	-0.29
	(1.05)	(1.10)	(1.08)	(0.93)	(0.95)
Part. Group pop. share in claimant	-2.68	-3.59	-3.29	-3.30	-1.45
	(2.44)	(2.57)	(2.55)	(2.48)	(2.32)
Partition & 1st leader * pop. share	5.15*	5.96*	4.97	5.39*	4.13
	(3.09)	(3.22)	(3.13)	(2.97)	(2.77)
Linear Combination 1-4 |	0.25	0.49	0.96	1.18**	1.75***
Pop. Share =.4	(0.67)	(0.67)	(0.61)	(0.54)	(0.55)
Border fractionalization claimant	-0.07**	-0.09**	-0.13**	-0.14**	-0.11*
	(0.03)	(0.04)	(0.05)	(0.06)	(0.06)
Part. Group pop. share in target	-1.62	-1.57	-1.43	-0.56	-0.49
	(2.08)	(2.02)	(2.06)	(1.45)	(1.34)
Border length (logged)	1.00*	0.99*	1.04**	1.05*	0.98*
	(0.54)	(0.55)	(0.53)	(0.54)	(0.55)
Constant	-7.46**	-7.60**	-8.03**	-8.07**	-8.49**
	(3.61)	(3.57)	(3.39)	(3.52)	(3.60)
Percent of segments that partition	93.2	90.1	81.8	71.3	49.7
Observations	153,370	153,370	153,370	153,370	153,370
Chi2 statistic	25.64	24.41	25.04	38.21	40.34
Pseudo R-squared	0.11	0.12	0.14	0.13	0.11

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In **Figure C6** we rely on Model 4 (the 10 percent sample) to graph the marginal effect of whether the partitioned group produced the first leader (at the birth of the dyad), conditional on share of the population. We see that the effect (whether the first leader is a member of the partitioned group or not) is statistically significant at the 5% level when the partitioned group constitutes more than 50% of the population in the claimant. (We report the marginal effects for the GREG sample in **Figure C7.**)

Figure C6: Murdock (Model 4): marginal effects of switching *1st leader* and group share



f) *The conditional effects of ethnic heterogeneity*

We next examine the conditional effects of ethnic heterogeneity in more detail. **Figure C7** combines **Figure 3** from the main text of the paper with a graph that presents the marginal effects of shifting from a partitioned group that produced the claimant country's first leader to a partitioned group that did not, as a function of the group's share of the population in the claimant.

Figure C7: The predicted (left) and marginal effect (right) of switching whether a partitioned group produced the claimant country's first leader

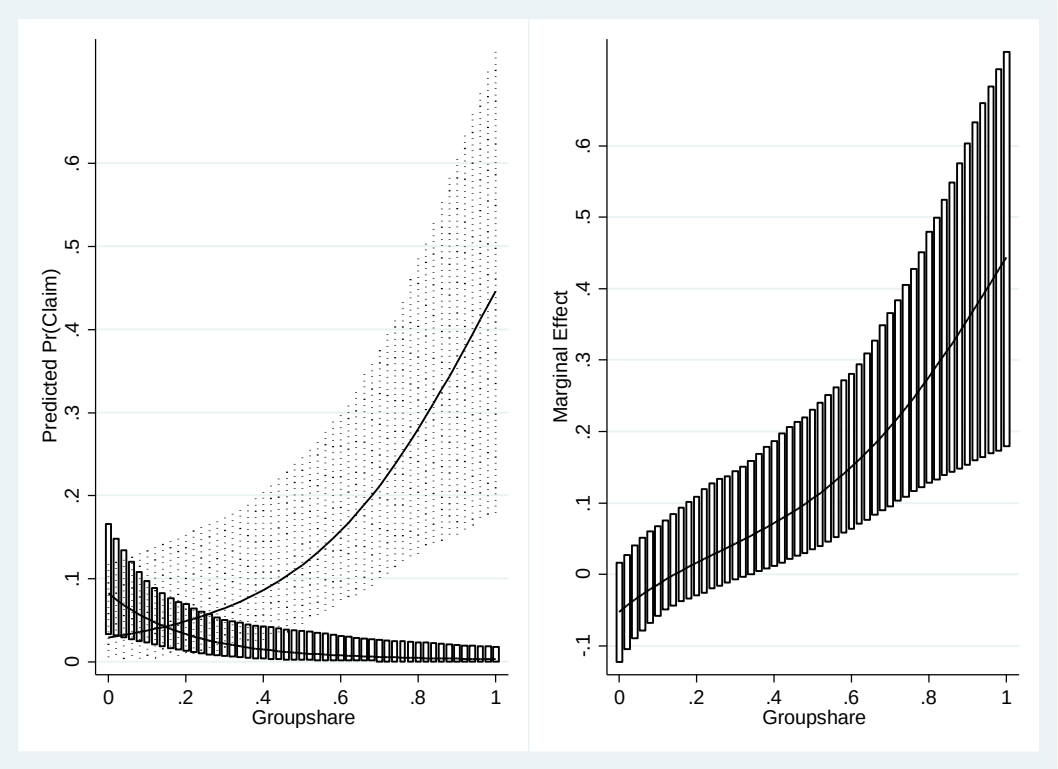
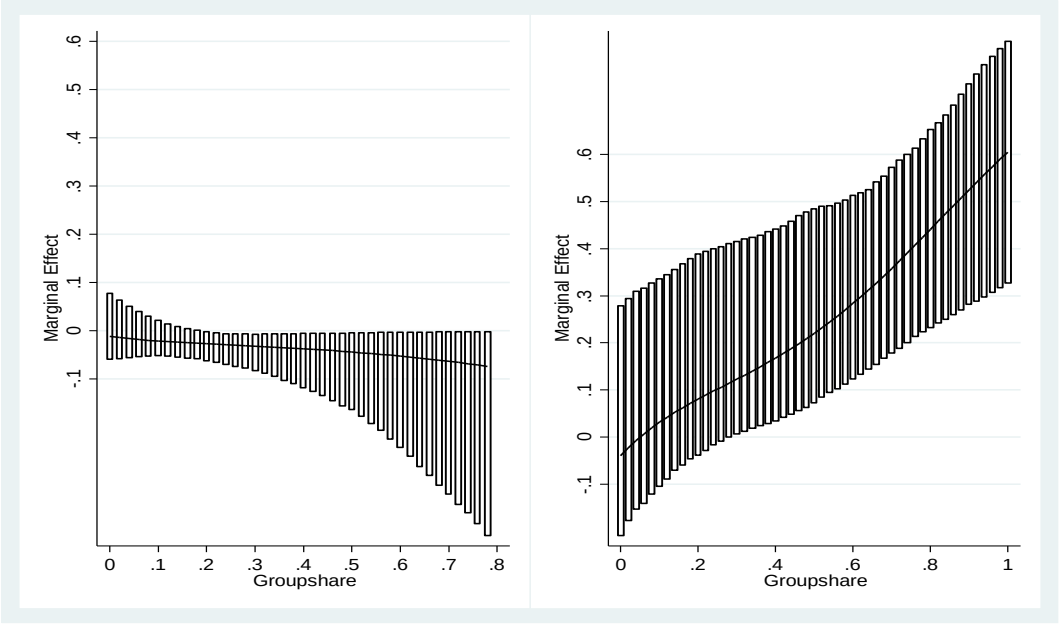


Figure C8 presents the marginal effects when we split the sample in two, one sub-sample for ethnically heterogeneous countries, the other sub-sample containing ethnically homogeneous countries, as measured by Fearon and Laitin's (2003, 2007) ethnolinguistic fractionalization (ELF) indicator. We split the samples along the median ethnic fractionalization. The graph on the left shows the marginal effects for ethnically heterogeneous countries, while the graph on the right shows the marginal effects for ethnically homogenous countries. (In each graph the x-axis covers the observed range of partitioned group share in each sub-sample.) As is clear, and in line with our expectations, all of the action is in the ethnically homogeneous countries. In ethnically heterogeneous countries, switching from a partitioned group that did not produce the first to one that did actually *reduces* the likelihood a segment will be disputed, an effect that is significant

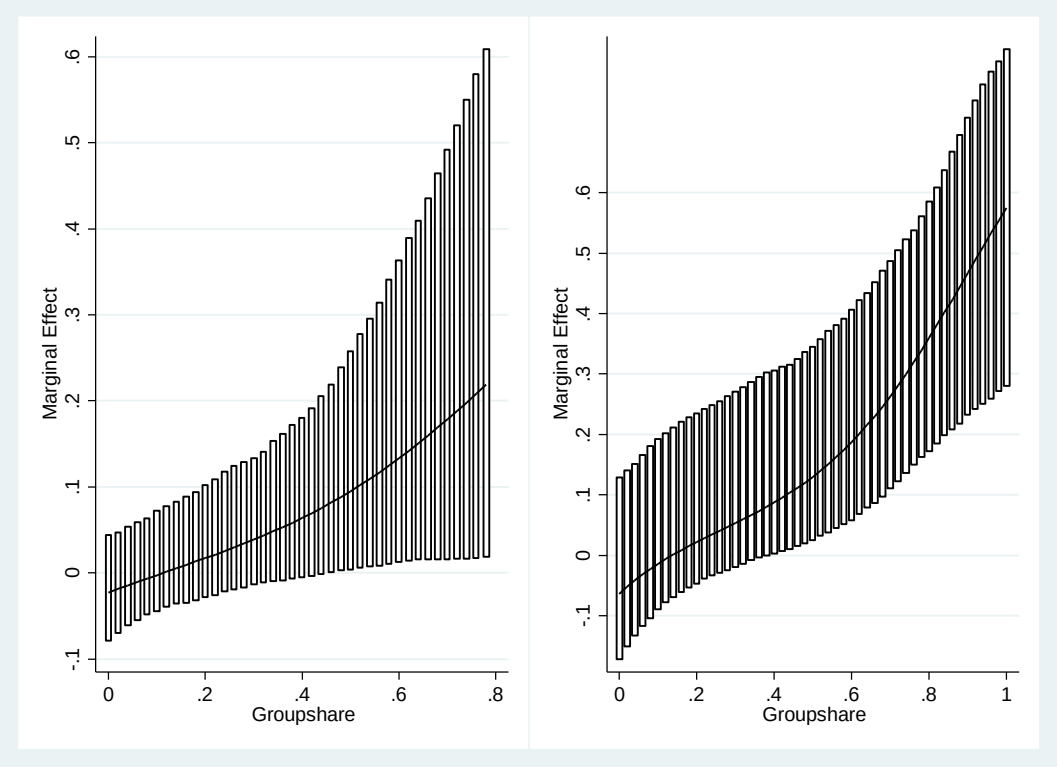
once the group is larger than about a fifth of the overall population. However, there is a strong and positive effect in ethnically homogeneous states, when the partitioned group is larger than about a third of the population. We conclude that ethnic fractionalization, a country-wide characteristic to be sure, is a major conditioning factor for the local characteristics of ethnic partitioning and ethnic political power in their effects on the likelihood a segment will be disputed.

Figure C8: Marginal effects of switching *1st leader* and the effects of group size in *ethnically heterogeneous* (left) and *ethnically homogeneous* (right) sub-samples as measured by ELF



We replicate these results with border fractionalization in **Figure C9**, again splitting the sample at the median (which is when the border partitions 9 groups). Again, the graph at the left presents the results from the heterogeneous sub-sample (high fractionalization) whereas the graph on the right represents the homogeneous sub-sample (low fractionalization).

Figure C9: Marginal effects of switching *1st leader* and the effects of group size in *heterogeneous* (left) and *homogeneous* (right) sub-samples as measured by # of groups partitioned by the border.



In both sub-samples the slope of the effect of increasing group share is positive and the effect of switching from a partitioned group that did not produce the first leader to that did is significant when the partitioned group's population share larger than about 40% in the homogenous sample and larger than about 46% in the heterogeneous sample. (Note that in the left – heterogenous sample – we scale the x-axis to .8 because the maximum groupshare in this sample = 0.784.) However, the slope of the effect of group share is much steeper in the homogeneous sub-sample.

We do not report similar tests on the Murdock map since there are very few partitioned 1st leaders using this map. Slicing the data one more time by splitting the sample makes the results highly unstable.

Alternative estimators

We now re-analyze the main models from the article with two sets of alternative estimators – spatial lag models and fixed effects models – to see how our hypotheses fare when subjected to these tests.

a) Spatial dependence

As noted in the text, the nature of territorial claims suggests that the probability of a claim on one segment is highly dependent on whether there is also a claim on adjacent segments. This spatial dependence can be captured by employing an estimator with a spatial lag. Formally, such an estimator take the form

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{e}\text{, or}$$

$$\mathbf{y} = \left(\mathbf{I} - \rho\mathbf{W}\right)^{-1}\left(\mathbf{X}\boldsymbol{\beta} + \mathbf{e}\right),$$

where **W** is an *n* x *n* spatial weight matrix that captures the dependence of each observation on every other observation and $\rho \in (-1,1)$ is the estimated spatial lag. For this test, we specify **W** such that each segment is spatially dependent on the two adjacent segments within the same dyadic border (unless the segment is at an end of the border, in which case it only has one neighbor).

Although this estimator captures the spatial dependence in our dependent variable, there are several compromises associated with implementing it. Because of limits on the maximize

matrix size, we cannot build a spatial weight matrix at the level of 1km border segments; instead, we combine these into 20km segments and aggregate the independent variables to that level by averaging them. The dependent variable remains dichotomous and indicates whether or not there was a claim anywhere on the segment. More importantly, because a probit model with a spatial lag is impractical, we use a linear regression model. Thus, while the dependent variable is bounded between zero and one, this constraint is not imposed by the estimator. We implement the spatial lag using the Stata plug-in **`spatreg`**.

Tables C8a and **C8b** re-analyze the ethnic politics models of **Table 1** in the main text, using the corresponding spatial lag models. Two broad conclusions follow. First, since ρ can range between 1 and -1, the estimated ρ of roughly 0.92 indicates a high degree of spatial dependence. Second, all our main results hold and with higher degrees of statistical significance. The results of the linear combinations are now all significant at the .005 level, with the exception of the linear combination of *Partition* and *In Coalition* in Model 6 which uses the EPR data, which also was not significant either in **Table C2b** above. The linear combination of *Partition* and *Powerless in the target* is now significant in Models 5 and 6, whereas it failed to reach significance in **Table C2b** above. In addition, *Border length* now is statistically significant in all models.

In **Table C9** we replicate the Realist and Institutional models of **Tables 3** and **4**. The results by and large mirror those in the text. In addition, compared to the Realist model (1) in **Table 3**, the results present two striking differences. First, *Distance to the capital of the claimant* – a measure of the claimant's ability to project power – is *positively* and significantly correlated with disputed status, while the presence of *Oil on the target side* is negatively and significantly associated with disputed segment status. Neither of these effects is consistent with the

hypotheses regarding those variables derived from the realist model. Second, comparing the Institutional model (2) to its counterpart in **Table 4**, *Straight line* segments now positively and significantly predict that a (20 km) segment will be disputed.

We complete our spatial lag re-analysis in **Tables C10a** and **C10b** by replicating the results run on the Full sample and the Early disputes in **Tables 5a** and **5b**. Comparing **Table C10a** with **Table 5a**, e.g., comparing results that rely on the GREG data to identify ethnic groups and their location, we see that the spatial lag model yields results very similar for the ethnic variables, with a somewhat higher degree of statistical stronger significance. It is noteworthy that the *Partitioned group's population share in the target* becomes significant in the spatial lag linear model, whereas it was not in the simple logit model. *Border fractionalization of the claimant* and *Distance to minerals on the target side* both are no longer significant for the early disputes. Switching to a comparison of **Table C10b** with **Table 5b**, e.g., focusing on EPR groups, we again find very similar results for the ethnic variables, with the same linear combinations significant at similar levels. Several other variables gain statistical significance in the spatial lag models when they failed to reach significance in the simple logit model in **Table 5b**. *Minerals in partitioned group in target*, *Partitioned group area share in target*, *Straight line segment*, and *Claimant lost territory in colonial period* all become significant; however, the first two are signed incorrectly for their respective hypotheses. Using the EPR indicators for ethnic groups surprisingly produces the corollary finding in both the Full and the Early disputes sample that border segments that follow straight lines do significantly increase the likelihood such segments will be disputed, in line with Englebert et al.'s (2002) earlier findings. In sum, the main results from the original analysis in the paper – strong support for the ethnic model and elements

of the institutional model, inconsistent results for the realist model – remain robust when we control for spatial dependence through the introduction of the spatial lag.

Table C8a. Ethnic Politics Models, Spatial Lag Regression (linear) model

Variables	(1) GREG	(2) GREG	(3) GREG	(4) EPR	(5) EPR	(6) EPR
1.Partition	0.002	-0.006***	-0.02	0.003*	-0.009***	-0.001
	(0.002)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)
Partitioned group status in claimant						
2.First leader		0.015***	-0.018***			
		(0.003)	(0.006)			
3.Population share		0.015***	-0.011***			
		(0.003)	(0.003)			
4.First leader * pop. share			0.069***			
			(0.010)			
5.In control					0.024***	0.023***
					(0.004)	(0.004)
6.In coalition					-0.001	-0.002
					(0.002)	(0.002)
Partitioned group status in target						
Powerless					0.015***	0.008***
					(0.003)	(0.003)
Population share		0.0002	-0.003			-0.014***
		(.003)	(0.003)			(0.003)
Border fractionalization			-0.001***		-0.005***	-0.005***
			(0.000)		(0.001)	(0.001)
Border length (logged)	0.005***	0.004***	0.004***	0.005***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Constant	-0.028***	-.020**	-0.012	-0.027***	-0.018*	-0.018*
	(0.0219	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)
ρ	0.921***	0.915***	0.911***	0.921***	0.916***	0.915***
	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)
Observations	7756	7756	7756	7756	7756	7756
Variance ratio	0.830	0.837	0.842	0.830	0.837	0.837

Note: ρ is the coefficient on the spatial weigh matrix. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10

Table C8b. Ethnic Politics Spatial Lag model - Linear Combinations

Linear Combinations	**(2) GREG**	**(3) GREG**	**(5) EPR**	**(6) EPR**
Linear Combination 1 & 2 |	0.008**	-0.020***		
	(0.003)	(0.006)		
Linear Combination 1 & 3 |	-0.006***	-0.010***		
Population share = .7	(0.002)	(0.002)		
Linear Combination 1, 2 & 3 |	0.019***	-0.028***		
Population share = .7	(0.003)	(0.006)		
Linear Combination 1, 2, 3 & 4 |		0.20***		
Population share = .7		(0.003)		
Linear Combination 1 & 5			0.015***	0.023***
In control			(0.003)	(0.003)
Linear Combination 1 & 6			-0.010***	-0.002
In coalition			((0.002)	(0.002)
Linear Combination 1 & 7			0.006***	0.008***
Powerless in target			(0.002)	(0.002)
Linear Combination 1 & 8				-0.015***
Pop. share in target = .7				(0.003)

Note: Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.10

Table C9: **Realist and Institutional Models, Spatial Lag Regression (linear) model**

Realist Variables	**(1) Realist**	**Institutional Variables**	**(2) Institutional (Dir. dyadic)**
Cap share of claimant	0.003	River border	-0.007***
	(0.003)		(0.002)
Area share of claimant	-0.015***	Watershed border	-0.010***
	(0.004)		(0.002)
Dist. to Cap of claimant	0.003***	Straight line segment	0.005***
	(0.001)		(0.002)
Oil on target side	-0.007***	Demarcated prior to independence	-0.002
	(0.002)		(0.002)
Dist to Oil in target	0.000	Border created after 1920	-0.005
	(0.000)		(0.003)
Minerals on target side	0.002	Border undefined or provisional	0.031***
	(0.002)		(0.006)
Dist to Min in target	-0.000***	Border change in col. period	
	(0.000)		
		Claimant lost territory	0.013***
			(0.003)
		Target lost territory	-0.001
			(0.003)
Border length (logged)	0.006***	Border length (logged)	0.004***
	(0.001)		(0.001)
Constant	-0.040***	Constant	-0.021**
	(0.012)		(0.009)
ρ	0.919***		0.917***
	(0.011)		(0.007)
Observations	7756		7756
Variance ratio	0.832		0.836

Note: Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table C10a: Combined GREG Models, Spatial Lag Regression (linear) model

Variables	(1) GREG-Full	(2) GREG-Early
1.Partition	0.001	0.002
	(0.002)	(0.002)
2.Partition and 1st leader	-0.020***	-0.018***
	(0.005)	(0.006)
3.Partitioned group population share in claimant	-0.010***	-0.009***
	(0.003)	(0.003)
4.Partition and 1st leader * population share	0.074**	0.074***
	(0.010)	(0.011)
Linear Combination 1-4 | population share =.70	0.025***	0.030***
	(0.004)	(0.004)
Partitioned group population share in target	-0.014***	-0.008**
	(0.003)	(0.003)
Border fractionalization of claimant	-0.000***	-0.000
	(0.000)	(0.000)
Area share of claimant	-0.003	0.000
	(0.004)	(0.003)
Oil on target side	-0.021***	-0.018***
	(0.003)	(0.002)
Distance to oil on target side	0.000***	0.000***
	(0.000)	(0.000)
Minerals on target side	-0.001	0.008***
	(0.003)	(0.003)
Distance to minerals on target side	-0.001	-0.000
	(0.003)	(0.000)
Minerals in partitioned group in target	0.001	-0.001
	(0.002)	(0.002)
Partitioned group area share in target	0.007*	0.003
	(0.004)	(0.004)
River border	-0.006***	-0.003
	(0.002)	(0.002)
Watershed border	-0.009***	-0.005***
	(0.002)	(0.002)
Straight line segment	0.002	0.001
	(0.002)	(0.001)
Border at independence undefined or provisional	0.039***	0.039***
	(0.006)	(0.006)
Claimant lost territory in colonial period	0.017***	0.016***
	(0.003)	(0.003)
ρ	0.903***	0.902
	(0.008)	(0.009)
Observations	7756	7756
Variance ratio	0.850	0.855

Note: Constant and variable for border length included but not reported. Robust standard errors are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table C10b: Combined EPR models, Spatial Lag Regression (linear) model

Variables	(1) EPR-Full	(2) EPR-Early
1.Partition	0.004	0.003
	(0.003)	(0.002)
2.Partition and in control in claimant	0.22***	0.023***
	(0.004)	(0.003)
Linear Combination 1 & 2	0.026***	0.025***
	(0.004)	(0.003)
3.Partition and in coalition in claimant	-0.000	0.000
	(0.002)	(0.002)
Linear Combination 1 & 3	0.004	0.003
	(0.003)	(0.002)
4.Partition and powerless in target	0.014***	0.009***
	(0.003)	(0.002)
Linear Combination 1 & 4	0.018***	0.012***
	(0.003)	(0.002)
Border fractionalization of claimant	-0.003***	-0.004**
	(0.001)	(0.001)
Area share of claimant	-0.009**	-0.002
	(0.004)	(0.004)
Oil on target side	-0.017***	-0.015***
	(0.002)	(0.002)
Distance to oil on target side	0.000***	0.000***
	(0.000)	(0.000)
Minerals on target side	0.005**	0.010***
	(0.002)	(0.002)
Distance to minerals on target side	-0.000**	-0.000**
	(0.000)	(0.000)
Minerals in partitioned group in target	-0.015***	-0.006**
	(0.003)	(0.002)
Partitioned group area share in target	-0.001	-0.012**
	(0.004)	(0.002)
River border	-0.006**	-0.003
	(0.002)	(0.002)
Watershed border	-0.011***	-0.007***
	(0.002)	(0.002)
Straight line segment	0.006***	0.005***
	(0.002)	(0.002)
Border at independence undefined or provisional	0.034***	0.034***
	(0.006)	(0.006)
Claimant lost territory in colonial period	0.011	0.009***
	(0.003)	(0.003)
ρ	0.909***	0.909***
	0.008	0.009

Observations	7756	7756
Variance ratio	0.845	0.848

Note: Constant and variable for border length included but not reported. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

b) Fixed Effects Models

A final potential concern is that the models omit features of the claimant state that are correlated with its propensity to challenge its borders. The ethnic model includes one such factor—ethnic homogeneity—but it is possible that there are other features of a state history, geography, or politics that make it more or less likely to issue a challenge anywhere. To assess this possibility, we re-estimate our results using claimant state fixed effects. The preferred estimator for fixed effects with a dichotomous dependent variable is the conditional logit. However, because of the very large number of segments per claimant state, the conditional logit model does not converge. Fortunately, the large number of observations per unit means that dummy variable indicators for each state can be included in a logit model, and the usual concern about bias in this context does not apply. Note that the fixed effects logit model drops all countries that did not dispute any segments. Therefore, all estimates are identified by within-country variation among the set of countries that disputed at least one segment. Given that only 35 states in our data set made a claim, this test reduces our effective sample quite dramatically, making it a very stringent test of our hypotheses. **Tables C11a** and **C11b** report estimates of the combined models from **Tables 5a** and **5b** in the text, using both the Full sample and the sample of Early disputes. (The fixed effect model of the late disputes on the EPR data failed to converge.)

When using the GREG data (**Table C11a**) the basic pattern of coefficients on the ethnic variables remain unchanged. We report the linear combinations of *Partition*, *Partitioned group*

and 1st leader, *Partitioned group's population share in the claimant* and the interaction *Partitioned group and 1st leader * Partitioned group's population share in the claimant* in the table, setting the partitioned group at 70% of the population (significance levels increase as the population share increases).[2] The fixed effect Full model also shows a positive effect of the presence of minerals in the partitioned group on the target side, a variable which was not significant in any other tests. Results based on the EPR groups (**Table C11b**) continue to show a positive and significant effect for the linear combination of *Partition* and *Partitioned group is powerless in the target*. *Partitioned group's area share in the target* is negatively correlated with dispute risk, further emphasizing the role of the partitioned group's status in the target. Segments that partition a partitioned group that is in control in the claimant, and *not* powerless in the target are not significantly different from segments that do not partition. However, in line with our hypotheses, a segment that partitions an ethnic group that is in control in the claimant and is powerless in the target is indeed significantly more likely to be disputed.

In conclusion, the ethnic model is robust across a large number of specification and estimators and passes the hardest tests we could devise. The only time the political status of the partitioned group in the claimant state fails to predict whether a segment is disputed is when we run the fixed effect model on the EPR operationalization of the ethnic variables of interest. Likewise, certain elements of the institutional model remain robust across specifications. By contrast, most of the variables associated with the realist hypotheses either have no effect, have an effect in the opposite direction from that hypothesize, or have the hypothesized effect only

[2] As we noted in the main text, in half the observed cases where a first leader's group is partitioned the group constitutes about 80% of the population.

inconsistently across models.

Table C11a: Combined GREG models, Fixed Effects Logit

Variables	(1) Full Sample Fixed Effects Logit	(2) Early disputes Fixed Effects Logit
1.Partition	0.11	0.70
	(0.61)	(0.82)
2.Partition and 1st leader	-1.19	-4.23***
	(0.98)	(1.06)
3.Part. group pop. share in claimant	-0.93	-16.22***
	(4.87)	(4.60)
4.Partition and 1st leader * pop share	4.04	24.50***
	(5.04)	(6.48)
Linear Combination 1-4 | Pop. Share=.7	1.09**	2.28*
	(0.53)	(1.30)
Part. group pop. share in target	-0.67	1.49
	(1.72)	(1.77)
Area share of claimant	8.46***	13.73***
	(3.25)	(4.68)
Oil on target side	-5.40***	-7.42***
	(1.61)	(2.61)
Distance to oil on target side	0.01***	0.01***
	(0.00)	(0.00)
Minerals on target side	-2.45**	1.50
	(1.21)	(3.16)
Distance to minerals on target side	-0.00	0.00
	(0.00)	(0.00)
Minerals in part. group in target	1.57***	1.51*
	(0.56)	(0.91)
Part. group area share in target	-1.77	-4.74***
	(1.57)	(1.19)
River border	-0.94*	-1.02
	(0.50)	(0.81)
Watershed border	-18.48***	-18.72***
	(1.08)	(2.28)
Straight line segment	0.48	-0.29
	(0.57)	(0.78)
Border undefined or provisional	4.64***	6.50***
	(1.55)	(1.73)
Claimant lost territory in col. period	3.72***	5.92***
	(1.07)	(1.93)
Observations	95,251	70,142
Pseudo -R-squared	0.61	0.71

Note: Constant and variable for border length included but not reported. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table C11b: Combined EPR models, Fixed Effects Logit

Variables	(3) Full Sample Fixed Effects Logit	(4) Early disputes Fixed Effects Logit
1.Partition	1.56	3.12
	(1.02)	(2.04)
2.Partition and in control in claimant	-0.82	-1.89*
	(0.77)	(0.98)
3.Part. and in coalition in claimant	-3.67***	-5.62***
	(1.36)	(1.76)
4.Partition and powerless in target	0.93	-0.15
	(0.94)	(1.74)
Linear combination 1 & 2	0.74	1.22
	(0.70)	(1.23)
Linear combination 1 & 3	-2.11	-2.50
	(1.52)	(2.83)
Linear combination 1 & 4	2.49***	2.96***
	(0.62)	(0.74)
Linear combination 1, 2 & 4	1.66*	1.07
	(0.96)	(0.82)
Linear combination 1, 3 & 4	-1.18	-2.66
	(1.35)	(1.71)
Areas share of claimant	7.71**	12.72**
	(3.34)	(5.77)
Oil on target side	-4.68***	-5.61***
	(1.19)	(1.86)
Distance to oil on target side	0.01***	0.01**
	(0.00)	(0.00)
Minerals on target side	-1.57	2.63
	(1.30)	(3.11)
Distance to minerals on target side	-0.00	0.00
	(0.00)	(0.00)
Minerals in part. group in target	1.07	2.22**
	(0.81)	(1.03)
Part. group area share in target	-5.80***	-7.95***
	(1.49)	(2.54)
River border	-0.94*	-0.78
	(0.52)	(0.86)
Watershed border	-19.00***	-18.25***
	(1.41)	(1.73)
Straight line segment	0.56	-0.01
	(0.58)	(0.70)

Border undefined or provisional	5.60***	6.17***
	(1.97)	(1.71)
Claimant lost territory in col. period	3.75***	5.31***
	(1.05)	(1.75)
Observations	95,251	70,142
Pseudo R-squared	0.61	0.70

Note: Constant and variable for border length included but not reported. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1