Complements or Substitutes? How Institutional Arrangements Bind Traditional Authorities and the State in Africa

Soeren J. Henn

Supplementary Materials

A Data Appendix

A.I Survey Questions

The **Perceptions of Traditional Leader Z-score** takes the standardized version of the following variables in the Afrobarometer survey and combines them in a z-score with mean 0 and standard deviation of 1:

- *Influence Traditional Leader*: "How much influence do traditional leaders currently have in governing your local community?" (Question 65 in Round 4)
- *Trust Traditional Leader*: "How much do you trust each of the following, or haven't you heard enough about them to say: Traditional leaders?" (Question 49I in Round 4, Q52K in Round 6)
- *Corruption Traditional Leader*: "How many of the following people do you think are involved in corruption, or haven't you heard enough about them to say: Traditional leaders?" (Question 50H in Round 4, Q53H in Round 6) (Inversed for the index)
- *Contact Traditional Leader*: "During the past year, how often have you contacted any of the following persons about some important problem or to give them your views: A traditional ruler?" (Question 23F in Round 3, Q27B in Round 4, Q24E in Round 6)

Note that each question offers the option of "Don't Know" or "Refuse to Answer". I code both cases as missing. There is no significantly different occurrence of these cases in the four variables across institutional settings.

The Afrobarometer surveys contain two additional questions about traditional authorities that are used in Table A5 to determine what drives the effect.

- *Performance of Traditional Leader*: "Do you approve or disapprove of the way the following people have performed their jobs over the past twelve months, or haven't you heard enough about them to say: Your Traditional Leader?" (Question Q68D in Round 6)
- *Traditional Leader Listens*: "How much of the time do you think the following try their best to listen to what people like you have to say: Traditional leaders?" (Question Q54C in Round 4)

Further the Afrobarometer survey contains several questions about local development, public goods provision, and tax payment from which I create a **Afrobarometer State Presence Index**. First, the sub-indexes are created by combining their standardized variables into a z-score. Second, the three sub-indexes (Development, Public Goods, and Taxation) are combined into the state capacity index with mean zero and standard deviation of 1.

- *Development*: Enumerators are asked whether the enumeration area contains public services.

  - "Are the following services present in the primary sampling unit/enumeration area: Electricity grid that most houses could access?" (Question EA-SVC-A in Rounds 3–6)
  - "Are the following services present in the primary sampling unit/enumeration area: Piped water system that most houses could access?" (Question EA-SVC-B in Rounds 3–6)
  - "Are the following services present in the primary sampling unit/enumeration area: Sewage system that most houses could access?" (Question EA-SVC-C in Rounds 3–6)

- *Public Goods*: Enumerators are asked whether the enumeration area contains state provided public goods:
  - "Are the following facilities present in the primary sampling unit/enumeration area, or within easy walking distance: Post-office?" (Question EA-FAC-A in Rounds 3–6)
  - "Are the following facilities present in the primary sampling unit/enumeration area, or within easy walking distance: School?" (Question EA-FAC-B in Rounds 3–6)
  - "Are the following facilities present in the primary sampling unit/enumeration area, or within easy walking distance: Police station?" (Question EA-FAC-C in Rounds 3–6)
  - "Are the following facilities present in the primary sampling unit/enumeration area, or within easy walking distance: Health clinic?" (Question EA-FAC-D in Rounds 3–6)
  - "Are the following facilities present in the primary sampling unit/enumeration area, or within easy walking distance: Market stalls (selling groceries and/or clothing)?" (Question EA-FAC-E in Rounds 3–6)
  - "In the PSU/EA, did you (or any of your colleagues) see: Any policemen or police vehicles?" Question EA-SEC-A in Rounds 3–6)

- *Taxation*: In round 4 respondents are asked whether they pay different types of taxes:
  - "Have you had to make any of the following payments during the past year: Fees for a government service such as education or health care?" (Question Q64A in Round 4)
  - "Have you had to make any of the following payments during the past year: Licence fees to local government e.g., for a bicycle, cart, or market stall?" (Question Q64B in Round 4)
  - "Have you had to make any of the following payments during the past year: Property rates or taxes?" (Question Q64C in Round 4)
  - "Have you had to make any of the following payments during the past year: Public utility fees, e.g., for water, electricity or telephone?" (Question Q64D in Round 4)

– "Have you had to make any of the following payments during the past year: Income taxes?" (Question Q64E in Round 4)

Similarly, the DHS survey allows us to create a **DHS State Presence Index** by combining the following standardized variables into a z-score with mean 0 and standard deviation of 1:

– *Electricity*: Whether the household has access to electricity. (HV206)

– *Piped Water*: Whether the household has access to piped water. (HV201)

– *Registered*: The percentage of children in each household that are registered with the state or have a birth certificate. (HV140)

– *Vaccination Card*: The percentage of children with a vaccination card in each household. (H1)

Additionally, we can create a **DHS Development Index** by combining the following standardized variables into a z-score with mean 0 and standard deviation of 1:

– *Literacy*: Whether the respondent can read a card shown by the enumerator. (H108)

– *Wealth*: Household wealth on a 1–5 scale. (HV270)

– *Piped Water*: Whether the household has access to piped water. (HV201)[16]

A.II Control Variables

– **Distance to the Capital**: The distance of a village from the capital city, measured in kilometers. *Source: OpenStreetMap*

– **Distance to the National Border**: The distance of a village from the national border, measured in kilometers. *Source: Digital Chart of the World*

– **Distance to the Coast**: The distance of a village from the nearest coastline, measured in kilometers. *Source: Digital Chart of the World*

– **Elevation**: Average value of elevation for grid cells of 30 Arc-Seconds (equivalent to 250 meters), measured in meters above sea level. *Source: SRTM version 4.1 (NASA)*

– **Ruggedness**: Averaging the Terrain Ruggedness Index of 30 by 30 arc-second cell. It is measured by dividing the millimeters of elevation difference by the area of the 30 by 30 arc-second cell. *Source: Nunn and Puga (2012)*

– **Land Suitability for Agriculture**: The fraction of each grid cell that is suitable to be used for agriculture. It is based on the temperature and soil conditions of each grid cell. *Source: Atlas of the Biosphere*

[16]Note that this variable is also included in DHS State Presence Index. This is to mirror the Afrobarometer Index which also includes piped water as a development outcome. Results remain when excluding "piped water" from either the state capacity or development index.

– **Distance to Historical Cities**: The distance of a village from the nearest historical city, measured in kilometers. *Source: Chandler (1987)*

– **Malaria Ecology Index:**: The index takes into account the prevalence and type of mosquitoes indigenous to a region, their human biting rate, their daily survival rate, and their incubation period. The index has been constructed for 0.5 degree by 0.5 degree grid-cells. *Source: Kiszewski et al. (2004)*

– **Distance to Catholic and Protestant mission stations**: The distance of a village from the nearest Catholic or Protestant mission station, measured in kilometers *Source: Nunn (2010)*

– **Distance to Railroad**: The distance of a village from the nearest railroad built before 1960, measured in kilometers. *Source: Jedwab and Moradi (2015)*

– **Administrative Unit Size**: Surface area of each administrative unit measured in square kilometers. *Source: GIS calculations by the author.*

B Description of Robustness Checks

The following section shows robustness to a range of different specifications and measurements; most notably, the validity of the assumptions underpinning the regression discontinuity design, different choices for the main specification, and the possible endogeneity of administrative borders and headquarters.

Throughout the robustness checks, the results remain qualitatively the same: distance to the state leads to an increased role of traditional leaders when the state and chiefs are institutionally separated. When both are linked, chiefs act as complements and their role decreases when the state is weak. I also rerun all robustness checks for the DHS data, the results of which can be seen in Tables B3–B6.

Testing the RDD assumption

Two underlying assumptions are crucial for the causal validity of any regression discontinuity specification: smooth variation of covariates and no sorting around the cutoff.

If treatment is indeed random around the border and not the result of confounding factors, treatment should not have an effect on pretreatment covariates. In the case of changes in state capacity, few potential variables are pretreatment. Therefore, to test the balance of my sample, I run the main specification on a set of geographical and historical variables. The results are reported in Table A6. Two out of ten are significantly different on the side of the border farther away from the state—distance to the national border and distance to colonial railways. A look at the observations on the map and sensitivity analysis finds that this is driven by observations from one country (Cameroon).[17] Still, all variables in the table

[17]Panel A in Figure A3 shows that dropping each country individually from the analysis does not affect the results.

and their interaction with institutionalization of chiefs are included as controls in the main analysis.[18]

For observations on both sides of the border to be comparable, there must be little or no sorting—i.e. chiefs and citizens should not move across internal borders to be closer or farther away from the state. One indication for sorting would be different densities on both sides of the border. To test for this, I perform McCrary tests on the Afrobarometer sample for the different bandwidth specification, the results of which can be seen in Figure B3. Unfortunately, neither the Afrobarometer nor the DHS data contains information on the population of the settlement. Consequently, these graphs only show the distribution of settlements around the bandwidth used in the specification to see whether settlements cluster close to administrative boundaries on the side closer or farther from the headquarters. Figure B3 shows no indication for significant variation in density on around the cutoff. Second, I use the DHS data to test whether the low state capacity treatment induces migration on either side of the border. Table A8 shows that neither migration by children, men, women, nor an indicator combining the three, is significantly different on one side of the border.

Different Specifications

The choice of optimal bandwidth is a crucial step in any regression discontinuity design. Various strategies exist to select an optimal bandwidth (Imbens and Kalyanaraman, 2012; Calonico, Cattaneo and Titiunik, 2014). The matched regression discontinuity design in this paper, however, creates inconsistent estimators for the optimal bandwidth.[19] In order to check the robustness of these results, I vary the bandwidth between 3 and 20 kilometers. Sample size restricts the possibility to use bandwidths smaller than 3 km, and larger bandwidths than 20 km become less meaningful from an identification standpoint, as villages can be up to 40 km away from each other and are thus less comparable. The results can be seen in Figure 4 in the paper. The results follow general regression discontinuity specifications, larger but less precise coefficients when using smaller bandwidths. No matter the bandwidth choice, chiefs remain substitutes from the state when not institutionalized by the constitution and they show the opposite relationship when being institutionalized. Still, the associated confidence intervals may not have correct coverage even if the estimator is unbiased, suggesting that it might be appropriate to use a higher critical value (Armstrong and Kolesar, 2017). Both the difference between treatment coefficients of the institutionalized and not institutionalized samples and the coefficient in the interaction specification surpass the most conservative critical value of 2.8.

The main specification uses an intensive treatment measure that indicates how much larger the distance to the administrative headquarters is on one side than on the other side of the internal administrative border. This intensive treatment measure is then scaled by the

[18] I also run the analysis without using controls in Column 2 in Table A9, and the results remain consistent.

[19] This is due to the matching aspect of the specification. In a normal RD setting, extending the bandwidth from X to X+1 only adds observations that are between X and X+1 from the cutoff. In this case, however, increasing the bandwidth from X to X+1 will not only add observations between X and X+1 from the cutoff but also their matched observations on the other side of the border, which could be anywhere from 0 to X+1 from the cutoff. Thus, the variance bias trade-off calculated by the standard optimal bandwidth algorithms is not consistent.

country and administrative division specific effect of distance on state capacity outcomes. The results hold when using the more rudimentary specification with a binary treatment indicator (Column (3) in Table A9). Using absolute log-distance to administrative headquarters instead of the treatment indicator returns similar results (Column (4)). Removing the scaling of treatment by the country and administrative division specific coefficient of distance on state capacity also does not change the findings (Column (5) in Table A9).

The main specification differs from some geographical regression discontinuity designs that use polynomial longitude-latitude specification (e.g., Dell. 2010). These studies estimate differences across a single geographical boundary. In that case using longitude and latitude offers a precise way of controlling for an observation's location vis-à-vis the boundary. However, when analyzing the differences across multiple boundaries, and in different countries as is the case here, using longitude and latitude becomes problematic. Since boundaries are in many different locations, longitude and latitude controls do not adequately capture an observation's location in relation to its boundary in this setting. Distance to the border, as used in this paper, represents a clean measure. It has the added benefit of closely mirroring the standard regression discontinuity specification that incorporates a control for the distance to the cutoff. Nevertheless, I show that using this specification results in the same heterogeneous pattern (Column (6) in Table A9). Furthermore, I also conservatively cluster the standard errors at the highest administrative division instead of the lowest (Column (7) in Table A9).

Lastly, Column (8) in Table A9 removes all observations who have a different treatment assignment when using their own distance to their administrative headquarters instead of the average distance on their side of the border region (56 out of 801 observations).

The specification could also be sensitive to the inclusion or exclusion of outliers, both in terms of extreme values of the explanatory variable as well as specific countries. To make sure the results are not driven by such outliers, I drop extreme outliers that are more than 100 km and 50 km away from the administrative headquarters in Columns (2) and (3) in Table A10, respectively. In Panel A in Figure A3, I show the results dropping one country at a time. Column (4) in Table A10 does not restrict to border segments by also including villagers whose nearest village on the other side of the border is farther than 30 km.

More generally, the results are also robust to different typical geographic regression discontinuity specification. While the logged distance is used in the main specification, the non-logged distance is used in Column (5) in Table A10. A more realistic measure of state capacity could be obtained by using travel time between villages and administrative headquarters. Travel time is linked to infrastructure investments that could be affected by state capacity or the state-chief interaction. Nevertheless, the results remain consistent when using logged travel time (Column (6) in Table A10).[20] I also restrict the analysis to rural observations since the dynamics between the state and chiefs might be different in an urban setting. Column (7) shows that the results hold when focusing on cases where observations on both sides of the border are classified as rural. Due to data availability, the samples for the Afrobarometer and DHS analysis are not identical. Column (3) in Table A12 shows that the results remain unchanged when limiting the sample to countries for which I have both

[20]Following methodology by Alegana et al. (2012) I use, altitude, land cover, rivers, and road network to calculate the travel time between a village and its administrative headquarters.

Afrobarometer and DHS geocoded data.

Endogenous Borders and Headquarters

Previous studies have found spillovers in state capacity (Acemoglu, Camilo and Robinson, 2015). If local state capacity spillovers were sizable in the African context, it would downward bias my results and reduce the potency of the regression discontinuity design. To test whether such spillovers influence the results, I control for a village's distance to administrative headquarters in the neighboring administrative unit (Column (2) in Table A11).

A concern in this particular regression discontinuity design might be that the locations of the administrative borders and headquarters are not random. Indeed, both the boundaries and the district capitals are likely to be the result of economic and political processes. Scholars have demonstrated, for example, that African governments routinely create more lower-level administrative units as part of political bargaining processes (Grossman and Lewis, 2014; Gottlieb et al., 2019). However, the endogeneity of borders and headquarters is unlikely to impact the results of this study, since both decisions are unlikely to be based on the particular villages and chiefs surveyed. Borders follow natural boundaries such as rivers or are straight lines and rarely altered for individual villages or chiefs. In other words, a strong local chief is unlikely to have the ability to influence the drawing of borders to put her village in a district with high or low state capacity.

Since the splitting of districts and the redrawing of boundaries is more prevalent in lower administrative divisions, I run the results separately for the first and second administrative divisions of the countries in my sample (Columns (3) and (4) in Table A11). Additionally, if borders were drawn to explicitly include or exclude a particular village, the boundary should be right next to the village. To exclude such potential cases I run a "Donut" RDD, where I exclude all villages within 1 km of the border (Column (5) in Table A11).

Another omitted factor in the analysis that could create discontinuity at the border is ethnicity. If administrative borders consistently coincide with ethnic demographics, the results and their interpretations could be affected. Column (6) in Table A11 indicates that this is not a concern. When controlling for ethnicity fixed effects based on the pre-colonial locations of ethnic groups, the results remain virtually unchanged.

Similarly to administrative boundaries, the location of headquarters is not based on the power of local chiefs but typically follows population density or economic activity: the biggest or economically most important village or town becomes the administrative capital. While these factors determine the location of the capital, they don't change discontinuously at the border. Controlling for the distance to the neighboring headquarters does not affect the results (Column (2) in Table A11) and there is no evidence of high levels of migration (Table A2 and A8).

Still, in some cases, the location of the capital might be influenced by a particular influential chief. To make sure the results are not driven by this phenomenon I use the most populated place in each district in 1960[21] to instrument for the location of the district capitals. Putting the distance to the instrumented capitals in the specification returns similar results (Column (7) in Table A11). Lastly, I also run a placebo test where I chose ran-

[21]Earlier data on population density is not disaggregated enough.

dom locations within administrative divisions as headquarters and estimate the effect of its distance on local chief power. The result can be seen in Column (8) in Table A11. Reassuringly, distance to these placebo headquarters does not result in sizable or significant effects, whether chiefs are institutionalized or not.

C Additional Figures

Figure A1: Bin-scatter between state capacity and distance

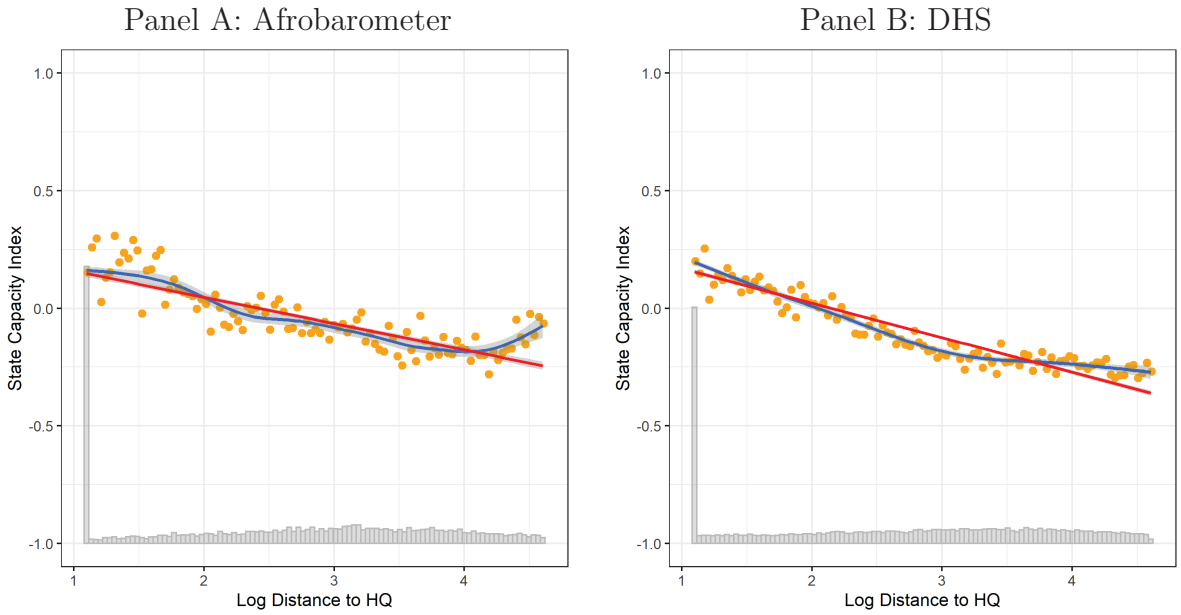


Figure A2: Illustration of Identification

Notes: This figure shows the boundaries of two states (Yobe in the West and Borno in the East) in Nigeria.

Figure A3: Results of Leaving out Countries

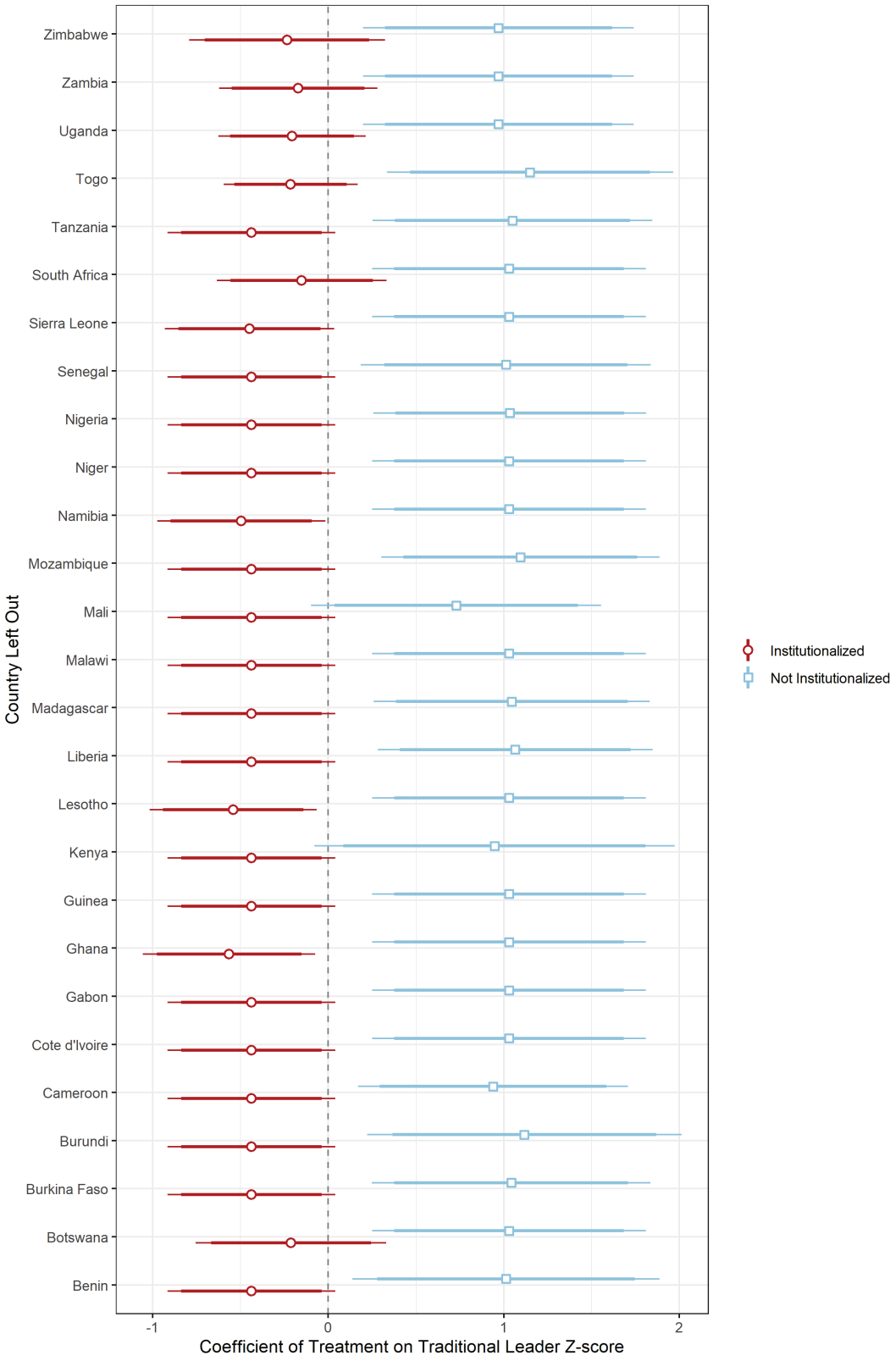
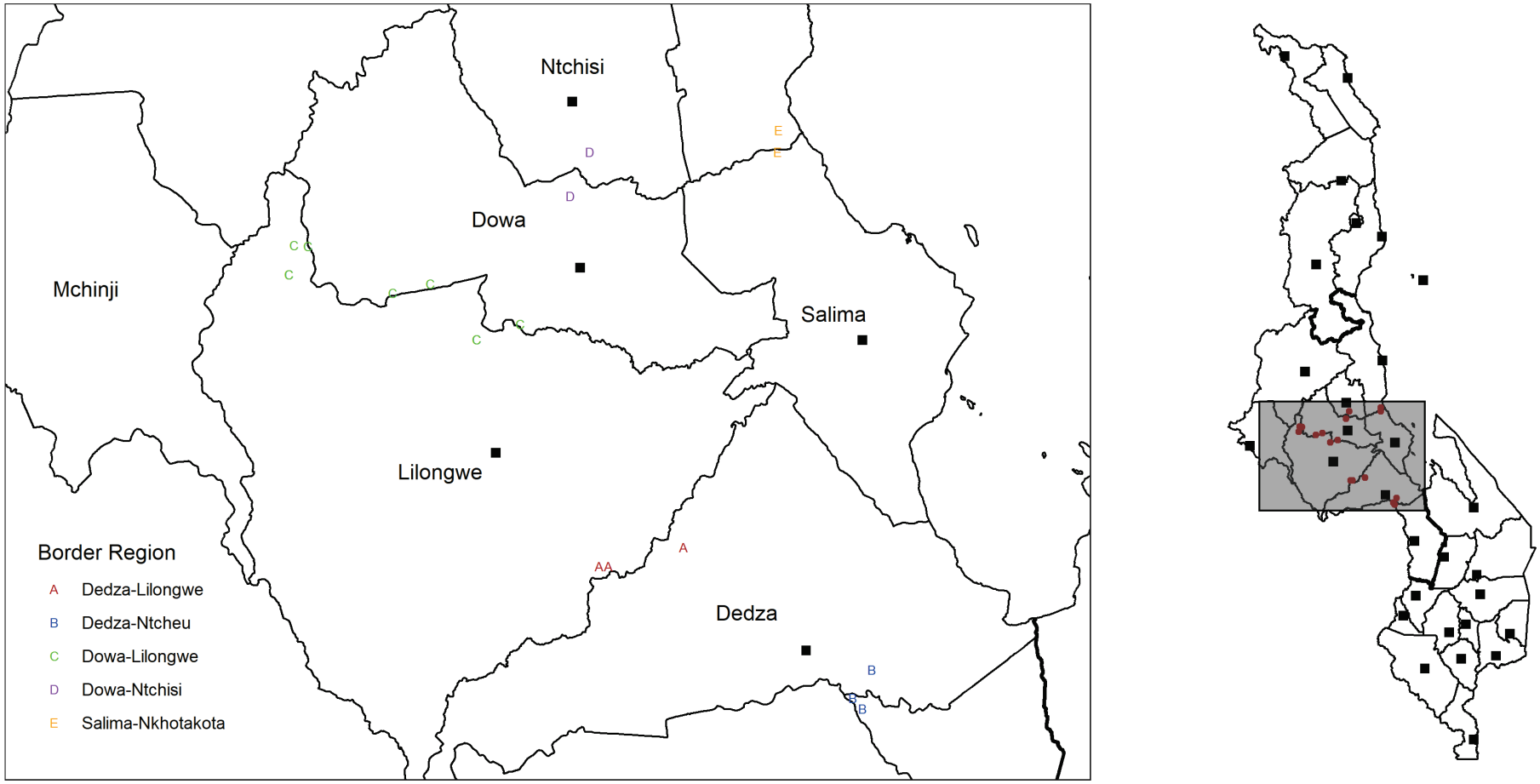


Figure A4: Example of Border Region Assignment



Notes: This figure shows the observations within 5 km of district boundaries in Malawi and their respective border region assignments marked by letters and colors. Headquarters are marked by squares.

D Additional Tables

Table A1: Administrative Divisions in Sample

Country	Admin Unit	# in 2002	# in 2005	# in 2008	# in 2012	# in 2015
Benin	department	12	12	12	12	12
Benin	commune	77	77	77	77	77
Botswana	district	15	15	16	16	16
Burkina Faso	province	45	45	45	45	45
Burkina Faso	department	351	351	351	351	351
Burundi	province	17	17	17	17	18
Burundi	commune	115	129	129	129	129
Cameroon	department	58	58	58	58	58
Cameroon	arrondissement	360	360	360	360	360
Cote d'Ivoire	department	58	70	81	107	108
Cote d'Ivoire	sub-prefectures				510	510
D.R.C	province	11	11	11	11	26
D.R.C	territory	166	166	166	166	166
Gabon	region	10	10	10	10	10
Gabon	department	48	48	50	50	49
Ghana	region	10	10	10	10	10
Ghana	district	110	110	170	216	216
Guinea	region	8	8	8	8	8
Guinea	prefecture	34	34	34	34	34
Kenya	province	8	8	8		
Kenya	county				46	46
Lesotho	district	10	10	10	10	10
Liberia	county	15	15	15	15	15
Madagascar	region		22	22	22	22
Madagascar	district	110	110	114	114	114
Malawi	region	3	3	3	3	3
Malawi	district	27	28	28	28	28
Mali	cercle	49	49	49	49	49
Mali	commune	701	701	701	701	701
Mozambique	province	10	10	10	10	10
Mozambique	district	128	128	128	128	151
Namibia	region	13	13	13	13	14
Namibia	constituency	102	107	107	107	121
Niger	region	7	7	7	7	7
Niger	department	36	36	36	63	63
Nigeria	state	36	36	36	36	36
Nigeria	lga	774	774	774	774	774
Senegal	region	11	11	14	14	14
Senegal	cr	364	364	364	431	431
Sierra Leone	district	14	14	14	14	14
Sierra Leone	chiefdom	149	149	149	149	149
South Africa	district	53	53	52	52	52
South Africa	Municipality	226	226	226	226	226
Tanzania	region	25	26	26	30	30
Tanzania	district	129	129	130	149	149
Togo	region	5	5	5	5	5
Togo	prefecture	31	31	31	36	36
Uganda	district	56	70	80	112	112
Zambia	province	9	9	9	10	10
Zambia	district	72	72	72	72	110
Zimbabwe	province	10	10	10	10	10
Zimbabwe	district	59	59	59	59	59

D.I Summary Statistics

Table A2: Summary Statistics for Full Regression Sample

Statistic	N	Mean	St. Dev.	Min	Max
Distance to Headquarter (km)	5,882	15.20	16.12	3.00	145.11
Distance to Admin. Border (km)	5,882	−0.49	2.56	−5.00	5.00
Distance to Village on Other Side (km)	5,882	8.16	6.16	0.10	29.99
Distance to Neighbouring HQ (km)	915	84.31	154.38	0.47	1,081.75
Traveltime to HQ (in min)	1,174	689.99	971.32	0.00	10,036.79
Treatment Intensity	5,611	0.48	1.00	0.00	8.14
Urban	5,882	0.50	0.50	0	1
Distance to National Capital (km)	5,787	170.13	221.02	0.15	1,789.27
Distance to National Border	5,787	74.75	73.34	0.02	378.52
Distance to Coast (km)	5,882	361.40	365.38	0.05	1,268.65
Elevation	5,882	625.50	620.92	−2	2,766
Ruggedness	5,882	0.08	0.12	0.00	1.30
Malaria Suitability	5,882	11.14	11.32	0	36
Agricultural Suitability	4,936	0.37	0.20	0.00	0.99
Distance to Christian Missions (km)	5,882	52.03	105.36	0.16	742.50
Distance to Histroical Cities (km)	5,882	450.13	378.87	0.00	1,940.92
Distance to Colonial Railroad (km)	5,882	73.40	109.15	0.00	968.55
Admin. Unit Size (sqkm)	5,787	2,657.73	7,771.85	2.22	175,770.30
Traditional Leader Z-score	810	−0.24	0.79	−2.62	2.84
Traditional Leader Influence	185	−0.10	0.97	−2.10	2.11
Trust in Traditional Leader	627	−0.31	1.06	−2.84	1.70
Corrupt Traditional Leader (Inverse)	627	−0.24	1.03	−3.96	1.94
Contact with Traditional Leader	810	−0.21	0.96	−1.04	4.05
State Presence Index	5,882	0.00	1.00	−2.96	3.00
Percentage of HH with Electricity	4,673	0.44	0.40	0.00	1.00
Percentage of Children Registered	3,551	0.52	0.32	0.00	1.00
Average Time to Water (min)	4,587	16.83	17.42	0.00	255.62
Literacy	3,655	0.53	0.31	0.00	1.00
Wealth Index	4,517	3.48	1.12	1.00	5.00
Infant Mortality	3,715	0.13	0.08	0.00	0.52
Traditional Medicine	4,006	−0.03	0.92	−0.28	9.74
Percentage of Kids Gone	3,715	0.24	0.11	0.00	0.75
Percentage of Men Born in Location	1,935	0.99	0.04	0.62	1.00
Percentage of Women Born in Location	1,929	0.98	0.04	0.55	1.00

Notes: This table shows the summary statistic of the regression sample. Only villages within 5 km of an administrative border, and which have a village on the other side of the border, are included. Villages farther than 150 km from their headquarters are dropped as are those where the neighboring village is more than 30 kilometers away. The sample for the DHS and Afrobarometer are pooled. Separate summary statistics can be found in Tables B1-B2 in the Online Appendix.

D.II Measurement

Table A3: Effect of Log Distance to HQ on Outcomes Related to State Presence

Panel A: Afrobarometer Data	*Dependent variable:*			
	Taxes paid	Local Dev	Public Goods	State Presence Index
	(1)	(2)	(3)	(4)
Log Distance to HQ	-0.155^{***}	-0.209^{***}	-0.094^{***}	-0.151^{***}
	(0.020)	(0.010)	(0.010)	(0.008)
Fixed effects	Border Region	Border Region	Border Region	Border Region
Controls	✓	✓	✓	✓
Observations	3,392	15,777	15,797	15,797
Adjusted R^2	0.240	0.602	0.329	0.476
Panel B: DHS Data	*Dependent variable:*			
	Registered	Electricity	Water Access	State Presence Index
	(1)	(2)	(3)	(4)
Log Distance to HQ	-0.148^{***}	-0.311^{***}	0.136^{***}	-0.216^{***}
	(0.009)	(0.010)	(0.013)	(0.007)
Fixed effects	Border Region	Border Region	Border Region	Border Region
Controls	✓	✓	✓	✓
Observations	21,178	30,239	29,150	30,239
Adjusted R^2	0.713	0.559	0.402	0.624

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of OLS regressions with log-distance to the administrative headquarters as the independent variable and individual components of the state presence index as the dependent variables. Standard errors, clustered at the district level, are shown in parentheses.

Table A4: Effect of Distance to State on Components of Development Index

	Dependent variable:		
	Literacy	Wealth	Piped Water
	(1)	(2)	(3)
Remoteness Treatment	-0.028^{**}	-0.070^{***}	-0.050^{**}
	(0.012)	(0.019)	(0.021)
Treatment × Recognized	-0.055^{**}	-0.106^{***}	-0.133^{***}
	(0.027)	(0.036)	(0.047)
Fixed effects?	Border Region	Border Region	Border Region
Controls	✓	✓	✓
Observations	3,061	3,516	3,563
Adjusted R^2	0.813	0.712	0.586

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of OLS regressions with individual components of the development index as the dependent variables. Standard errors, clustered at the district level, are shown in parentheses.

Table A5: Effect of Distance to State on Components of Traditional Leader Z-Score and Additional Variables

	Dependent variable:						
	Traditional Leader Z-Score	Influence of TL	Contact with TL	Trust in TL	TL not Corrupt	Performance of TL	TL Listens
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Remoteness Treatment	0.154***	0.297	0.159**	0.058	0.246***	0.074	0.449*
	(0.055)	(0.195)	(0.071)	(0.068)	(0.079)	(0.073)	(0.268)
Treatment × Recognized	−0.219***	−0.332	−0.162*	−0.074	−0.285***	−0.239**	−0.457
	(0.067)	(0.271)	(0.084)	(0.082)	(0.101)	(0.098)	(0.293)
Fixed effects?	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	703	157	703	536	536	375	155
Adjusted R^2	0.639	0.513	0.607	0.540	0.432	0.535	0.516

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of OLS regressions with the traditional leader z-score (1), its components (2-5), and two additional questions (6-7) as the dependent variable. Variables are described in Section A.I. Standard errors, clustered at the district level, are shown in parentheses.

D.III Geographic Outcomes

Table A6: Effect of Treatment on Historical and Geographical Controls using Afrobarometer and DHS Data

	Dependent variable:									
	Dist Capital	Dist Nat Border	Dist Coast	Elevation	Ruggedness	Agriculture	Hist Cities	Malaria	Missions	Dist Rail
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Remoteness Treatment	0.00005	0.035***	0.002	0.004	0.131	0.053	0.002	0.016	−0.002	0.027***
	(0.003)	(0.011)	(0.003)	(0.024)	(0.133)	(0.040)	(0.003)	(0.073)	(0.007)	(0.007)
Treatment × Recognized	0.008	−0.023	0.001	−0.003	−0.159	−0.064	0.011*	0.023	0.002	−0.008
	(0.006)	(0.027)	(0.006)	(0.039)	(0.213)	(0.071)	(0.006)	(0.078)	(0.013)	(0.012)
Fixed effects?	Border Region	BR	BR	BR	BR	BR	BR	BR	BR	BR
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	4,595	4,595	4,595	4,595	4,595	4,595	4,595	4,595	4,595	4,595
Adjusted R^2	0.999	0.994	1.000	0.986	0.644	0.936	1.000	0.955	0.998	0.997

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of OLS regressions with various geographical and historical variables as dependent variables. Variables are described in Section A.II. Standard errors, clustered at the district level, are shown in parentheses.

D.IV Robustness Checks

Table A7: Robustness: Different Measures of Institutional Context

	Dependent variable:			
	Traditional Leader Z-Score			
	(1)	(2)	(3)	(4)
Remoteness Treatment	0.154***	0.185***	0.133***	0.063
	(0.055)	(0.058)	(0.049)	(0.048)
Treatment × Recognized	−0.219***			
	(0.067)			
Treatment × Mentioned		−0.253***		
		(0.069)		
Treatment × Protected			−0.227***	
			(0.062)	
Treatment × Salary				−0.125**
				(0.060)
Fixed effects	Border Region	Border Region	Border Region	Border Region
Controls	✓	✓	✓	✓
Observations	703	703	703	703
Adjusted R^2	0.639	0.637	0.636	0.641

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of the same specification as Table 3 in Column (1). Instead of using whether traditional authorities are institutionalized in the constitution, it interacts treatment with Baldwin (2016) measure of whether traditional authorities are mentioned in the constitution (Column 2) or protected in the constitution (Column 3). Column (4) interacts treatment with an indicator if traditional leaders in the country receive an official salary from the state. Standard errors, clustered at the district level, are shown in parentheses.

Table A8: Effect of Treatment on Migration

	Dependent variable:			
	Migration			
	Children	Men	Women	Z-score
	(1)	(2)	(3)	(4)
Remoteness Treatment	0.019	−0.047	−0.017	−0.035
	(0.024)	(0.053)	(0.039)	(0.022)
Treatment × Recognized	−0.012	0.056	−0.016	0.041
	(0.050)	(0.064)	(0.067)	(0.037)
Fixed effects	Border Region	Border Region	Border Region	Border Region
Controls	✓	✓	✓	✓
Observations	3,088	1,519	1,621	3,135
Adjusted R^2	0.346	0.122	0.204	0.566

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of OLS regressions with various measures of migration as outcomes. It follows the same specification as Table 4. The following dependent variables from the DHS survey are used: Column (1): Percentage of children that do not live at home. Column (2): Percentage of men that have always lived in their current location. Column (3): Percentage of women that have always lived in their current location. Column (4): Z-score combination of the three measures. Standard errors, clustered at the district level, are shown in parentheses.

Table A9: Robustness: Different Specifications

	Dependent variable:							
	Traditional Leader Z-Score							
	Main	No Controls	Binary Treatment	Absolute Distance	No Scaling	Long/Lat	Cluster	Drop Misassigned
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Remoteness Treatment	0.154***	0.113**	0.413**	0.113*	0.164***	0.094**	0.154**	0.136**
	(0.055)	(0.047)	(0.187)	(0.066)	(0.059)	(0.045)	(0.068)	(0.057)
Treatment × Institutionalized	−0.219***	−0.145**	−0.677***	−0.229**	−0.208***	−0.103*	−0.219***	−0.203***
	(0.067)	(0.057)	(0.245)	(0.077)	(0.075)	(0.053)	(0.079)	(0.071)
Fixed effects	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Observations	703	801	703	703	703	703	703	654
Adjusted R^2	0.639	0.637	0.640	0.641	0.638	0.632	0.639	0.615

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of the same specification as Table 3 for Column (1). Column (2) removes geographical controls. Column (3) uses only a binary treatment. Column (4) uses absolute log-distance to the hq instead of the treatment indicator. Column (5) does not scale the treatment variable. Column (6) uses a long-lat specification similar to Dell (2010). Column (7) clusters at the highest admin. division. Column (8) shows the result after removing observations that have a different treatment assignment when using their own distance as opposed to the average distance on their side. Standard errors, clustered at the district level, are shown in parentheses.

Table A10: Robustness: Different Measurement

	Dependent variable:						
	Traditional Leader Z-Score						
	Main	Drop 100 km	Drop 50 km	No Restriction	Non-Logged	Traveltime	Rural Only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Remoteness Treatment	0.154***	0.179***	0.189***	0.174***	0.108*	0.139**	0.200**
	(0.055)	(0.069)	(0.070)	(0.048)	(0.055)	(0.065)	(0.091)
Treatment × Recognized	−0.219***	−0.238***	−0.250***	−0.233***	−0.194***	−0.223***	−0.276*
	(0.067)	(0.080)	(0.095)	(0.061)	(0.062)	(0.075)	(0.148)
Fixed effects	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	703	695	660	749	703	683	306
Adjusted R^2	0.639	0.644	0.643	0.638	0.639	0.636	0.593

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of the same specification as Table 3 for Column (1). Column (2) drops outliers farther than 100 km away from their administrative headquarters. Column (3) drops observations more than 50 km away. Column (4) includes observations that do not have an observation on the other side of the border within 30 km. Column (5) uses non-logged distance. Column (6) uses travel time to the administrative headquarters instead of straight distance. Column (7) restricts to rural observations. Standard errors, clustered at the district level, are shown in parentheses.

Table A11: Robustness: Headquarters and Boundaries

	Dependent variable:							
	Traditional Leader Z-Score							
	Main	Neighbor HQ	Admin 1	Admin 2	Donut RD	Ethnicity FE	Instrumented HQs	Placebo
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Remoteness Treatment	0.154***	0.176**	0.181***	0.097	0.056	0.162**	0.140***	0.084
	(0.055)	(0.072)	(0.064)	(0.087)	(0.057)	(0.065)	(0.053)	(0.061)
Treatment × Recognized	−0.219***	−0.229**	−0.249***	−0.149	−0.133*	−0.215***	−0.134**	−0.113
	(0.067)	(0.094)	(0.083)	(0.105)	(0.079)	(0.076)	(0.064)	(0.076)
Fixed effects	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Observations	703	558	392	311	560	701	726	731
Adjusted R^2	0.639	0.609	0.655	0.611	0.621	0.641	0.627	0.633

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of the same specification as Table 3 for Column (1). Column (2) controls for distance to the neighboring headquarters. Columns (3) and (4) only uses the first and second administrative division in each country respectively. Column (5) includes ethnic homeland fixed effects. Column (6) uses instrumented locations for the administrative headquarters based on 1960 population density. Column (7) shows the effect of distance to randomly assigned "placebo" headquarters. Standard errors, clustered at the district level, are shown in parentheses.

Table A12: Additional Robustness

	Dependent variable:		
	Traditional Leader Z-Score		
	Main	British Colonies	Drop Non-DHS Countries
	(1)	(2)	(3)
Remoteness Treatment	0.154^{***}	0.196^{**}	0.186^{***}
	(0.055)	(0.082)	(0.069)
Treatment × Recognized	-0.219^{***}	-0.261^{***}	-0.156^{*}
	(0.067)	(0.090)	(0.088)
Fixed effects?	Border Region	Border Region	Border Region
Controls	✓	✓	✓
Observations	703	527	452
Adjusted R^2	0.639	0.642	0.632

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of the same specification as Table 3 for Column (1). Column (2) restricts the sample to former British colonies. Column (3) drops countries which are not included in the DHS data from the Afrobarometer sample.

D.V Endogenous Institutions

Table A13: Covariate Balance—Country-Level Variables

	Not Recognized		Recognized		
Covariates (country level)	**N**	**Mean**	**N**	**Mean**	**p-value**
Historical Centralization	16	0.33	10	0.59	0.02
Year of Independence	16	1,954.31	11	1,961.82	0.43
Violent Independence?	16	0.19	11	0.27	0.63
Slave Exports	16	376,818.21	11	169,121.67	0.29
Population in 1400	16	973,040.63	11	439,638.09	0.16
Log Settler Mortality	15	6.07	5	5.41	0.40
British Colony	16	0.19	11	0.91	0.00
British Legal Origins	16	0.25	11	0.91	0.00
Settler Colony	16	0.13	11	0.36	0.19
Colonial Railroads (km)	16	962.36	11	921.50	0.91
Gemstones	16	2,014.94	11	40,045.45	0.11
Soil Quality	16	39.19	11	29.79	0.24
Average Distance to Coast	16	19.07	11	9.77	0.18
Land area (1000 Ha)	16	43,710.94	11	51,110.18	0.65
Ruggedness	16	0.51	11	1.24	0.20
Oil Production in 2000	16	8,285.33	11	60.62	0.26
Malaria Suitability	16	16.70	11	7.97	0.01
Rule of Law	16	-0.88	11	-0.37	0.03
GDP 1950	16	893.60	11	924.22	0.91
Failed State Index 2006	15	86.11	10	81.02	0.43
Taxes as % of GDP 2010	13	13.46	8	16.87	0.24
Democracy Index 2017	16	4.72	11	5.68	0.10
Political Decentralization	13	1.92	8	2.24	0.55

Notes: Difference in means between countries where traditional leaders are recognized and where they are not. All reported p-values are from two-sided t-tests.

Table A14: Robustness: Interaction with Country Variables

	Dependent variable:						
	Traditional Leader Z-Score						
	Hist. Central.	Year Indep.	Violent Indep.	Slave Export	Settler Mortality	Colonial Rail	Soil Quality
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Low Local State Capacity	0.099**	0.084*	0.125***	0.112***	0.072	0.115**	0.124***
	(0.047)	(0.050)	(0.043)	(0.043)	(0.047)	(0.045)	(0.045)
Treatment × Recognized	−0.144**	−0.165***	−0.163***	−0.142**	−0.075	−0.159***	−0.177***
	(0.063)	(0.056)	(0.054)	(0.059)	(0.084)	(0.055)	(0.054)
Treatment × CountryVariable	−0.031	0.087	−0.037*	0.046	0.120*	−0.001	−0.007
	(0.022)	(0.055)	(0.021)	(0.049)	(0.061)	(0.021)	(0.028)
Fixed effects	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	668	703	703	703	577	703	703
Adjusted R^2	0.600	0.641	0.639	0.638	0.594	0.638	0.637

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of the main specification but also includes the interaction of treatment with several country-level variables to control for possible confounding factors. This results in the following specification: $Y_{i,s,r} = \beta_0 + \beta_1 Tint_s + \beta_2 DB_i + \beta_3 T_s \times DB_i + \beta_4 Tint_s \times Recognized + \beta_5 DB_i \times Recognized + \beta_6 T_s \times DB_i \times Recognized + \beta_7 Tint_s \times CountryVariable + \beta_8 DB_i \times CountryVariable + \beta_9 T_s \times DB_i \times CountryVariable + \beta_{10}\chi_i + \beta_{11} BR_r + \epsilon$. Standard errors, clustered at the district level, are shown in parentheses.

Table A15: Robustness: Interaction with Country Variables

	Dependent variable:						
	Near Coast	Land Area	Oil Production	Traditional Leader Z-Score RGDP 1950	Years Schooling	Fragile State Index	Tax Revenue over GDP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Low Local State Capacity	0.107**	0.124***	0.126***	0.046	0.075	0.115**	0.081*
	(0.043)	(0.047)	(0.044)	(0.053)	(0.057)	(0.046)	(0.045)
Treatment × Recognized	−0.214***	−0.175***	−0.173***	−0.139***	−0.098	−0.185***	−0.167**
	(0.063)	(0.052)	(0.053)	(0.053)	(0.066)	(0.058)	(0.067)
Treatment × CountryVariable	−0.088*	−0.047	0.020	−0.133**	−0.080	0.024	−0.106*
	(0.048)	(0.042)	(0.046)	(0.059)	(0.060)	(0.026)	(0.062)
Fixed effects	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region	Border Region
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	703	703	703	703	663	666	617
Adjusted R^2	0.639	0.639	0.639	0.642	0.646	0.595	0.590

Notes: *p<0.1; **p<0.05; ***p<0.01. This table shows the results of the main specification but also includes the interaction of treatment with several country-level variables to control for possible confounding factors. The specification as in Table A14 is used. Standard errors, clustered at the district level, are shown in parentheses.

E References

Acemoglu, Daron, Garcia-Jimeno Camilo and James A. Robinson. 2015. "State Capacity and Economic Development: A Network Approach." *American Economic Review* 105(8):2364–2409.

Alegana, Victor A., Jim A. Wright, Uusiku Pentrina, Abdisalan M. Noor, Robert W. Snow and Peter M. Atkinson. 2012. "Spatial Modelling of Healthcare Utilisation for Treatment of Fever in Namibia." *International Journal of Health Geographics* 11(1):1–13.

Armstrong, Timothy B. and Michal Kolesar. 2017. "A Simple Adjustment for Bandwidth Snooping." *Review of Economics and Statistics* 85(2):732–765.

Calonico, Sebastian, Matias D. Cattaneo and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6):2295–2326.

Chandler, Tertius. 1987. *Four Thousand Years of Urban Growth: An Historical Census*. Lewiston, NY: St. David's University Press.

Gottlieb, Jessica, Guy Grossman, Horacio Larreguy and Benjamin Marx. 2019. "The Effect of Administrative Unit Creation on Electoral Behavior: Evidence from Senegal." *Journal of Politics* 6(3): 192–198.

Grossman, Guy and Janet I. Lewis. 2014. "Administrative Unit Proliferation." *American Political Science Review* 108(1):196–217.

Imbens, Guido and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economics and Statistic*s 79(3):933–959.

Jedwab, Remi and Alexander Moradi. 2015. "The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa." *Review of Economics and Statistics* 98(2):268–284.

Kiszewski, Anthony, Andrew Mellinger, Andrew Spielman, Pia Malaney, Sonia Ehrlich Sachs and Jeffrey Sachs. 2004. "A Global Index Representing the Stability of Malaria Transmission." *American Journal of Tropical Medicine and Hygiene* 70(5):486–498.

Nunn, Nathan. 2010. "Religious Conversion in Colonial Africa." *American Economic Review* 100(2):147–152.

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