

"The Employment Effects of Ethnic Politics"

by Francesco Amodio, Giorgio Chiovelli and Sebastian Hohmann

Online Appendix

Table A.1: Characteristics of Ethnic Groups With and Without Party Links

	No Party Link	Has Party Link	Difference
<i>Panel A: Contemporary Characteristics</i>			
Population Share	0.010	0.049	0.038 (0.007)
Employed (Share)	0.595	0.540	-0.055 (0.022)
Age	29.079	29.286	0.207 (0.314)
Rural (Share)	0.625	0.634	0.009 (0.025)
Completed Primary (Share)	0.249	0.313	0.064 (0.019)
Secondary Educ. or Higher (Share)	0.522	0.399	-0.123 (0.025)
Employed in Agriculture (Share)	0.189	0.254	0.065 (0.019)
Employed in Manufacturing (Share)	0.119	0.089	-0.030 (0.014)
Employed in Services (Share)	0.286	0.207	-0.079 (0.018)
Total	289	247	
<i>Panel B: Pre-colonial Characteristics</i>			
Gathering	0.446	0.368	-0.078 (0.039)
Hunting	0.547	0.747	0.200 (0.037)
Fishing	0.539	0.587	0.047 (0.040)
Animal Husbandry	1.790	1.640	-0.150 (0.083)
Agriculture	6.326	6.188	-0.138 (0.105)
Jurisdictional Hierarchy of Local Community	1.797	1.843	0.046 (0.031)
Jurisdictional Hierarchy Beyond Local Community	1.191	1.248	0.057 (0.073)
Class Stratification	0.632	0.503	-0.129 (0.043)
Slavery	0.980	0.921	-0.059 (0.019)
Elections	0.074	0.090	0.016 (0.025)
Total	267	375	

Notes. The table reports the average characteristics of ethnic groups across countries and separately for ethnicities without and with political party linkages as determined by the labeling algorithm described in Section 2. It also reports the difference in means across the two groups and the corresponding estimated standard error in parenthesis. Panel A focuses on contemporary characteristics as derived from DHS individual survey data. Panel B focuses on pre-colonial characteristics (Murdock 1967; Michalopoulos and Papaioannou 2013).

Table A.2: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	Observations
<i>Panel A: Overall Sample</i>					
Vote Margin (Adjusted)	0.092	0.388	-0.992	1	408187
Elected Local Ethnic Politician	0.586	0.493	0	1	408187
Employed	0.577	0.494	0	1	411199
Employed in Agriculture	0.287	0.452	0	1	411199
Employed in Manufacturing	0.115	0.319	0	1	411199
Employed in Services	0.199	0.399	0	1	411199
Employed in Public Sector	0.018	0.134	0	1	411199
Agriculture, Self-employed	0.238	0.426	0	1	411199
Agriculture, for Someone Else	0.049	0.215	0	1	411199
Owens Land	0.663	0.473	0	1	350153
In-kind Pay	0.022	0.148	0	1	411199
Age	29.044	10.122	15	64	411199
Rural	0.636	0.481	0	1	411199
Female	0.722	0.448	0	1	411199
Completed Primary School	0.334	0.472	0	1	411180
Secondary Education or Higher	0.352	0.478	0	1	411199
Latitude	1.734	11.132	-26.817	16.656	411199
Longitude	11.209	19.521	-17.498	41.877	411199
Distance from Improved Roads (km)	15.789	27.125	0	306.802	411199
Distance from Cities (km)	33.081	27.555	0.077	161.295	411199
Elevation (m)	543.165	551.151	-0.556	3193.583	411199
Terrain Ruggedness	53.255	71.41	1.024	1165.691	411199
Agricultural Suitability	0.389	0.212	0	0.987	394734
Malaria Suitability	14.955	10.014	0	37.609	411199
<i>Panel B: Observations Within 20% Right of the Threshold</i>					
Vote Margin (Adjusted)	0.094	0.056	0	0.2	72712
Elected Local Ethnic Politician	1	0	1	1	72712
Employed	0.573	0.495	0	1	72712
Employed in Agriculture	0.289	0.453	0	1	72712
Employed in Manufacturing	0.109	0.311	0	1	72712
Employed in Services	0.201	0.401	0	1	72712
Employed in Public Sector	0.019	0.137	0	1	72712
Agriculture, Self-employed	0.243	0.429	0	1	72712
Agriculture, for Someone Else	0.046	0.209	0	1	72712
Owens Land	0.683	0.465	0	1	61668
In-kind Pay	0.02	0.14	0	1	72712
Age	29.176	10.148	15	64	72712
Rural	0.656	0.475	0	1	72712
Female	0.722	0.448	0	1	72712
Completed Primary School	0.369	0.483	0	1	72707
Secondary Education or Higher	0.353	0.478	0	1	72712
Latitude	-0.577	11.111	-26.046	16.532	72712
Longitude	15.645	18.544	-16.73	41.83	72712
Distance from Improved Roads (km)	18.859	31.942	0	302.652	72712
Distance from Cities (km)	32.96	28.888	0.077	155.746	72712
Elevation (m)	590.933	542.853	-0.556	3193.583	72712
Terrain Ruggedness	63.956	88.446	1.024	1165.691	72712
Agricultural Suitability	0.421	0.21	0	0.987	69747
Malaria Suitability	14.274	10.15	0	37.609	72712
<i>Panel C: Observations Within 20% Left of the Threshold</i>					
Vote Margin (Adjusted)	-0.097	0.054	-0.2	0	69828
Elected Local Ethnic Politician	0	0	0	0	69828
Employed	0.559	0.497	0	1	69828
Employed in Agriculture	0.261	0.439	0	1	69828
Employed in Manufacturing	0.118	0.323	0	1	69828
Employed in Services	0.2	0.4	0	1	69828
Employed in Public Sector	0.021	0.142	0	1	69828
Agriculture, Self-employed	0.227	0.419	0	1	69828
Agriculture, for Someone Else	0.035	0.183	0	1	69828
Owens Land	0.644	0.479	0	1	55611
In-kind Pay	0.02	0.138	0	1	69828
Age	28.926	10.032	15	64	69828
Rural	0.63	0.483	0	1	69828
Female	0.724	0.447	0	1	69828
Completed Primary School	0.367	0.482	0	1	69825
Secondary Education or Higher	0.365	0.481	0	1	69828
Latitude	0.315	10.729	-25.977	16.532	69828
Longitude	13.747	19.302	-16.723	41.249	69828
Distance from Improved Roads (km)	16.811	28.447	0	300.631	69828
Distance from Cities (km)	32.055	26.665	0.124	161.295	69828
Elevation (m)	612.534	556.289	0.286	2804.455	69828
Terrain Ruggedness	57.266	70.892	1.664	846.678	69828
Agricultural Suitability	0.38	0.22	0	0.925	67143
Malaria Suitability	13.554	9.747	0	35.064	69828

Notes. The table reports the summary statistics of all variables used in the empirical analysis, described in Section 2 and with additional details in Supplementary Appendix B.5. Panel A shows the summary statistics for the overall final sample, while Panel B and C provide summary statistics for the subsample of observations within an adjusted vote share of 20% right and left respectively of the threshold that is relevant for gaining a local ethnic party representative from the constituency in the national assembly.

Table A.3: Sample Distribution by Country and Election Year

Countries	Election Years	DHS Survey Years	Observations	Frequency (%)
Benin	1999	2001	5357	1.3
	2011	2012	5731	1.39
Burkina Faso	2002	2003	9399	2.29
	2007	2010	10721	2.61
	2012	2014	5280	1.28
Cameroon	2002	2004	8748	2.13
	2007	2011	11048	2.69
Ghana	1996	1998, 1999	4819	1.17
	2000	2003	7792	1.89
	2004	2008	7584	1.84
	2012	2014	13216	3.21
Ivory Coast	2011	2012	7160	1.74
Kenya	2002	2003	872	0.21
	2007	2008, 2009	3508	0.85
	2013	2014, 2015	35866	8.72
Liberia	2005	2008, 2009	3486	0.85
	2011	2013	11394	2.77
Malawi	1999	2000	6252	1.52
	2004	2005	1423	0.35
	2009	2010, 2012	24576	5.98
	2014	2015, 2016, 2017	33306	8.1
Mali	2013	2015	5325	1.29
Mozambique	2009	2011	14965	3.64
Nigeria	2007	2008, 2010	4139	1.01
	2011	2013, 2015	49458	12.03
Senegal	2007	2008, 2009, 2010, 2011	1890	0.46
	2012	2013, 2014, 2015, 2016	47637	11.58
Sierra Leone	2007	2008	9226	2.24
	2012	2013	21610	5.26
Uganda	2011	2014, 2015	3859	0.94
Zambia	2006	2007	7739	1.88
	2011	2013, 2014	27813	6.76
Total			411199	100

Notes. The first and second columns show the set of countries and election years respectively in our sample. The third column shows the DHS survey years that we match to each election round. The fourth column shows the total number of matched DHS observations of individuals belong to ethnic groups for which our linking procedure described in Section 2 identifies a link with a political party, using a labeling cutoff of 0.01%. The fifth column shows the observations frequency distribution.

Table A.4: Test of Balance I – Village Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Latitude	Longitude	Dist. to Roads	Dist. to Cities	Elevation	Ruggedness	Suitability	Malaria								
$\hat{\beta}_{OLS}$	0.0944 (0.1276)	-0.0019 (0.0084)	0.0066 (0.009)	2.3976 (1.9574)	0.9637 (0.5671)	3.7241 (3.0254)	-0.2168 (0.6204)	-56.4171 (19.1693)	-10.0596 (3.765)	-9.5233 (5.4449)	-4.0321 (2.3744)	0.0098 (0.0202)	-0.0012 (0.0045)	0.2928 (0.4906)	0.0819 (0.0862)	
$\hat{\beta}_{IK}$	0.0564 (0.1250)	0.0024 (0.0181)	0.0185 (0.0122)	3.3238 (2.0843)	0.3350 (0.4277)	2.3901 (2.7191)	-0.0883 (0.4246)	-43.7314 (29.1767)	-2.9228 (2.6675)	-1.6778 (5.8640)	-0.4233 (1.2418)	0.0078 (0.0171)	-0.0027 (0.0029)	0.1312 (0.4940)	-0.0516 (0.0932)	
$\hat{\beta}_{CCT}$	0.0364 (0.1507)	0.0047 (0.0204)	0.0183 (0.0142)	3.7608 (2.4623)	0.3881 (0.5339)	2.4058 (3.2300)	-0.0173 (0.5020)	-47.5113 (34.2539)	-3.0362 (3.0099)	-1.2884 (6.6965)	-0.0146 (1.4162)	0.0068 (0.0196)	-0.0032 (0.0034)	0.1646 (0.5428)	-0.0760 (0.1121)	
Bandwidth	0.247	0.165	0.198	0.203	0.277	0.164	0.237	0.281	0.231	0.219	0.144	0.151	0.241	0.173	0.239	0.330
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	170214	121027	141431	144900	186465	120395	166731	189055	163109	155360	109654	113416	162029	121141	168170	216745

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable each village-level covariate as indicated in the column header. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanam 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calónico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calónico, Cattaneo, and Titiunik (2014). Standard errors are clustered at the ethnic group level.

Table A.5: Test of Balance I – Individual Characteristics

	Age		Rural		Female		Primary School	Secondary School or Higher		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\hat{\beta}_{OLS}$	0.1526 (0.1448)	0.0120 (0.1102)	0.0434 (0.0346)	0.0258 (0.0161)	0.0043 (0.0061)	-0.0008 (0.0067)	0.0034 (0.0087)	-0.0036 (0.0072)	-0.0291 (0.0191)	0.0030 (0.0122)
$\hat{\beta}_{IK}$	0.2222 (0.1389)	0.0514 (0.0755)	0.0492 (0.0343)	0.0130 (0.0092)	0.0084 (0.0062)	0.0010 (0.0035)	0.0013 (0.0092)	-0.0002 (0.0046)	-0.0243 (0.0170)	-0.0072 (0.0070)
$\hat{\beta}_{CCT}$	0.2179 (0.1717)	0.0463 (0.0910)	0.0589 (0.0371)	0.0149 (0.0109)	0.0110 (0.0071)	0.0010 (0.0043)	0.0004 (0.0110)	0.0001 (0.0057)	-0.0281 (0.0194)	-0.0078 (0.0085)
Bandwidth	0.209	0.332	0.198	0.229	0.154	0.249	0.279	0.256	0.211	0.208
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	149794	216856	141431	162586	114666	171140	188273	175044	151329	149420

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable each village-level covariate as indicated in the column header. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaram 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calónico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calónico, Cattaneo, and Titiunik (2014). Standard errors are clustered at the ethnic group level.

Table A.6: Test of Balance II – Village Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Latitude	Longitude	Dist. to Roads	Dist. to Cities	Elevation	Ruggedness	Suitability	Malaria								
$\hat{\beta}_{OLS}$	-0.2417 (0.1332)	-0.0200 (0.0064)	0.0416 (0.1187)	-0.0188 (0.0135)	1.0812 (3.7435)	1.1861 (1.0678)	-3.8480 (3.4314)	-0.7937 (0.7040)	-13.3612 (72.7809)	-9.0937 (8.8748)	11.5885 (11.9049)	-6.6231 (5.2478)	-0.0059 (0.0183)	-0.0076 (0.0055)	1.1415 (0.5978)	-0.2247 (0.1113)
$\hat{\beta}_{FK}$	-0.4177 (0.2161)	-0.0032 (0.0232)	0.0360 (0.1037)	-0.0217 (0.0695)	0.8335 (3.0686)	-0.0731 (0.9891)	-3.4558 (3.7027)	-0.8353 (1.0334)	-24.2662 (62.7203)	-1.0487 (7.5415)	14.1007 (13.0946)	2.5947 (5.0448)	0.0390 (0.0177)	0.0042 (0.0047)	0.8328 (0.6646)	-0.1342 (0.1729)
$\hat{\beta}_{CCT}$	-0.4732 (0.2393)	0.0027 (0.0276)	0.0426 (0.1134)	-0.0193 (0.0887)	0.1359 (3.4560)	-0.2667 (1.1983)	-4.9780 (4.3488)	-0.9225 (1.2732)	-36.7876 (77.0324)	-0.1795 (8.6844)	12.4854 (15.0734)	2.5966 (5.9418)	0.0428 (0.0202)	0.0047 (0.0052)	0.8138 (0.7521)	-0.1682 (0.2176)
Bandwidth	0.157	0.119	0.221	0.120	0.212	0.158	0.195	0.191	0.196	0.152	0.172	0.288	0.200	0.240	0.223	0.191
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	27277	21021	37439	21047	36014	27425	33294	32768	33484	26346	29715	47211	32700	38287	37502	32588

Notes. The unit of observation is an individual as surveyed in the DHS. Sample is restricted to those observations used in the placebo subsample, i.e. for whom it is possible to locate the adjusted vote share of the same party in the same constituency in the next election in our dataset. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable each village-level covariate as indicated in the column header. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{FK}$ is the conventional RD estimate (Imbens and Kalyanam 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calónico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calónico, Cattaneo, and Titiunik (2014). Standard errors are clustered at the ethnic group level.

Table A.7: Test of Balance II – Individual Characteristics

	Age		Rural		Female		Primary School		Secondary School or Higher	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\hat{\beta}_{OLS}$	0.2252 (0.2207)	0.0183 (0.492)	-0.0341 (0.0335)	0.0290 (0.0271)	-0.0071 (0.0092)	-0.0094 (0.0121)	-0.0244 (0.0199)	0.0135 (0.0162)	0.0118 (0.0217)	-0.0218 (0.0151)
$\hat{\beta}_{TK}$	0.3303 (0.2297)	0.3036 (0.2507)	-0.0233 (0.0489)	0.0267 (0.0284)	-0.0021 (0.0113)	0.0002 (0.0091)	-0.0258 (0.0211)	-0.0200 (0.0112)	-0.0105 (0.0262)	0.0011 (0.0140)
$\hat{\beta}_{CCT}$	0.3495 (0.2744)	0.3633 (0.3010)	-0.0223 (0.0610)	0.0364 (0.0348)	0.0013 (0.0145)	0.0032 (0.0112)	-0.0286 (0.0244)	-0.0235 (0.0132)	-0.0113 (0.0301)	-0.0002 (0.0168)
Bandwidth	0.194	0.202	0.282	0.197	0.266	0.252	0.171	0.171	0.207	0.240
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	33244	34762	45833	33538	43000	41475	29553	29553	35397	39833

Notes. The unit of observation is an individual as surveyed in the DHS. Sample is restricted to those observations used in the placebo subsample, i.e. for whom it is possible to locate the adjusted vote share of the same party in the same constituency in the next election in our dataset. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable each village-level covariate as indicated in the column header. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{TK}$ is the conventional RD estimate (Imbens and Kalyanam 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Standard errors are clustered at the ethnic group level.

Table A.8: Correlates of Political Competition

	Number of Parties					HHI
	(1)	(2)	(3)	(4)	(5)	(6)
Fractionalization	0.8588 (0.2148)		1.4403 (0.4261)	1.1791 (0.2867)	0.0578 (0.3147)	-0.0973 (0.0314)
Polarization		0.5038 (0.1899)	-0.5941 (0.3761)	-0.3900 (0.2525)	0.1096 (0.2546)	0.0263 (0.0254)
Share of Employed					0.5846 (0.3834)	-0.0908 (0.0382)
Share of Agric.					-1.0636 (0.3465)	0.0495 (0.0345)
Share of Services					-2.3951 (0.4173)	0.0408 (0.0416)
Population					0.0010 (0.0004)	-0.0000 (0.0000)
Share of Rural Pop.					-0.3868 (0.1451)	-0.0188 (0.0145)
Primary School					1.0026 (0.2788)	-0.0173 (0.0278)
Secondary School					1.8146 (0.2519)	-0.0331 (0.0251)
Country FE	No	No	No	Yes	Yes	Yes
Observations	2292	2292	2292	2292	2292	2292
R^2	0.0069	0.0031	0.0080	0.5899	0.6156	0.2343

Notes. The unit of observation is the electoral constituency in an election year. In columns 1 to 5, the dependent variable is the number of parties running in the constituency. In column 6, the dependent variable is a Herfindahl-type index of political competition calculated using vote shares by party. The independent variables included are derived from the individual-level DHS observations in the full sample, keeping only those belonging to the year that is after and closest to the election year.

Table A.9: Effect on Employment – Including Constituency × Election Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Employment Dummy			
$\hat{\beta}_{OLS}$	0.0194 (0.0246)	0.0219 (0.016)	0.0236 (0.0157)	0.0273 (0.0126)	0.0205 (0.013)	0.0169 (0.013)	0.0074 (0.0091)
$\hat{\beta}_{IK}$	0.0281 (0.0408)	0.0288 (0.0159)	0.0275 (0.0111)	0.0255 (0.0101)	0.0224 (0.0101)	0.0158 (0.0073)	0.0027 (0.0033)
$\hat{\beta}_{CCT}$	0.0382 (0.0455)	0.0340 (0.0179)	0.0302 (0.0128)	0.0293 (0.0115)	0.0250 (0.0117)	0.0179 (0.0085)	0.0030 (0.0040)
Bandwidth	0.226	0.205	0.218	0.255	0.220	0.190	0.189
Country × Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes	Yes
Constituency FE	No	No	No	No	No	Yes	n.a.
Constituency × Election FE	No	No	No	No	No	No	Yes
Observations	15861	145736	154040	166888	149387	132033	131579

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of $\hat{\beta}$ and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table A.10: Effect on Employment – Robustness

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0195 (0.0116)	0.0169 (0.013)	0.0191 (0.0238)	0.0306 (0.0209)	0.0061 (0.0194)	0.0144 (0.0232)
$\hat{\beta}_{IK}$	0.0193 (0.0075)	0.0158 (0.0073)	0.0127 (0.0053)	0.0113 (0.0059)	0.0143 (0.0085)	0.0162 (0.0090)
$\hat{\beta}_{CCT}$	0.0217 (0.0088)	0.0179 (0.0085)	0.0140 (0.0065)	0.0124 (0.0071)	0.0162 (0.0100)	0.0187 (0.0106)
Bandwidth	0.203	0.190	0.271	0.244	0.220	0.181
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	n.a.	n.a.	Yes	Yes
Constituency FE	Yes	Yes	n.a.	n.a.	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Ethnicity \times Const. FE	No	No	Yes	Yes	No	No
Observations	144310	132033	168243	149775	111409	90994

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. In column 3 and 4, sample is restricted to those countries for which we have information on more than one election round. In column 5 and 6, sample is restricted to those individuals belonging to ethnic groups other than the historically dominant one in the area as determined by matching the constituency centroid with the historical ethnic homelands in the Murdock map. Standard errors are clustered at the ethnic group level.

Table A.11: Effect on Employment – Polity IV Democracies

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0155 (0.0203)	0.0211 (0.0197)	0.0143 (0.0183)	0.0212 (0.0152)	0.0162 (0.0142)	0.0263 (0.0137)
$\hat{\beta}_{IK}$	0.0202 (0.0376)	0.0314 (0.0181)	0.0225 (0.0140)	0.0242 (0.0145)	0.0202 (0.0125)	0.0160 (0.0102)
$\hat{\beta}_{CCT}$	0.0248 (0.0417)	0.0357 (0.0202)	0.0260 (0.0164)	0.0285 (0.0171)	0.0247 (0.0144)	0.0176 (0.0120)
Bandwidth	0.309	0.218	0.222	0.213	0.236	0.216
Country \times Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	135639	100970	102975	94092	103762	94942

Notes. The unit of observation is an individual as surveyed in the DHS. Sample is restricted to those countries and election years with Polity IV score equal to 6 or higher. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table A.12: Effect on Agricultural Employment

	Agricultural Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0362 (0.0262)	0.0358 (0.0227)	0.0178 (0.0197)	0.0165 (0.0189)	0.0047 (0.0169)	0.0189 (0.0153)
$\hat{\beta}_{IK}$	0.0352 (0.0391)	0.0353 (0.0269)	0.0264 (0.0190)	0.0201 (0.0186)	0.0109 (0.0157)	0.0169 (0.0081)
$\hat{\beta}_{CCT}$	0.0429 (0.0440)	0.0393 (0.0315)	0.0278 (0.0234)	0.0205 (0.0236)	0.0094 (0.0188)	0.0185 (0.0096)
Bandwidth	0.226	0.196	0.190	0.232	0.208	0.203
Country \times Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	159426	139960	137495	157966	142693	138955

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working in agriculture. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table A.13: Agricultural Employment Type, Land Ownership, and In-kind Pay

	Agriculture Self-employed		Agriculture for Someone Else		Owns Land		In-kind Pay	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	0.0182 (0.0163)	0.0163 (0.0179)	0.0092 (0.0048)	0.0076 (0.0056)	-0.0066 (0.0111)	-0.0227 (0.0104)	0.0044 (0.0032)	0.0050 (0.0031)
$\hat{\beta}_{IK}$	0.0153 (0.0092)	0.0107 (0.0087)	0.0066 (0.0037)	0.0059 (0.0039)	-0.0038 (0.0093)	-0.0139 (0.0085)	0.0033 (0.0026)	0.0030 (0.0027)
$\hat{\beta}_{CCT}$	0.0164 (0.0108)	0.0115 (0.0102)	0.0077 (0.0044)	0.0069 (0.0046)	-0.0048 (0.0111)	-0.0161 (0.0097)	0.0040 (0.0033)	0.0036 (0.0034)
Bandwidth	0.187	0.185	0.164	0.170	0.177	0.170	0.248	0.247
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	136487	130203	120514	120093	106954	100516	170389	163588

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be self-employed in agriculture (columns 1 and 2), employed in the agriculture working for someone else (columns 3 and 4), own land (columns 5 and 6), and receive in-kind pay (columns 7 and 8). $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calónico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calónico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table A.14: Disfavoritism – Effect on Employment by Sector

	Agriculture		Manufacturing		Services		Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	-0.0372 (0.011)	-0.0250 (0.0081)	-0.0017 (0.0077)	-0.0032 (0.0088)	0.0245 (0.0116)	0.0160 (0.0086)	0.0018 (0.0042)	0.0018 (0.0038)
$\hat{\beta}_{IK}$	-0.0124 (0.0078)	-0.0044 (0.0081)	-0.0017 (0.0038)	-0.0023 (0.0040)	0.0130 (0.0056)	0.0094 (0.0047)	0.0025 (0.0017)	0.0023 (0.0016)
$\hat{\beta}_{CCT}$	-0.0149 (0.0083)	-0.0059 (0.0091)	-0.0030 (0.0049)	-0.0039 (0.0052)	0.0149 (0.0064)	0.0097 (0.0054)	0.0031 (0.0020)	0.0027 (0.0019)
Bandwidth	0.191	0.188	0.310	0.295	0.260	0.324	0.286	0.278
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	81782	75409	127307	114589	109440	127708	117573	108418

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of β and standard errors that we obtain when estimating equation 2, but considering the vote share of the party that is (the most) under-represented, relative to its overall vote share in the country, among voters from ethnic group e , and using a cutoff of 5%. The dependent variable is a dummy equal to one if the individual reports to be working in agriculture (columns 1 and 2), manufacturing (columns 3 and 4), services (columns 5 and 6), and the public sector (columns 7 and 8). $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calónico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calónico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table A.15: Disfavoritism – Effect on Employment

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0242 (0.0359)	0.0227 (0.0205)	0.0080 (0.0208)	0.0045 (0.023)	0.0095 (0.0221)	0.0080 (0.0092)
$\hat{\beta}_{IK}$	0.0366 (0.0514)	0.0132 (0.0203)	0.0061 (0.0191)	0.0072 (0.0188)	0.0110 (0.0184)	0.0053 (0.0068)
$\hat{\beta}_{CCT}$	0.0385 (0.0533)	0.0146 (0.0238)	-0.0003 (0.0222)	0.0035 (0.0222)	0.0059 (0.0218)	0.0060 (0.0073)
Bandwidth	0.237	0.289	0.253	0.234	0.232	0.311
Country \times Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	100259	119295	105838	92642	91900	120105

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of β and standard errors that we obtain when estimating equation 2, but considering the vote share of the party that is (the most) under-represented, relative to its overall vote share in the country, among voters from ethnic group e , and using a cutoff of 5%. The dependent variable is a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table A.16: Effect on Employment for Alternative Party Linkages

	Religion		Age		Literacy		Education	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	0.0219 (0.0153)	0.0285 (0.014)	0.0288 (0.0126)	0.0214 (0.0107)	-0.0012 (0.0142)	-0.0113 (0.0116)	0.0094 (0.0155)	0.0038 (0.0108)
$\hat{\beta}_{IK}$	0.0114 (0.0061)	0.0124 (0.0062)	0.0192 (0.0147)	0.0215 (0.0089)	-0.0002 (0.0070)	-0.0033 (0.0061)	0.0030 (0.0090)	0.0023 (0.0071)
$\hat{\beta}_{CCT}$	0.0116 (0.0073)	0.0130 (0.0073)	0.0196 (0.0158)	0.0229 (0.0098)	-0.0015 (0.0083)	-0.0048 (0.0071)	0.0013 (0.0105)	0.0013 (0.0086)
Bandwidth	0.307	0.314	0.269	0.195	0.317	0.308	0.235	0.284
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	184997	181817	154079	110641	179166	166540	120573	135342

Notes. The unit of observation is an individual as surveyed in the DHS. Party linkages are derived using the over-representation method discussed in Section 2 for groups defined by religion (column 1 and 2), age (column 3 and 4), literacy (column 5 and 6), and education (column 7 and 8). The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table A.17: Effect on Employment and Decentralization

	Employment Dummy							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>Not Decentralized</i>				<i>Decentralized</i>			
$\hat{\beta}_{OLS}$	0.0276 (0.0313)	0.0217 (0.0225)	0.0099 (0.0212)	0.0010 (0.0092)	0.0501 (0.0965)	0.0260 (0.0242)	0.0210 (0.0383)	0.0491 (0.0259)
$\hat{\beta}_{TK}$	0.0119 (0.0192)	0.0166 (0.0162)	0.0120 (0.0142)	0.0068 (0.0073)	0.1101 (0.0392)	0.0736 (0.0296)	0.0323 (0.0341)	0.0253 (0.0209)
$\hat{\beta}_{CCT}$	0.0139 (0.0198)	0.0198 (0.0175)	0.0152 (0.0153)	0.0102 (0.0087)	0.1219 (0.0500)	0.0876 (0.0389)	0.0423 (0.0389)	0.0311 (0.0255)
Difference (p-value)					[0.0003]	[0.0021]	[0.0638]	[0.4813]
Bandwidth	0.232	0.253	0.203	0.208	0.095	0.129	0.169	0.146
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Individual Controls	No	No	Yes	Yes	No	No	Yes	Yes
Constituency FE	No	No	No	Yes	No	No	No	Yes
Observations	45392	46482	39976	41322	6228	10296	12302	11366

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{TK}$ is the conventional RD estimate (Imbens and Kalyanam 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Columns 1 to 4 are estimated using the subsample of countries where local governments do not have jurisdiction over taxing, spending or legislation. Columns 5 to 8 report the same estimates using the subsample of countries where local governments have jurisdiction over taxing, spending or legislation. They also report in square brackets the p-value from a test of equality of $\hat{\beta}_{CCT}$ from the same specification across the two subsamples. Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table A.18: Characteristics of Agricultural Workers

	Some Primary			Age			Female			Rural		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)				
$\hat{\beta}_{OLS}$	-0.0299 (0.0149)	-0.0272 (0.0158)	-0.0043 (0.0209)	-0.0022 (0.0195)	-0.0001 (0.0178)	0.0127 (0.0163)	0.0120 (0.0202)	-0.0057 (0.0197)				
$\hat{\beta}_{IK}$	-0.0300 (0.0135)	-0.0231 (0.0134)	-0.0044 (0.0191)	-0.0017 (0.0194)	0.0062 (0.0157)	0.0094 (0.0141)	0.0112 (0.0206)	0.0019 (0.0182)				
$\hat{\beta}_{CCT}$	-0.0320 (0.0157)	-0.0250 (0.0159)	0.0017 (0.0225)	0.0050 (0.0228)	0.0047 (0.0193)	0.0102 (0.0176)	0.0157 (0.0246)	0.0040 (0.0221)				
Bandwidth	0.289	0.332	0.317	0.306	0.330	0.347	0.224	0.253				
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes				
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes				
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes				
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes				
Observations	53915	59028	57606	55130	59585	61376	43382	47655				

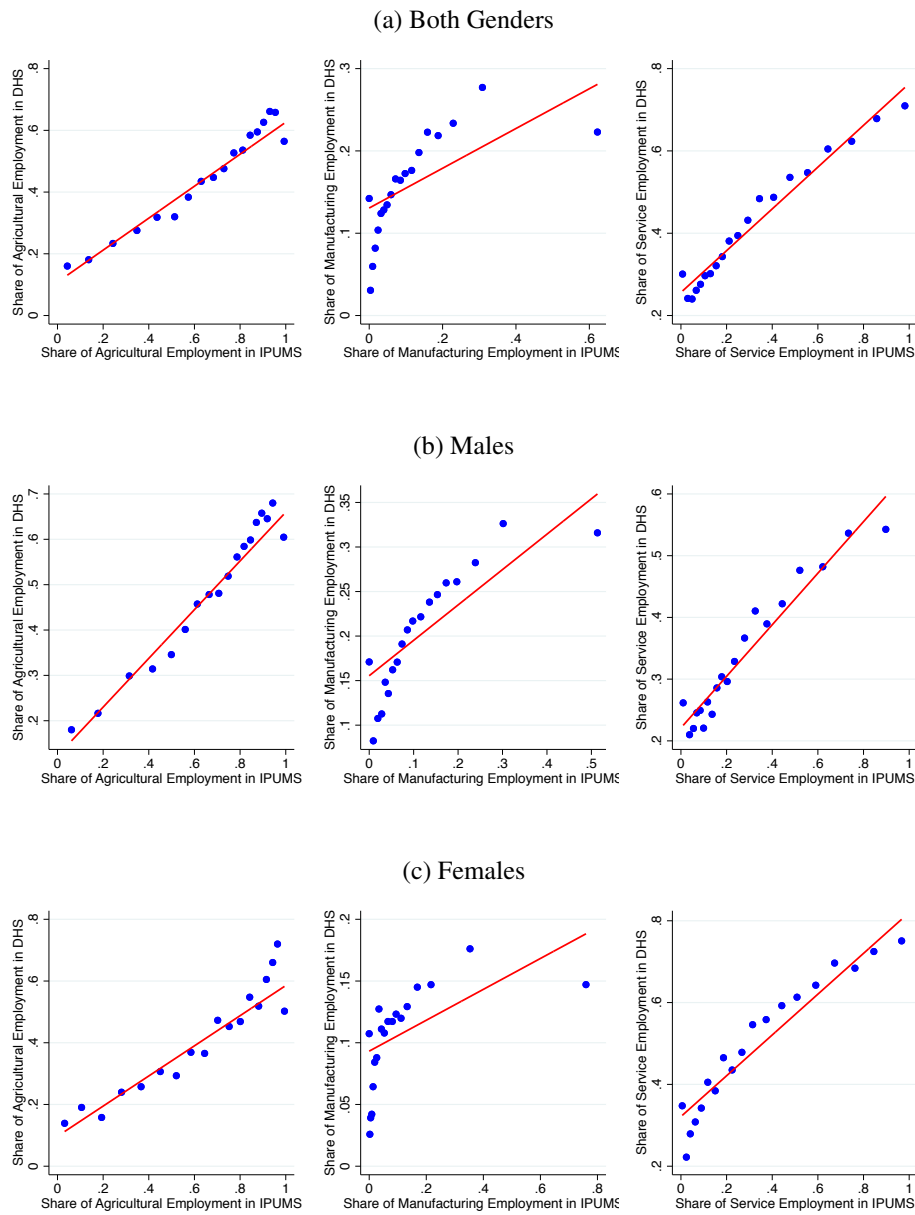
Notes. The unit of observation is an individual as surveyed in the DHS and reporting to be working in the agricultural sector. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to have some primary education (columns 1 and 2), individual's age in standard deviation (columns 3 and 4), a dummy for female (columns 5 and 6), and a dummy for rural (columns 7 and 8). $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calónico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calónico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Standard errors are clustered at the ethnic group level.

Table A.19: Characteristics of Workers in Services

	Some Primary		Age		Female		Rural	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	-0.0140 (0.0314)	-0.0101 (0.0233)	-0.0042 (0.0332)	0.0084 (0.0331)	0.0043 (0.0167)	0.0030 (0.0168)	0.0462 (0.0523)	0.0114 (0.0355)
$\hat{\beta}_{TK}$	-0.0180 (0.0208)	-0.0102 (0.0205)	-0.0070 (0.0292)	0.0052 (0.0269)	-0.0007 (0.0155)	-0.0012 (0.0160)	0.0362 (0.0488)	0.0084 (0.0376)
$\hat{\beta}_{CCT}$	-0.0233 (0.0240)	-0.0152 (0.0239)	-0.0134 (0.0331)	0.0037 (0.0310)	-0.0043 (0.0186)	-0.0039 (0.0193)	0.0507 (0.0517)	0.0147 (0.0407)
Bandwidth	0.127	0.141	0.255	0.300	0.275	0.264	0.219	0.229
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	16803	17718	30117	32839	32341	28838	27133	26638

Notes. The unit of observation is an individual as surveyed in the DHS and reporting to be working in the service sector. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to have some primary education (columns 1 and 2), individual's age in standard deviation (columns 3 and 4), a dummy for female (columns 5 and 6), and a dummy for rural (columns 7 and 8). $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{TK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calónico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calónico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Standard errors are clustered at the ethnic group level.

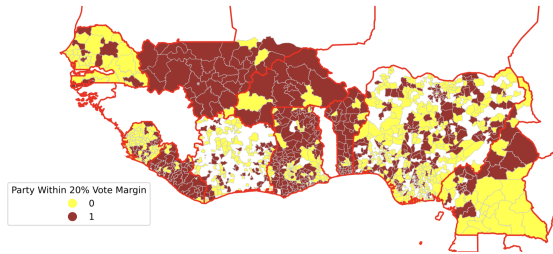
Figure A.1: Employment Shares by Sector: DHS vs. IPUMS



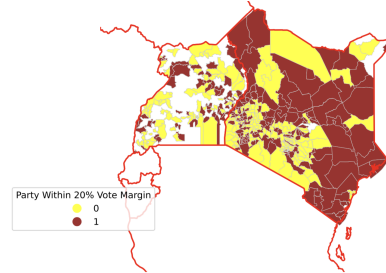
Notes. The figures plot the share of employment in each of the three main sectors – agriculture, manufacturing, and services – at the level of administrative subdivisions or regions and calculated from DHS data against the same variable calculated from IPUMS census data. They do so for those countries in our sample for which IPUMS data are available – Benin, Cameroon, Ghana, Liberia, Mozambique, Malawi, Nigeria, Sierra Leone and Zambia – and considering those DHS survey years that are closest in time to census years. The figures show binned scatter plots together with the corresponding linear regression line.

Figure A.2: Spatial Distribution of Contested Constituencies

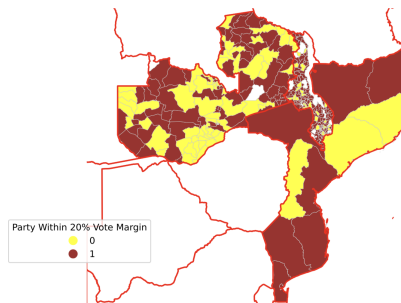
(a) West Africa



(b) East Africa

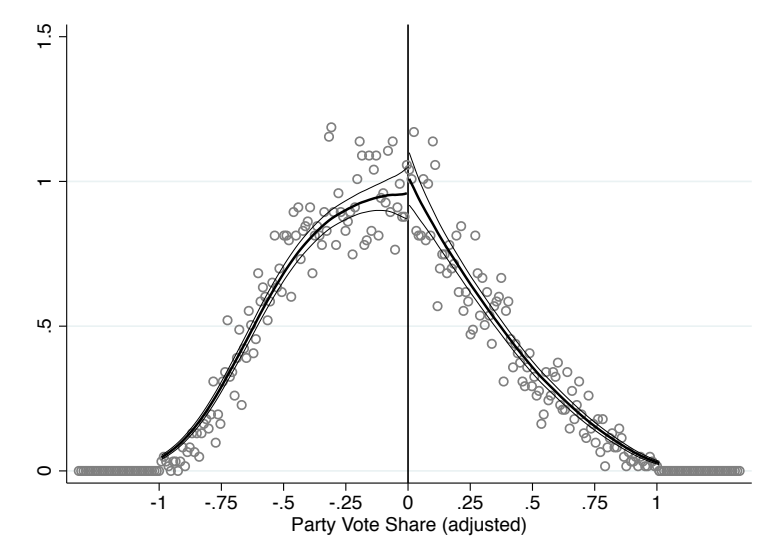


(c) Southern Africa



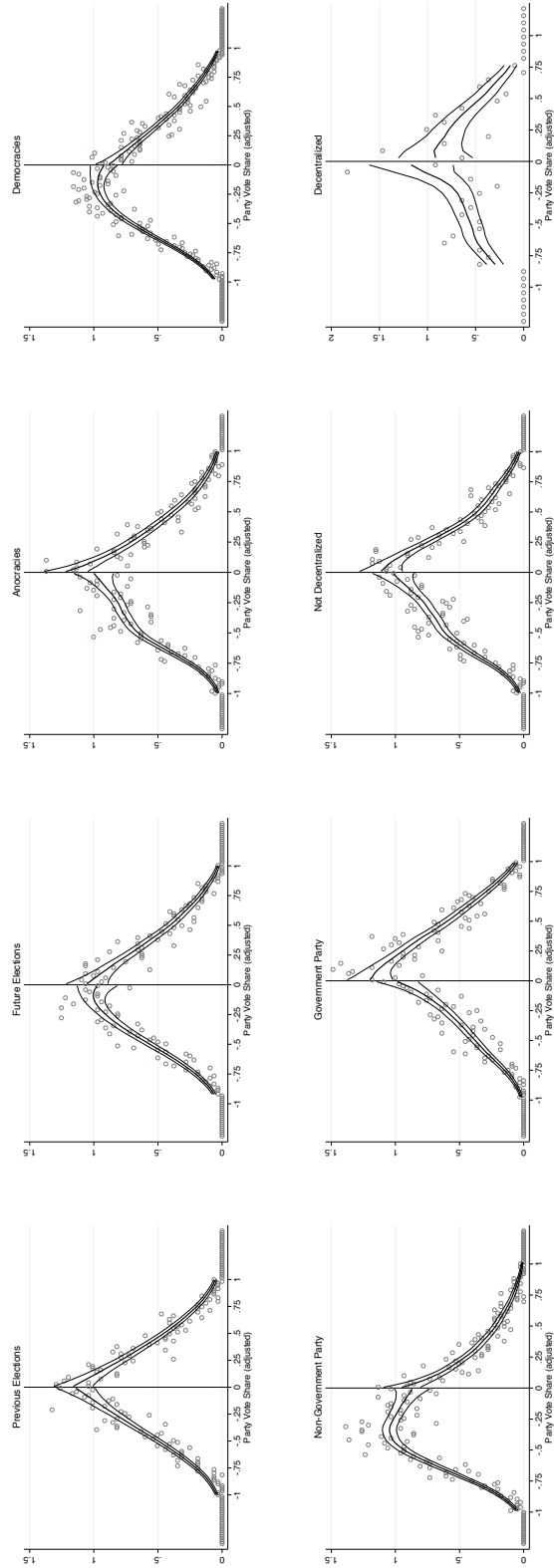
Notes. The figures show the spatial distribution of constituencies for West Africa (a), East Africa (b), and Southern Africa (c). We highlight in brown those constituencies in which, across all election rounds in our sample, we observe at least one party with a vote margin of less than 20% from the relevant threshold for winning a seat in parliament.

Figure A.3: Density Plot – McCrary Test Results



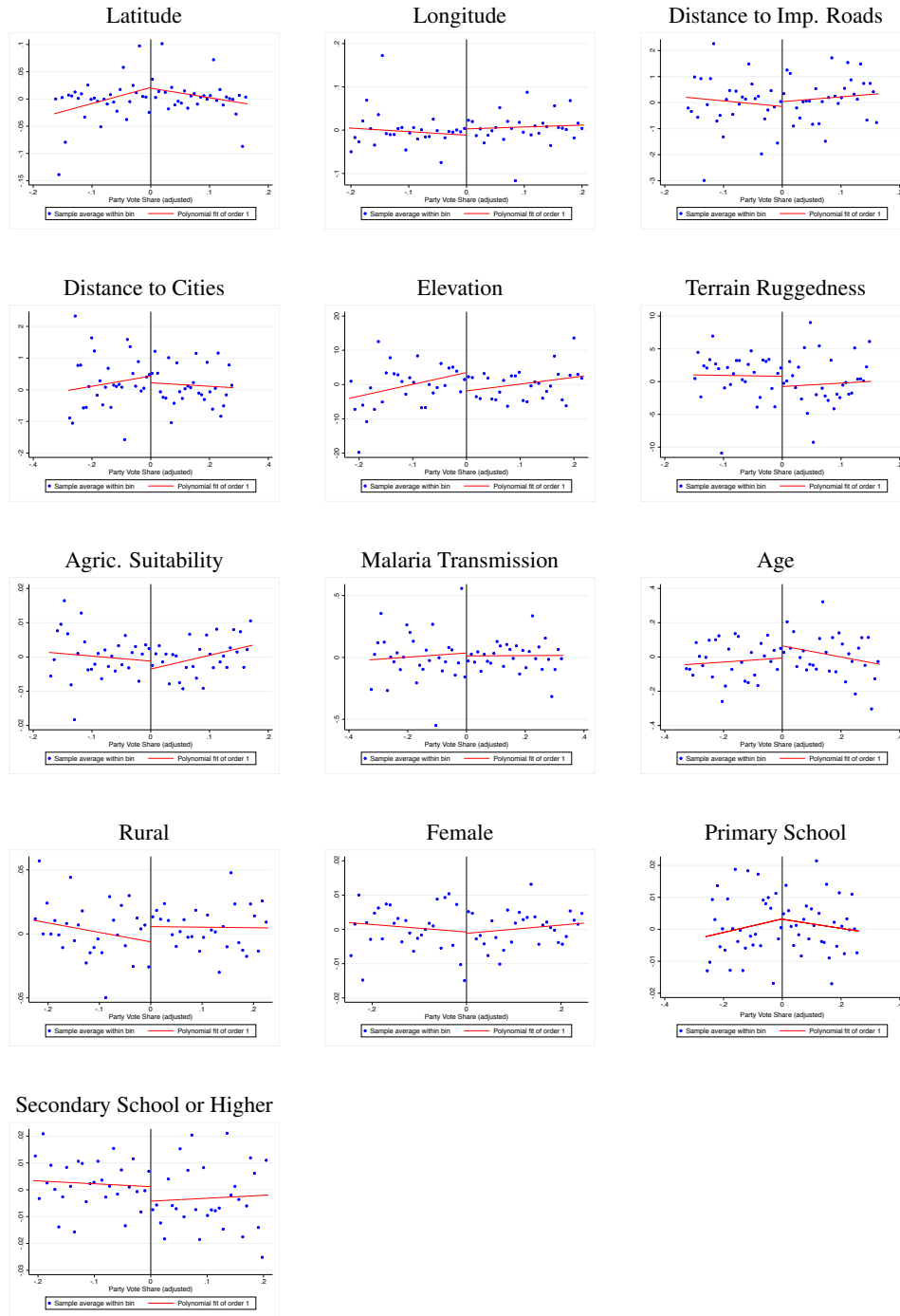
Notes. The figure shows the distribution of the running variable and its density function as estimated separately on both sides of the threshold, together with 95% confidence intervals. The p-value from a test of equality of the value of the density function on the left and right side of the threshold is equal to 0.4759 using the test by [McCrary \(2008\)](#), and equal to 0.3817 using the test by [Cattaneo, Jansson, and Ma \(2020\)](#). We therefore cannot reject the hypothesis of no discontinuity in the density of the running variable at the threshold.

Figure A.4: Density Plots – Additional McCrary Test Results



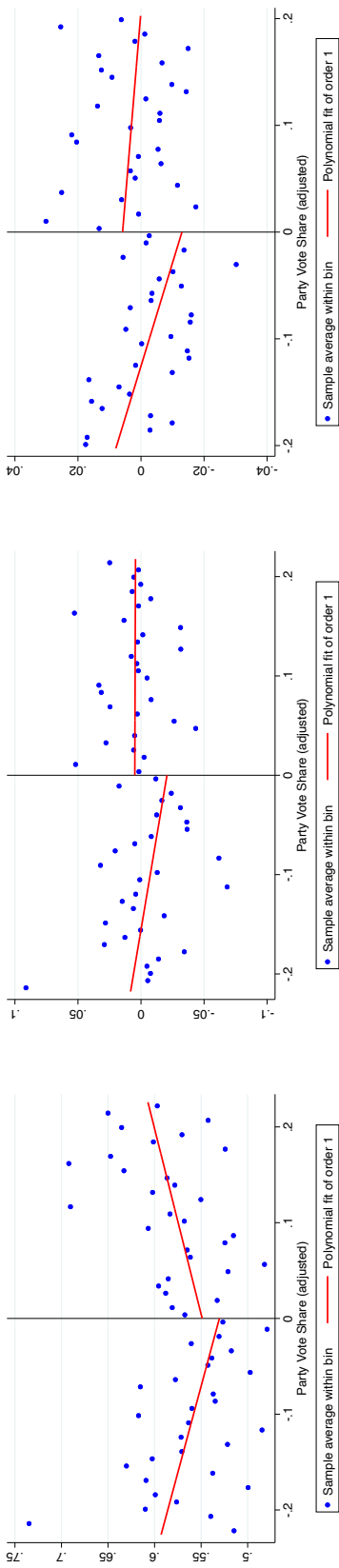
Notes. The figures show the distribution of the running variable and its density function as estimated separately on both sides of the threshold, together with 95% confidence intervals, and separately in different subsamples. In the subsample of observations used in the placebo subsample, i.e. for whom it is possible to locate the adjusted vote share of the same party in the same constituency in the next election in our dataset, considering the (adjusted) vote share in the last election, the p-value from a test of equality of the value of the density function on the left and right side of the threshold is equal to 0.5540 using the test by McCrary (2008), and equal to 0.5666 using the test by Cattaneo, Jansson, and Ma (2020). When considering the (adjusted) vote share in the next election, the corresponding p-values are 0.3234 and 0.1882. In the subsample of countries and election years with Polity IV score between -5 and +5 (anocracies), the corresponding p-values are 0.0573 and 0.1890. In the subsample of countries and election years with Polity IV score equal to 6 or higher (democracies), the corresponding p-values are 0.5723 and 0.8752. For parties that do not support the central government in the aftermath of elections, the corresponding p-values are 0.5161 and 0.0808. For parties that support the central government, the corresponding p-values are 0.0172 and 0.6996. In the subsample of countries where local governments do not have authority over taxing, spending, or legislating (not decentralized), the corresponding p-values are 0.4721 and 0.6424. In the subsample of countries where local governments do have authority over taxing, spending, or legislating (decentralized), the corresponding p-values are 0.4032 and 0.9871.

Figure A.5: Balancedness at the Threshold



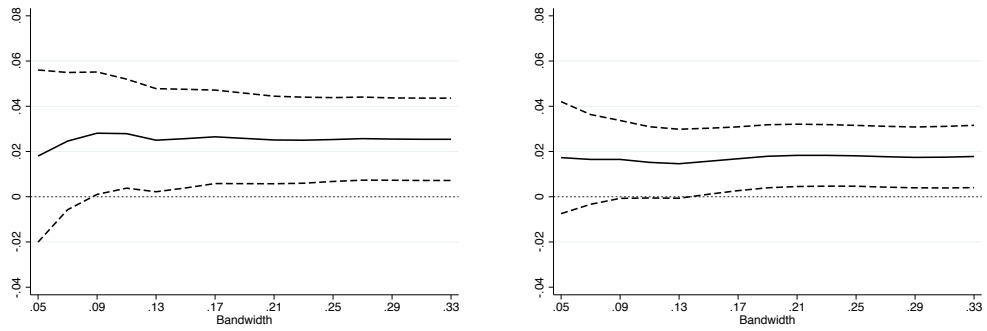
Notes. The figures provide a graphical representation of the bias-corrected local linear regression fit on both sides of the threshold (Calónico, Cattaneo, and Titiunik 2014). They also plot a scatterplot showing the average value of each average residual covariate net of country \times year, ethnicity, and constituency fixed effects in 30 bins right and left of the threshold. The figures show the absence of any meaningful discontinuities in the values of the covariates.

Figure A.6: Effect on Employment



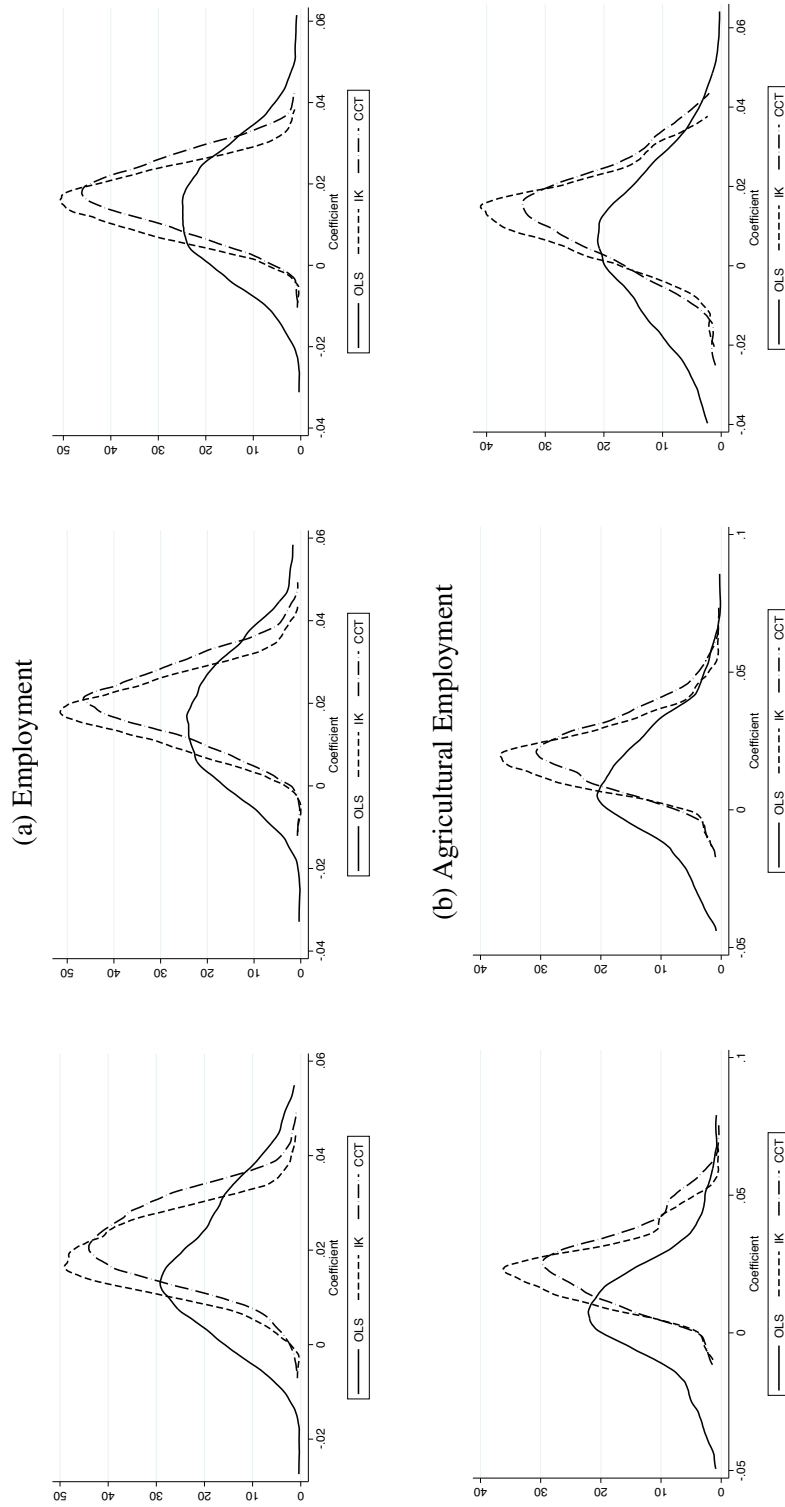
Notes. The figures plot the relationship between individual employment probabilities as derived from the DHS and electoral outcomes. The figures provide a graphical representation of the bias-corrected (Cattaneo, and Titiunik 2014) local linear regression fit on both sides of the threshold that determines whether the party linked to the ethnic group the individual belongs to gains a local representative in the national assembly. They also plot a scatterplot showing employment probabilities within 30 bins right and left of the threshold. The first graph shows unconditional probabilities, the second shows average residual probabilities net of country \times year and ethnicity fixed effects. The third also nets out constituency fixed effects.

Figure A.7: Effect on Employment – Alternative Bandwidth



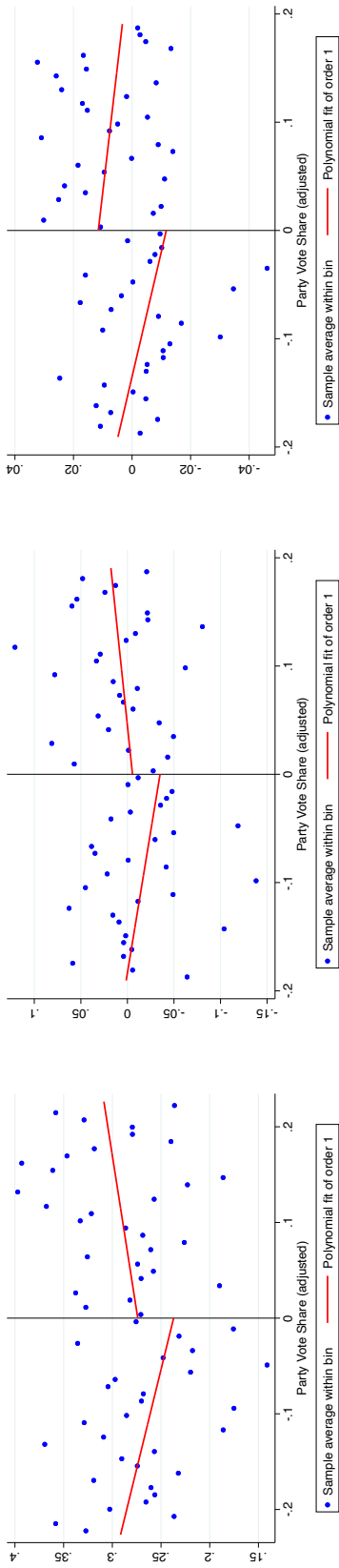
Notes. The figure reports the estimated coefficients and 90% confidence intervals that we obtain when implementing the specification in column 5 (left graph) and column 6 (right graph) of Table 2 using the estimator proposed by [Calonico, Cattaneo, and Titiunik \(2014\)](#) and across a wide range of alternative estimation bandwidths. Point estimates are remarkably stable irrespectively of bandwidth choice, although insignificant when considering the smallest bandwidths due to low statistical power.

Figure A.8: Effect Within Neighboring Constituencies



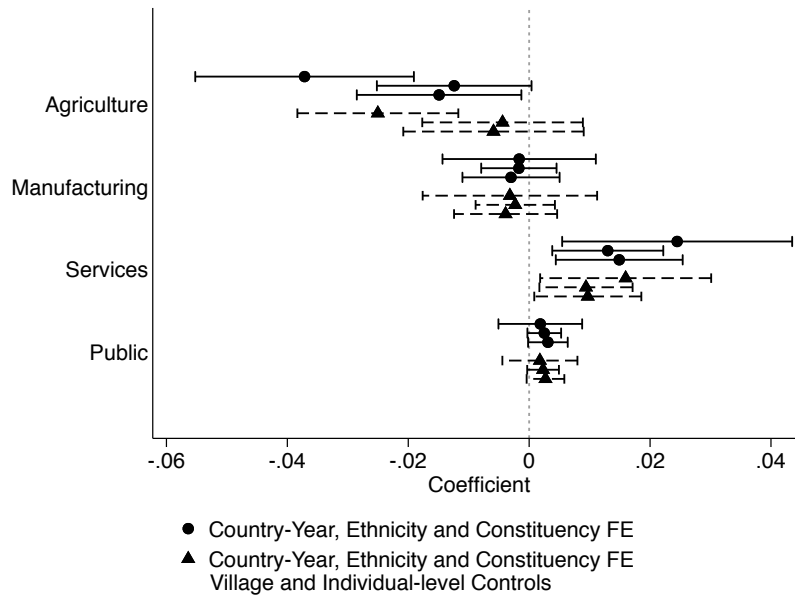
Notes. The figures plot the empirical distribution of coefficients obtained by implementing the regression specification in equation 2 with the addition of a full set of fixed effects that identify neighboring constituencies. Distributions are obtained following a bootstrap-type procedure that identifies pairs of neighboring constituencies at random in each of 200 repetitions. The left figure shows the empirical distributions obtained when implementing a specification that also includes country \times year and ethnicity fixed effects. The central figure shows the ones obtained from implementing a specification that also includes village-level controls, and the right figure the ones obtained from implementing a specification that also includes individual-level controls.

Figure A.9: Effect on Agricultural Employment



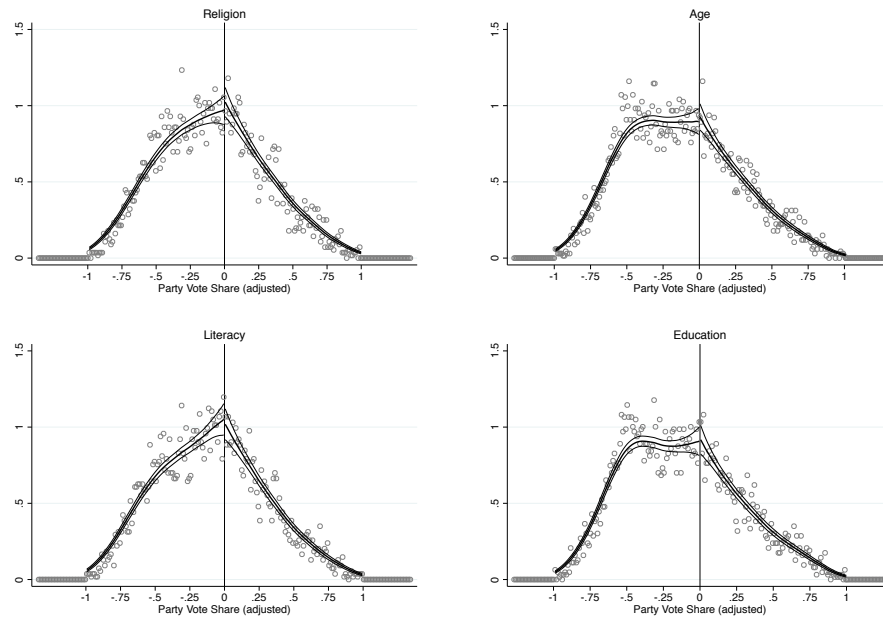
Notes. The figures plot the relationship between individual probabilities of employment in the agricultural sector as derived from the DHS and electoral outcomes. The figures provide a graphical representation of the bias-corrected (Calónico, Cattaneo, and Titiunik 2014) local linear regression fit on both sides of the threshold that determines whether the party linked to the ethnic group the individual belongs to gains a local representative in the national assembly. They also plot a scatterplot showing employment probabilities within 30 bins right and left of the threshold. The first graph shows unconditional probabilities, the second shows average residual probabilities net of country \times year and ethnicity fixed effects. The third also nets out constituency fixed effects.

Figure A.10: Disfavoritism - Effect on Employment by Sector



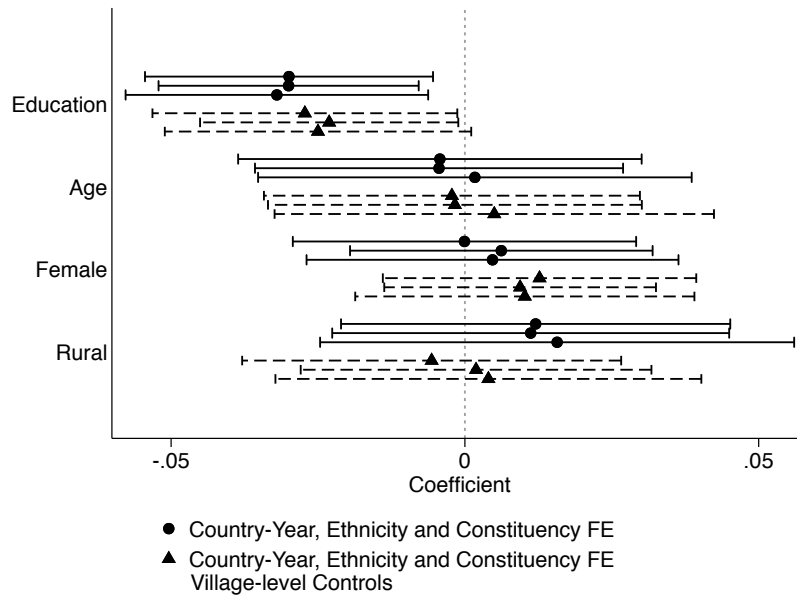
Notes. The figure reports the estimated coefficients and 90% confidence intervals that we obtain when estimating equation 2, but considering the vote share of the party that is (the most) under-represented, relative to its overall vote share in the country, among voters from ethnic group e , and using a cutoff of 5%. The dependent variable is a dummy equal to one if the individual reports to be working in agriculture, manufacturing, services, and the public sector. Online Appendix Table A.14 shows the corresponding results. For each sector, the figure reports two sets of three coefficient estimates: $\hat{\beta}_{OLS}$, obtained using OLS; $\hat{\beta}_{IK}$, obtained using the estimator discussed by [Imbens and Kalyanaraman \(2012\)](#); and $\hat{\beta}_{CCT}$, obtained using the estimator proposed by [Calonico, Cattaneo, and Titiunik \(2014\)](#). The estimation bandwidth is the one obtained using the selector proposed by [Calonico, Cattaneo, and Titiunik \(2014\)](#). The first set of three estimates belongs to a specification that includes country \times year, ethnicity, and constituency fixed effects. The second set of three estimates belongs to a specification that also includes village-level and individual-level controls.

Figure A.11: Density Plots – McCrary Test for Alternative Party Linkages



Notes. The figures show the distribution of the running variable and its density function as estimated separately on both sides of the threshold, together with 95% confidence intervals. The running variable is defined considering alternative party linkages. For religion-based party linkages, the p-value from a test of equality of the value of the density function on the left and right side of the threshold is equal to 0.4163 using the test by [McCrary \(2008\)](#), and equal to 0.6655 using the test by [Cattaneo, Jansson, and Ma \(2020\)](#). For age-based party linkages, the corresponding p-values are 0.3079 and 0.7136. For literacy-based party linkages, the corresponding p-values are 0.8643 and 0.9600. For education-based party linkages, the corresponding p-values are 0.8145 and 0.8991.

Figure A.12: Characteristics of Agricultural Workers



Notes. The figure reports the estimated coefficients and 90% confidence intervals that we obtain when restricting the sample to individuals reporting to be working in the agricultural sector and estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to have some primary education, individual's age (in standard deviation), a dummy for female, and a dummy for rural. Online Appendix Table A.18 shows the corresponding results. For each sector, the figure reports two sets of three coefficient estimates: $\hat{\beta}_{OLS}$, obtained using OLS; $\hat{\beta}_{IK}$, obtained using the estimator discussed by Imbens and Kalyanaraman (2012); and $\hat{\beta}_{CCT}$, obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). The first set of three estimates belongs to a specification that includes country \times year, ethnicity, and constituency fixed effects. The second set of three estimates belongs to a specification that also includes village-level controls.

"The Employment Effects of Ethnic Politics"

by Francesco Amodio, Giorgio Chiovelli and Sebastian Hohmann

Supplementary Appendix

B.1 Results with Alternative Ethnic Party Labeling Cutoffs

Table B.1: Effect on Employment – Over-representation Cutoff of 5%

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0205 (0.0229)	0.0332 (0.0145)	0.0334 (0.0131)	0.0240 (0.0107)	0.0267 (0.011)	0.0138 (0.0124)
$\hat{\beta}_{IK}$	0.0233 (0.0385)	0.0336 (0.0145)	0.0271 (0.0104)	0.0260 (0.0093)	0.0228 (0.0095)	0.0141 (0.0072)
$\hat{\beta}_{CCT}$	0.0264 (0.0427)	0.0366 (0.0166)	0.0297 (0.0121)	0.0300 (0.0107)	0.0254 (0.0111)	0.0156 (0.0086)
Bandwidth	0.273	0.232	0.242	0.289	0.244	0.223
Country \times Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	163027	146179	149781	164775	143863	133530

Notes. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the over-representation method discussed in Section 2 and using a cutoff of 5%. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table B.2: Effect on Employment – Estimates in Placebo Sample – Over-representation Cutoff of 5%

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: Effect of Last Elections</i>						
$\hat{\beta}_{OLS}$	0.0488 (0.0409)	0.0447 (0.025)	0.0262 (0.0208)	0.0237 (0.02)	0.0120 (0.0204)	0.0115 (0.0118)
$\hat{\beta}_{IK}$	0.0467 (0.0577)	0.0543 (0.0324)	0.0333 (0.0206)	0.0343 (0.0262)	0.0265 (0.0241)	0.0078 (0.0092)
$\hat{\beta}_{CCT}$	0.0419 (0.0641)	0.0560 (0.0382)	0.0352 (0.0240)	0.0329 (0.0305)	0.0264 (0.0278)	0.0091 (0.0110)
Bandwidth	0.207	0.244	0.268	0.203	0.191	0.193
Observations	34128	39672	41372	31917	30501	30815
<i>Panel B: Placebo Effect of Next Elections</i>						
$\hat{\beta}_{OLS}$	0.0180 (0.0298)	0.0105 (0.0207)	0.0016 (0.0223)	-0.0077 (0.0265)	-0.0124 (0.0249)	-0.0102 (0.0117)
$\hat{\beta}_{IK}$	0.0087 (0.0554)	0.0020 (0.0260)	0.0011 (0.0215)	0.0049 (0.0260)	0.0012 (0.0244)	0.0027 (0.0198)
$\hat{\beta}_{CCT}$	0.0096 (0.0604)	-0.0008 (0.0300)	-0.0012 (0.0251)	0.0104 (0.0304)	0.0068 (0.0290)	0.0074 (0.0234)
Bandwidth	0.239	0.257	0.266	0.213	0.203	0.196
Observations	37548	39352	40338	32435	31094	30033
Country \times Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes

Notes. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the over-representation method discussed in Section 2 and using a cutoff of 5%. Sample is restricted to those countries and political parties for which we have data on more than one election round. Panel A reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. Panel B reports the estimate we obtain when considering the placebo adjusted vote share, equal to the one obtained by the same party in the same constituency in the next election in our dataset. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table B.3: Effect on Employment by Sector – Over-representation Cutoff of 5%

	Agriculture		Manufacturing		Services		Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	0.0226 (0.0155)	0.0210 (0.0163)	0.0072 (0.008)	0.0062 (0.0083)	-0.0097 (0.0077)	-0.0121 (0.0063)	-0.0002 (0.0031)	-0.0017 (0.0028)
$\hat{\beta}_{IK}$	0.0179 (0.0092)	0.0143 (0.0085)	0.0040 (0.0040)	0.0041 (0.0041)	-0.0041 (0.0048)	-0.0030 (0.0038)	-0.0004 (0.0017)	-0.0006 (0.0015)
$\hat{\beta}_{CCT}$	0.0200 (0.0106)	0.0163 (0.0100)	0.0052 (0.0050)	0.0054 (0.0051)	-0.0048 (0.0056)	-0.0025 (0.0046)	-0.0008 (0.0021)	-0.0011 (0.0018)
Bandwidth	0.198	0.205	0.234	0.236	0.277	0.298	0.236	0.234
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	125062	123592	147188	141132	164867	168159	147563	140636

Notes. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the over-representation method discussed in Section 2 and using a cutoff of 5%. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working in agriculture (columns 1 and 2), manufacturing (columns 3 and 4), services (columns 5 and 6), and the public sector (columns 7 and 8). $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calónico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calónico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table B.4: Effect on Employment and Affiliation with Government – Over-representation Cutoff of 5%

	Employment Dummy							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>Non-Government Party</i>				<i>Government Party</i>			
$\hat{\beta}_{OLS}$	0.0337 (0.0236)	0.0231 (0.0161)	0.0183 (0.0171)	-0.0057 (0.0367)	0.0397 (0.0198)	0.0440 (0.0205)	0.0431 (0.0198)	0.1110 (0.054)
$\hat{\beta}_{TK}$	0.0230 (0.0156)	0.0200 (0.0130)	0.0128 (0.0129)	0.0021 (0.0081)	0.0472 (0.0224)	0.0595 (0.0248)	0.0535 (0.0214)	0.0295 (0.0116)
$\hat{\beta}_{CCT}$	0.0245 (0.0177)	0.0205 (0.0150)	0.0134 (0.0144)	0.0012 (0.0099)	0.0506 (0.0267)	0.0655 (0.0283)	0.0587 (0.0243)	0.0323 (0.0132)
Difference (p-value)					[0.1366]	[0.0407]	[0.0678]	[0.0693]
Bandwidth	0.258	0.310	0.262	0.255	0.221	0.173	0.178	0.145
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Individual Controls	No	No	Yes	Yes	No	No	Yes	Yes
Constituency FE	No	No	No	Yes	No	No	No	Yes
Observations	76068	84018	74960	73877	35305	28294	28706	25114

Notes. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the over-representation method discussed in Section 2 and using a cutoff of 5%. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{TK}$ is the conventional RD estimate (Imbens and Kalyanaram 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Columns 1 to 4 are estimated using the subsample of parties that do not support the central government in the aftermath of elections. Columns 5 to 8 report the same estimates using the subsample of parties that do not support the central government in the aftermath of elections. They also report in square brackets the p-value from a test of equality of $\hat{\beta}_{CCT}$ from the same specification across the two subsamples. Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table B.5: Effect on Employment – Polity IV Democracies – Over-representation Cutoff of 5%

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0237 (0.0228)	0.0280 (0.0167)	0.0252 (0.0158)	0.0281 (0.0165)	0.0218 (0.0152)	0.0147 (0.0188)
$\hat{\beta}_{IK}$	0.0210 (0.0406)	0.0334 (0.0181)	0.0249 (0.0132)	0.0261 (0.0139)	0.0239 (0.0129)	0.0162 (0.0098)
$\hat{\beta}_{CCT}$	0.0232 (0.0454)	0.0362 (0.0206)	0.0267 (0.0161)	0.0282 (0.0173)	0.0259 (0.0158)	0.0171 (0.0116)
Bandwidth	0.291	0.225	0.237	0.227	0.234	0.261
Country \times Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	117968	95629	101059	91655	94824	101934

Notes. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the over-representation method discussed in Section 2 and using a cutoff of 5%. Sample is restricted to those countries and election years with Polity IV score equal to 6 or higher. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table B.6: Effect on Agricultural Employment – Over-representation Cutoff of 5%

	Agricultural Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0109 (0.0199)	0.0322 (0.0173)	0.0221 (0.0167)	0.0165 (0.0141)	0.0069 (0.0141)	0.0210 (0.0163)
$\hat{\beta}_{IK}$	0.0236 (0.0389)	0.0332 (0.0269)	0.0240 (0.0184)	0.0218 (0.0189)	0.0119 (0.0160)	0.0143 (0.0085)
$\hat{\beta}_{CCT}$	0.0274 (0.0438)	0.0371 (0.0322)	0.0271 (0.0232)	0.0256 (0.0250)	0.0157 (0.0202)	0.0163 (0.0100)
Bandwidth	0.269	0.221	0.204	0.240	0.219	0.205
Country \times Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	160786	139536	129135	142739	131860	123592

Notes. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the over-representation method discussed in Section 2 and using a cutoff of 5%. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working in agriculture. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table B.7: Agricultural Employment Type, Land Ownership, and In-kind Pay – Over-representation Cutoff of 5%

	Agriculture Self-employed (1)	(2)	Agriculture for Someone Else (3)	(4)	Owns Land (5)	(6)	In-kind Pay (7)	(8)
$\hat{\beta}_{OLS}$	0.0155 (0.0184)	0.0203 (0.0199)	0.0071 (0.0055)	0.0070 (0.0064)	0.0049 (0.0096)	-0.0106 (0.0088)	0.0026 (0.0037)	0.0043 (0.0034)
$\hat{\beta}_{TK}$	0.0122 (0.0096)	0.0081 (0.0093)	0.0062 (0.0039)	0.0057 (0.0041)	-0.0001 (0.0087)	-0.0074 (0.0077)	0.0029 (0.0023)	0.0029 (0.0025)
$\hat{\beta}_{CCT}$	0.0131 (0.0113)	0.0089 (0.0108)	0.0071 (0.0045)	0.0067 (0.0048)	-0.0005 (0.0105)	-0.0087 (0.0090)	0.0033 (0.0029)	0.0033 (0.0032)
Bandwidth	0.196	0.187	0.167	0.170	0.197	0.205	0.257	0.238
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	123880	115035	109174	105718	101138	102646	155974	141959

Notes. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the over-representation method discussed in Section 2 and using a cutoff of 5%. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be self-employed in agriculture (columns 1 and 2), employed in the agriculture working for someone else (columns 3 and 4), own land (columns 5 and 6), and receive in-kind pay (columns 7 and 8). $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{TK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calónico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calónico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table B.8: Disfavoritism – Effect on Employment by Sector – Underrepresentation Cutoff of 0.01%

	Agriculture		Manufacturing		Services		Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	-0.0080 (0.0085)	-0.0057 (0.0074)	0.0028 (0.0077)	0.0073 (0.0077)	0.0130 (0.0142)	0.0073 (0.0091)	0.0003 (0.0036)	0.0026 (0.0031)
$\hat{\beta}_{TK}$	-0.0033 (0.0064)	0.0007 (0.0062)	-0.0030 (0.0038)	-0.0032 (0.0041)	0.0106 (0.0049)	0.0056 (0.0042)	0.0021 (0.0014)	0.0014 (0.0012)
$\hat{\beta}_{CCT}$	-0.0025 (0.0071)	0.0015 (0.0071)	-0.0049 (0.0046)	-0.0053 (0.0048)	0.0122 (0.0058)	0.0060 (0.0051)	0.0025 (0.0017)	0.0016 (0.0014)
Bandwidth	0.300	0.245	0.229	0.213	0.217	0.260	0.254	0.304
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	148299	115865	116715	103877	111169	123071	127242	141062

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of β and standard errors that we obtain when estimating equation 2, but considering the vote share of the party that is (the most) under-represented, relative to its overall vote share in the country, among voters from ethnic group e , and using a cutoff of 0.01%. The dependent variable is a dummy equal to one if the individual reports to be working in agriculture (columns 1 and 2), manufacturing (columns 3 and 4), services (columns 5 and 6), and the public sector (columns 7 and 8). $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{TK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table B.9: Disfavoritism – Effect on Employment – Underrepresentation Cutoff of 0.01%

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0288 (0.0318)	0.0423 (0.0251)	0.0139 (0.019)	0.0084 (0.0201)	0.0165 (0.0192)	-0.0005 (0.0077)
$\hat{\beta}_{IK}$	0.0281 (0.0539)	0.0306 (0.0258)	0.0084 (0.0155)	0.0091 (0.0152)	0.0122 (0.0139)	0.0030 (0.0056)
$\hat{\beta}_{CCT}$	0.0319 (0.0580)	0.0307 (0.0321)	0.0020 (0.0177)	0.0038 (0.0174)	0.0063 (0.0161)	0.0027 (0.0060)
Bandwidth	0.300	0.263	0.261	0.245	0.262	0.266
Country \times Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	148321	131023	130500	115745	123836	126417

Notes. The unit of observation is an individual as surveyed in the DHS. The table reports the estimates of β and standard errors that we obtain when estimating equation 2, but considering the vote share of the party that is (the most) under-represented, relative to its overall vote share in the country, among voters from ethnic group e , and using a cutoff of 0.01%. The dependent variable is a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table B.10: Effect on Employment for Alternative Party Linkages – Over-representation Cutoff of 5%

	Religion		Age		Literacy		Education	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	0.0461 (0.0257)	0.0477 (0.0192)	0.2051 (0.046)	0.0137 (0.0229)	0.0144 (0.0272)	0.0227 (0.0278)	0.0009 (0.0193)	0.0012 (0.0145)
$\hat{\beta}_{TK}$	0.0091 (0.0080)	0.0077 (0.0063)	0.1334 (0.0258)	0.0173 (0.0159)	0.0147 (0.0146)	0.0146 (0.0136)	0.0033 (0.0124)	0.0097 (0.0111)
$\hat{\beta}_{CCT}$	0.0102 (0.0099)	0.0086 (0.0078)	0.1338 (0.0303)	0.0142 (0.0200)	0.0155 (0.0174)	0.0168 (0.0166)	0.0003 (0.0155)	0.0114 (0.0139)
Bandwidth	0.343	0.380	0.238	0.284	0.286	0.340	0.262	0.325
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	64389	67694	19378	20262	21687	26216	34699	41840

Notes. The unit of observation is an individual as surveyed in the DHS. Party linkages are derived using the over-representation method discussed in Section 2 using a cutoff of 5% and for groups defined by religion (column 1 and 2), age (column 3 and 4), literacy (column 5 and 6), and education (column 7 and 8). The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{TK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calónico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calónico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

Table B.11: Effect on Employment and Decentralization – Over-representation Cutoff of 5%

	Employment Dummy							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>Not Decentralized</i>				<i>Decentralized</i>			
$\hat{\beta}_{OLS}$	0.0353 (0.0317)	0.0195 (0.019)	0.0148 (0.0195)	0.0007 (0.0089)	-0.0025 (0.0717)	0.0116 (0.0716)	0.0839 (0.0526)	0.0702 (0.0459)
$\hat{\beta}_{TK}$	0.0184 (0.0192)	0.0228 (0.0163)	0.0165 (0.0140)	0.0065 (0.0073)	0.0353 (0.0651)	0.0927 (0.0374)	0.0847 (0.0395)	0.0395 (0.0117)
$\hat{\beta}_{CCT}$	0.0211 (0.0198)	0.0258 (0.0178)	0.0199 (0.0149)	0.0095 (0.0087)	0.0470 (0.0775)	0.1007 (0.0459)	0.0857 (0.0564)	0.0466 (0.0145)
Difference (p-value)					[0.9011]	[0.4662]	[0.2022]	[0.0955]
Bandwidth	0.231	0.286	0.207	0.208	0.180	0.133	0.142	0.270
Country \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Individual Controls	No	No	Yes	Yes	No	No	Yes	Yes
Constituency FE	No	No	No	Yes	No	No	No	Yes
Observations	42188	47798	38060	38282	5745	3603	4373	8293

Notes. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the over-representation method discussed in Section 2 and using a cutoff of 5%. The table reports the estimates of β and standard errors that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{TK}$ is the conventional RD estimate (Imbens and Kalyanaraman 2012), and $\hat{\beta}_{CCT}$ is the bias-corrected RD estimate (Calonico, Cattaneo, and Titiunik 2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Columns 1 to 4 are estimated using the subsample of countries where local governments do not have jurisdiction over taxing, spending or legislation. Columns 5 to 8 report the same estimates using the subsample of countries where local governments have jurisdiction over taxing, spending or legislation. They also report in square brackets the p-value from a test of equality of $\hat{\beta}_{CCT}$ from the same specification across the two subsamples. Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the ethnic group level.

B.2 Elections and Data Sources by Country

Country	Data Description
Benin	<p><i>Electoral System:</i> Candidates are elected proportionally to the unicameral National Assembly (<i>Assemblée nationale</i>) on party lists. Each constituency corresponds to one of the country's departments. A party list has as many candidates as there are seats to be filled, with seats being allocated using a departmental (simple) quotient. Remaining seats are then filled using the greatest remainder rule.</p> <p><i>Election Years:</i> 1999, 2011.</p> <p><i>Election Data Sources:</i> CLEA.</p>
Burkina Faso	<p><i>Electoral System:</i> Candidates are elected proportionally to the National Assembly (<i>Assemblée nationale</i>) on party lists within multi-member constituencies. Seats are allotted using a simple electoral quotient and highest remainders.</p> <p><i>Election Years:</i> 2002, 2007, 2012.</p> <p><i>Election Data Sources:</i> 2002 - Carr; 2007 - CENI (http://www.ceni.bf/?q=resultats-provinciaux-legislatives-2007.html); 2012 - CLEA.</p>
Cameroon	<p><i>Electoral System:</i> Candidates are elected to the unicameral National Assembly (<i>Assemblée nationale</i>, the lower house of parliament) in a mixed system. Single-member districts use a plurality system. The remaining districts use a list proportional system. In the proportional system, a list which gains an absolute majority wins. If no list gains an absolute majority, the one with the highest number of votes gets half of the seats and the remainder is allocated to the other best-placed lists. There is a 5% threshold for representation in these districts, and seats are awarded to candidates in the order they appear in the lists.</p> <p><i>Election Years:</i> 2002, 2007.</p> <p><i>Election Data Sources:</i> CLEA.</p>
Ghana	<p><i>Electoral System:</i> Candidates are elected to the unicameral Parliament using a simple plurality rule in single-member constituencies.</p> <p><i>Election Years:</i> 1996, 2000, 2004, 2012.</p> <p><i>Election Data Sources:</i> 1996 - Election Passport; 2000, 2004, 2012 - CLEA.</p>
Ivory Coast	<p><i>Electoral System:</i> Candidates are elected to the National Assembly (<i>Assemblée nationale</i>, the lower house of the legislature) using a plurality rule in single- and multi-member constituencies. In the multi-member constituencies, this amounts to block voting.</p> <p><i>Election Years:</i> 2011.</p> <p><i>Election Data Sources:</i> CLEA.</p>

Kenya	<p><i>Electoral System:</i> Candidates are elected to the National Assembly, the lower chamber of the Kenyan legislature, in a plurality system.⁴⁰ Some seats in the Assembly are elected from single member districts while others are seats reserved for women, who are elected at the county level. Other members are nominated and the position of the speaker is <i>ex officio</i>.</p> <p><i>Election Years:</i> 2002, 2007, 2013.</p> <p><i>Election Data Sources:</i> 2002 - EU Election Observation Mission Final Report; 2007 - Carr; 2013 - CLEA.</p>
Liberia	<p><i>Electoral System:</i> Candidates are elected using a plurality rule to the House of Representatives, the lower house of the Liberian legislature, in single-member constituencies.</p> <p><i>Election Years:</i> 2005, 2011.</p> <p><i>Election Data Sources:</i> CLEA.</p>
Malawi	<p><i>Electoral System:</i> Candidates are elected to the unicameral National Assembly using a plurality rule in single-member constituencies.</p> <p><i>Election Years:</i> 1999, 2004, 2009, 2014.</p> <p><i>Election Data Sources:</i> CLEA.</p>
Mali	<p><i>Electoral System:</i> Candidates are elected to the unicameral National Assembly (<i>Assemblée nationale</i>) in two rounds within single- and multi-member constituencies. In the first round, an absolute majority is required to win. The two best-placed candidates or lists in the first round advance to the second round, where only a simple majority (plurality) is required.</p> <p><i>Election Years:</i> 2013.</p> <p><i>Election Data Sources:</i> Carr.</p>
Mozambique	<p><i>Electoral System:</i> Candidates are elected proportionally on party lists to the unicameral Assembly of the Republic (<i>Assembleia da Republica</i>) in single and multi-member constituencies. The d'Hondt method is used to allocate seats in the multi-member districts. There is a 5% electoral threshold to be represented in the Assembly.</p> <p><i>Election Years:</i> 2009.</p> <p><i>Election Data Sources:</i> CLEA.</p>
Nigeria	<p><i>Electoral System:</i> Candidates are elected using a plurality rule to the lower house of Nigeria's National Assembly, the House of Representatives, in single-member constituencies.</p> <p><i>Election Years:</i> 2007, 2011.</p>

⁴⁰From 1966 to 2010 the Kenyan legislature was unicameral, lacking a Senate.

Election Data Sources: 2007 - Carr; 2011 - CLEA.

Senegal *Electoral System:* Candidates are elected in a mixed system to the unicameral National Assembly (*Assemblée nationale*) in single- and multi-member constituencies. Block voting is used in a portion of the districts, while in the remaining constituencies candidates are chosen proportionally at the national level from lists of candidates provided by parties or coalitions.

Election Years: 2007, 2012.

Election Data Sources: CLEA.

Sierra Leone *Electoral System:* Candidates are elected using a plurality rule to the unicameral Parliament in single-member constituencies. Several other members are indirectly elected.

Election Years: 2007, 2012.

Election Data Sources: CLEA.

Uganda *Electoral System:* Most candidates are elected using a plurality rule to the unicameral Parliament in single-member constituencies. Further seats are reserved for women, who are elected directly, and appointees.

Election Years: 2011.

Election Data Sources: CLEA.

Zambia *Electoral System:* Most candidates are elected using a plurality rule to the unicameral National Assembly in single-member constituencies. Other members are appointed by the President. The Vice President is also considered a member of the body.

Election Years: 2006, 2011.

Election Data Sources: CLEA.

B.3 Labeling Algorithm: Design and Implementation

Political parties may function as the representatives of special interests, such as ethnic groups. Which interest groups are represented by which party is typically well understood by participants of any given political system, but assigning a “group label” to a party still involves a certain amount of subjectivity. To overcome this challenge and tie our hands, we use survey-data that has information on individuals’ ethnic identification as well as the parties individuals voted for (both self-reported) to “learn” the ethnic affiliation of parties from data. The data we use come from Afrobarometer, waves 1-6. Table B.13 gives an example.

Table B.13: Example of ethnicity-voting data from Afrobarometer

country	Afrobarometer round	ethnic group	party vote
Ghana	5	Akan	NPP
Ghana	5	Akan	NDC
Ghana	5	Ewe	NDC

These data can be used in two ways.

- (a) We can try to predict which ethnic groups are disproportionately likely to be among the voters for some party to assign ethnicity-labels to parties
- (b) We can try to predict which party is disproportionately likely to be the one chosen by the voters of a particular ethnic group to assign party-labels to ethnicities.

This appendix describes different methods for both of these tasks and compares them in practise using our data.

Harmonizing ethnicity names Reported ethnicities in Afrobarometer sometimes fall into many (50+ per country) individual groups. Many of these groups are tiny (3-5 individuals only in the data) and are subgroups of other, larger groups. If groups are very small, predicted labels may be very sensitive to outliers, e.g. if all individuals sampled from that group happen to vote for one party, which, given the clustered design of Afrobarometer, is more likely to happen. Moreover, the DHS ethnicity data is less granular than the Afrobarometer data. Since we observe our outcome of interest (employment) for the DHS data, we can only use the level of detail on ethnic groups available in the DHS data. As a first step, we therefore harmonize groups appearing in Afrobarometer to their DHS equivalent and assign smaller subgroups to their larger ethnic kin. We create ethnicity link files for each country to automate the matching. We do not change the level of ethnic detail on the DHS side but we do harmonize spellings of groups where they differ between DHS survey rounds. For Afrobarometer, we use an individual’s language (in the variables “What is your home language” (round 1) and “Language of respondent” (round 2), where explicit ethnicity information is not available. For rounds 3-6, we have the variable “Tribe or ethnic group”, and only use the language variable where this variable is missing.

Harmonizing party names Afrobarometer records the affiliation to political parties in great detail. In rounds 1-2, respondents are asked “Do you feel close to any particular party?” “If yes, which party?” In rounds 3-6, they are additionally asked “If the election were held tomorrow, which party would you vote for?” We use the direct voting question where available and use the closeness to party only where the answer to the direct voting question is not available. We then match all answers given in Afrobarometer to the names of parties in our voting data. There are several parties appearing in Afrobarometer that do not appear in our voting data. In running the algorithms, we therefore constrain the set of parties to be that in our election data.

Missing data We exclude from the dataset used for labeling those individuals for which either the ethnicity or party information is missing.

B.3.1 The algorithms

(a) Assigning ethnicity labels to parties

Vector-distance over-representation The first algorithm assigns ethnicity labels to parties by checking which ethnicities are over-represented relative to their national population share among the voters of each party. First, define the following notation:

- N number of individuals in the country
- N^e number of individuals of ethnic group e
- N_p number of votes cast for party p
- N_p^e number of votes cast by individuals of ethnic group e for party p

It carries out the following steps

- 1) Compute ethnicity-party vote share $s_p^e = \frac{N_p^e}{N_p}$, party vote share $s_p = \frac{N_p}{N}$, and ethnicity population shares $s^e = \frac{N^e}{N}$
- 2) Compute ethnicity-party-over-representation measure $\beta_p^e = s_p^e - s^e$
- 3) Keep only party×ethnicity observations with $\beta_p^e > \bar{\beta}$ (this cutoff can be varied).
- 4) Sort the labels by β_p^e such that the ethnicity with the highest β_p^e is the first label of each party.

The end result is a set of labels for each party. As an alternative to simple over-representation, we also compute a measure of “percentage” over-representation, replacing β_p^e with $\gamma_p^e = \frac{s_p^e - s^e}{s^e}$. Intuitively, when an ethnicity is small, the simple difference $s_p^e - s^e$ can never be very large. Normalizing by group size puts all ethnicities on an equal footing, but makes the measure assign large “meaning” to very small groups that are over-represented among voters of certain parties.

Dummy regressions This algorithm predicts the ethnic affiliation of parties by

- 1) running a set of $k = 1, \dots, K$ regressions (one set of K regressions for every party p , where each individual regression includes only an ethnicity dummy for ethnic group k):

$$\mathbb{I}(\text{vote for party } p)_{is} = F(\alpha_s, E^k; \boldsymbol{\theta}), \quad (1)$$

where E^k is a dummy equal to one if individual i observed in survey(-year) s belongs to ethnicity k and zero otherwise, α_s is a survey fixed effect, $\boldsymbol{\theta}$ is a vector of coefficients, and $F(\cdot)$ is a function. The code implements this regression as either LPM or Probit.

- 2) Obtain the estimated coefficient $\hat{\theta}_p^k$ on E^k as well as its standard error, and compute the absolute value of the t -statistic as $t_p^k \equiv \frac{\hat{\theta}_p^k}{\widehat{se}(\hat{\theta}_p^k)}$.
- 3) Discard party \times ethnicities with $|\hat{\theta}_p^k| < t_{0.05, N-n_s-1}$, where $t_{0.05, N-n_s-1}$ is the critical value for a two-sided t -test from a distribution with $N - n_s - 1$ d.f. (number of individuals minus number of survey years minus 1; corresponds to significance at the 5% level).
- 4) Among the significant party \times ethnicities, retain only those with $\hat{\theta}_p^k > \bar{\theta}$ (the cutoff can be varied).
- 5) Sort the labels by $\hat{\theta}_p^k$ such that the ethnicity with the highest $\hat{\theta}_p^k$ is the first label of each party.

The end result is a set of labels for each party.

(b) Assigning party labels to ethnicities

Vector-distance over-representation This algorithm works similarly to the vector-distance over-representation algorithm assigning ethnicity labels to parties. It asks: “Are parties over-represented relative to their overall vote share among the voters from certain ethnic groups?” and carries out the following steps:

- 1) Compute party-ethnicity vote share $s_e^p = \frac{N_p^e}{N_e}$, and party vote share $s^p = \frac{N_p}{N}$
- 2) Compute party-ethnicity over-representation measure $\beta_e^p = s_e^p - s^p$
- 3) Keep only ethnicity \times party observations with $\beta_e^p > \bar{\beta}$ (this cutoff can be varied).
- 4) Sort the labels by β_e^p such that the party with the highest β_e^p is the first label of each ethnicity.

The end result is a set of party-labels for each ethnic group. As for assigning ethnicity-labels to parties, we also compute a version of the labels based on a normalized over-representation measure $\gamma_e^p = \frac{s_e^p - s^p}{s^p}$

Dummy regressions This algorithm predicts the party affiliation of ethnic groups by

- 1) running a set of regressions (one for every ethnicity e):

$$\mathbb{I}(\text{individual belongs to ethnicity } e)_i = F(\alpha_s, P^k; \boldsymbol{\theta}), \quad (2)$$

where P^k is a dummy equal to one if individual i observed in survey(-year) s voted for party k and zero otherwise, α_s is a survey fixed effect, $\boldsymbol{\theta}$ is a vector of coefficients, and $F()$ is a function. The code implements this regression as either LPM or Probit.

- 2) Obtain the estimated $\hat{\theta}_e^k$ on P^k as well as its standard error, and compute the absolute value of the t -statistic as $t_e^k \equiv \frac{\hat{\theta}_e^k}{\widehat{se}(\hat{\theta}_e^k)}$.
- 3) Discard ethnicity \times parties with $|\hat{\theta}_e^k| < t_{0.05, N-n_s-1}$, where $t_{0.05, N-n_s-1}$ is the critical value for a two-sided t -test from a distribution with $N - n_s - 1$ d.f. (number of individuals minus number of survey years minus 1; corresponds to significance at the 5% level).
- 4) Among the significant ethnicity \times parties, retain only those with $\hat{\theta}_e^k > \bar{\theta}$ (the cutoff can be varied).
- 5) Sort the labels by $\hat{\theta}_e^k$ such that the party with the highest $\hat{\theta}_e^k$ is the first label of each ethnicity.

The end result is a set of party-labels for each ethnic group.

B.3.2 Comparing the algorithms

Our baseline results are based on the simple over-representation measure, direction (b). There are two reasons for choosing (b). The first is practical. Ideally, we would like to obtain the party affiliation of each individual in the DHS data based on their ethnicity. In going in direction (a), one ethnic group may be “the most” over-represented group for several parties, while other groups are the most over-represented group for no party. For those latter groups, we would then not have a label. The second reason is more conceptual. For direction (a), an ethnic group with a 5 percent population share may represent 7 percent of the voters for a certain party. At the same time, a party may capture a greater share of the voters from that ethnic group.

B.3.3 Labeling Output - Over-representation with 0.01 Cutoff

Country	Ethnic Group	Linked Party
Benin	adja	parti social-democrate (psd)
	bariba	forces cauris pour un benin emergent
	betamaribe	front d'action pour le renouveau et le developpement (fard-alafia)
	dendi	forces cauris pour un benin emergent
	fon	parti de la renaissance du benin
	peulh	forces cauris pour un benin emergent
	yoa and lokpa	front d'action pour le renouveau et le developpement (fard-alafia)
	yoruba	forces cauris pour un benin emergent
	Burkina Faso	bissa
bobo		alliance pour la démocratie et la fédération / rassemblement démocratique africain
dagaaba		union pour le progrès et le changement
dioula		union pour le progrès et le changement
gurunsi		union pour le progrès et le changement

	lobi	union pour le progrès et le changement
	mossi	congrès pour la démocratie et le progrès
	peulh	parti pour la démocratie et le socialisme/parti des bâtisseurs
	touareg	congrès pour la démocratie et le progrès
Cameroon	adamaoua-oubangui	union nationale pour la démocratie et le progrès
	arab-choa/peulh/haoussa/kanuri	union nationale pour la démocratie et le progrès
	bamilike/bamoun	front social-démocratique
	bantoide south-west	front social-démocratique
	beti/bassa/mbam	rassemblement démocratique du peuple camerounais
	biu-mandara	union nationale pour la démocratie et le progrès
	cotier/ngoe/oroko	rassemblement démocratique du peuple camerounais
	grassfields	front social-démocratique
Ghana	akan	new patriotic party
	dagaaba	national democratic congress
	ewe	national democratic congress
	ga-dangme	national democratic congress
	grusi	national democratic congress
	guan	national democratic congress
	gurma	national democratic congress
	hausa	national democratic congress
	mande	national democratic congress
	mole-dagbani	national democratic congress
Ivory Coast	akan	parti démocratique de côte d'ivoire
	krou	parti démocratique de côte d'ivoire
	mande, north	rassemblement des républicains
	mande, south	union pour la démocratie et la paix en côte d'ivoire
	voltaic, south	rassemblement des républicains
Kenya	borana	kenya african national union
	gabbra	party of national unity
	iteso	orange democratic movement
	kalenjin	united republican party
	kamba	national rainbow coalition (narc)
	kikuyu	the national alliance
	kisii	forum for the restoration of democracy-people
	kuria	national rainbow coalition (narc)
	luhya	orange democratic movement
	luo	orange democratic movement
	maragoli	national rainbow coalition (narc)
	masai/samburu	the national alliance
	meru/embu	national rainbow coalition (narc)
	mijikenda/swahili	national rainbow coalition (narc)
	oromo	united republican party
	pokot	united republican party
	rendille	kenya african national union
	somali	orange democratic movement
	taita/taveta	orange democratic movement
	turkana	orange democratic movement
Liberia	bassa	liberty party
	belle	freedom alliance party of liberia
	dey	congress for democratic change
	gbandi	noational democratic coalition
	gio	national union for democratic progress
	gola	congress for democratic change
	grebo	congress for democratic change
	kissi	unity party
	kpelle	unity party
	krahn	congress for democratic change
	kru	congress for democratic change
	lorma	unity party
	mandingo	all liberian coalition party
	mano	national union for democratic progress
	mende	liberty party
	sarpo	liberty party
	vai	congress for democratic change
Malawi	amanganja/anyanja	democratic progressive party
	chewa	malawi congress party
	lambya	democratic progressive party
	lomwe	democratic progressive party
	ngoni	democratic progressive party

	nkhonde	people's party
	sena	democratic progressive party
	tonga	people's party
	tumbuka	people's party
	yao	united democratic front
Mali	bambara	congres national d'initiative democratique
	bobo	union pour democracie et developpement
	malinke	rally for mali
	peul/toucouleur	union pour la republique et la democratie
	sarakole/soninke/marka	party for national rebirth
	senoufou/miniaka	african solidarity for democracy and independence
	sonra	union pour la republique et la democratie
	tamachek/bella	alliance pour la democratie au mali
Mozambique	bitonga	liberation front of mozambique
	chitewe	liberation front of mozambique
	ciceva	mozambican national resistance
	cichopi	liberation front of mozambique
	cindau	mozambican national resistance
	cinyungwe	mozambican national resistance
	cisena	mozambican national resistance
	ciyao	mozambican national resistance
	coti	liberation front of mozambique
	echuwabo	democratic movement of mozambique
	elomwe	mozambican national resistance
	emakhuwa	liberation front of mozambique
	kimwane	mozambican national resistance
	shimakonde	liberation front of mozambique
	shona	liberation front of mozambique
	xichangana	liberation front of mozambique
	xirhonga	liberation front of mozambique
	xitswa	liberation front of mozambique
Nigeria	abua	people's democratic party
	adun	people's democratic party
	alago	all progressive congress (apc) / all progressives grand alliance
	angas	all nigeria people's party
	annang	people's democratic party
	bajju	people's democratic party
	bassa	people's democratic party
	berom	people's democratic party
	bini/edo	action congress of nigeria
	bura/babur	all nigeria people's party
	chamba	action congress of nigeria
	chip	people's democratic party
	djerma	people's democratic party
	ebira/igbira	all progressive congress (apc) / all progressives grand alliance
	efik	people's democratic party
	eggon	people's democratic party
	ekoi	people's democratic party
	ekpeye/afisa/epie	people's democratic party
	eleme	people's democratic party
	esan	action congress of nigeria
	etsako	people's democratic party
	fulani	all nigeria people's party
	gerawa	people's democratic party
	gwari	people's democratic party
	gwoza	action congress of nigeria
	hausa	all nigeria people's party
	ibibio	people's democratic party
	idoma	people's democratic party
	igala	people's democratic party
	igbo/ibo	all progressive congress (apc) / all progressives grand alliance
	igede	people's democratic party
	ijaw/izon	people's democratic party
	ika	people's democratic party
	ikwere	people's democratic party
	irigwe	people's democratic party
	ishan	people's democratic party
	isoko	people's democratic party
	itsekiri	all nigeria people's party

	iyala	people's democratic party
	jaba	people's democratic party
	jarawa	people's democratic party
	jukun	people's democratic party
	kadara	people's democratic party
	kagoro	people's democratic party
	kalabari	people's democratic party
	kambari	all nigeria people's party
	kanuri/beriberi	all nigeria people's party
	kataf/atyap	people's democratic party
	kibaku	people's democratic party
	kilba	all nigeria people's party
	koro	people's democratic party
	lantang	people's democratic party
	laru	all progressive congress (apc) / all progressives grand alliance
	mada	all nigeria people's party
	mandara	people's democratic party
	marghi/mangi	all nigeria people's party
	migili	people's democratic party
	mumuye	people's democratic party
	mwagavol	people's democratic party
	nupe	all progressive congress (apc) / all progressives grand alliance
	ogba	all nigeria people's party
	ogaja	all nigeria people's party
	ogoni	people's democratic party
	okirika	people's democratic party
	pyem	people's democratic party
	rukuba	people's democratic party
	sayawa/siyawa	people's democratic party
	shuwa	people's democratic party
	tangale	all progressive congress (apc) / all progressives grand alliance
	tarok	people's democratic party
	tiv	people's democratic party
	ugeb	people's democratic party
	ukwuani	people's democratic party
	urhobo	people's democratic party
	waja	all progressive congress (apc) / all progressives grand alliance
	yakurr	people's democratic party
	yoruba	action congress of nigeria
Senegal	diola	front pour le socialisme et la democratie/benno jubel
	mandingue/soce/malnk	parti démocratique sénégalais
	poular	benno bokk yakaar
	sarakole /soninke	benno bokk yakaar
	serer	benno bokk yakaar
	wolof/lebou	parti démocratique sénégalais
Sierra Leone	creole	sierra leone people's party (slpp)
	fullah	national democratic alliance (nda)
	kono	people's movement for democratic change (pmdc)
	koranko	all people's congress (apc)
	limba	all people's congress (apc)
	loko	all people's congress (apc)
	mandingo	sierra leone people's party (slpp)
	mende	sierra leone people's party (slpp)
	sherbro	sierra leone people's party (slpp)
	temne	all people's congress (apc)
Uganda	acholi	forum for democratic change
	akonjo/ konzo	forum for democratic change
	alur/jopadhola	forum for democratic change
	fumbira	national resistance movement
	ganda	democratic party
	gisu/sabiny	national resistance movement
	gwere	national resistance movement
	iteso	national resistance movement
	karimojong	national resistance movement
	kiga	national resistance movement
	kumam	forum for democratic change
	lango	ugandan people's congress
	lendu-lugwara-chope-madi	forum for democratic change
	nyankole	national resistance movement

	nyarwanda	national resistance movement
	nyoro	national resistance movement
	soga	national resistance movement
	toro	national resistance movement
Zambia	bemba	patriotic front
	bisa	patriotic front
	bwile	patriotic front
	chewa	movement for multiparty democracy
	chikunda	movement for multiparty democracy
	chokwe	united party for national development
	gowa	patriotic front
	ila	united party for national development
	kaonde	united party for national development
	kunda	patriotic front
	lala	patriotic front
	lamba	patriotic front
	lenje	united party for national development
	lozi	united party for national development
	luchazi	movement for multiparty democracy
	lunda (northwestern)	united party for national development
	lungu	patriotic front
	luvale	united party for national development
	mambwe	patriotic front
	mashi	united national independence party
	mbunda	united party for national development
	namwanga	patriotic front
	ngoni	movement for multiparty democracy
	ngumbo	patriotic front
	nkoya	united party for national development
	nsenga	movement for multiparty democracy
	nyanja	movement for multiparty democracy
	solli	united party for national development
	tabwa	patriotic front
	toka-leya	united party for national development
	tonga	united party for national development
	tumbuka	movement for multiparty democracy
	ushi	patriotic front

B.4 Verification of Ethnicity-Party Linkages

To further validate our matching algorithm, we gathered qualitative information from a number of sources. This sometimes consisted of finding existing academic literature which noted correspondences between a party and a certain ethnic group. More often, however, national and local newspapers provided suggestions as to which parties are supported by or linked to particular ethnicities. In some cases, such articles outright state that there is a connection, while in many other cases there is only a suggested relationship. In yet other cases, the match is based on a party leader hailing from a particular ethnic group. All reliable verifications are summarized in the table below. After the table, a key of party abbreviations by country and complete bibliography of sources is given.

Country	Party	Linkage Descriptions
Benin	FARD-A	<i>Linked Ethnicities:</i> Betamaribe <i>Sources:</i> Nicaise (2004)
	PSD	<i>Linked Ethnicities:</i> Adja <i>Sources:</i> Basedau and Stroh (2011), Somé (2009)
	RB	<i>Linked Ethnicities:</i> Fon <i>Sources:</i> "Soglo [of the RB] held sway in the central Abomey region, which is populated by fellow Fon." (Battle and Seely 2010, p. 50)
Burkina Faso	ADF-RDA	<i>Linked Ethnicities:</i> Bobo <i>Sources:</i> Party congress held in the homeland of the Bobo (Lefaso.net 2019)
	CDP	<i>Linked Ethnicities:</i> Mossi <i>Sources:</i> Basedau and Stroh (2011)
	UPC	<i>Linked Ethnicities:</i> Bissa <i>Sources:</i> Party leader is of Bissa descent (Voaafrique 2015)
Cameroon	CPDM	<i>Linked Ethnicities:</i> Beti <i>Sources:</i> Suggests that a prior CPDM leader favoured his co-ethnics in the Beti (Morse 2017)
	NUDP	<i>Linked Ethnicities:</i> Adamaoua, Peulh <i>Sources:</i> Notes that the region of the Adamaoua is a bastion of the NUDP (Awalson 2013) Notes that a leader of the NUDP is Peulh (Universalis.fr 1992)
	SDF	<i>Linked Ethnicities:</i> Grassfields, Bamileke <i>Sources:</i> "In the villages of the Cameroon grassfields... the dominant Social Democratic Front Party... has come to enjoy the status of a religion." (Chi 1998, p. 366) "Most of the [SDF's] approximately 60-per-cent Francophone membership belongs to the entrepreneurial Bamileke..." (Konings 2004, p. 372)

Côte d'Ivoire	RDR	<p><i>Linked Ethnicities:</i> Northern Mandé, Voltaic</p> <p><i>Sources:</i> "Significant positive correlations exist in the Northern Mandé and Voltaic votes for Ouattara [of the RDR]. (Bassett 2011, p. 476)</p>
	UPDCI	<p><i>Linked Ethnicities:</i> Southern Mandé</p> <p><i>Sources:</i> "the [UDPCI]... is primarily supported by followers from the general's [General Gueï, founder of the UDPCI] Yacouba ethnic group as well as other groupings of the Southern Mandés." (The World Bank 2008, p. 8)</p>
Ghana	NDC	<p><i>Linked Ethnicities:</i> Ga-Dangme, Guan, Mole-Dagbani</p> <p><i>Sources:</i> "... the fact that the Ga-Dangme constituencies have been industrious, faithful and loyal to the progress of the National Democratic Congress and Ghana as a whole since its inception." (AdomOnline.com 2019)</p> <p>"... these results suggest that districts with large percentages of Guan speakers lean heavily towards the NDC." (Fridy 2009, p. 12, footnote 24)</p> <p>"... northern groups, especially the Mole-Dagbani, aligning with the NDC." (Hoffman and Long 2013, p. 131)</p>
	NPP	<p><i>Linked Ethnicities:</i> Akan</p> <p><i>Sources:</i> "What I am clearly and unabashedly driving at here is that the key operatives of the Akan-dominated [NPP]..." (Okoampa-Ahoofe 2018)</p>
Kenya	NARC	<p><i>Linked Ethnicities:</i> Kamba</p> <p><i>Sources:</i> "The various leaders of NARC, and in particular... Musyoka [of the Kamba]... appealed to their communities to turn out in large numbers and to support NARC." (Elscher 2013, p. 81-82)</p>
	ODM	<p><i>Linked Ethnicities:</i> Luo</p> <p><i>Sources:</i> "... [Odinga's, who ran on the ODM ticket in 2007] strongest following is among his Luo community in the south-western region of Nyanza..." (BBC 2017)</p>
	URP	<p><i>Linked Ethnicities:</i> Kalenjin</p> <p><i>Sources:</i> "The Jubilee Party, formed in September 2016, followed a merger between The National Alliance and the [URP] representing... the Kalenjin." (Maina 2017)</p>
	TNA	<p><i>Linked Ethnicities:</i> Kikuyu</p> <p><i>Sources:</i> "... while the Kikuyus support President Uhuru Kenyatta [former member of the National Alliance]." (Nyambura 2017)</p>
Liberia	CDC	<p><i>Linked Ethnicities:</i> Grebo</p> <p><i>Sources:</i> "For this election, Mr Tubman switched to the [CDC] party... His mother came from the Grebo ethnic group..." (BBC 2011)</p>
	NUDP	<p><i>Linked Ethnicities:</i> Gio</p> <p><i>Sources:</i> "'The emphasis on ethnicity is declining,' Guannu said. 'But take the Johnson [of the National Union for Democratic Progress] case in Nimba... Where you see him receiving votes outside of Nimba, it is because his two groups, the Mano and Gio have sizable communities.'" (MacDougall 2011)</p>

	UP	<i>Linked Ethnicities:</i> Lorma, Kpelle <i>Sources:</i> "Opposing parties formed along tribal lines...for the Loma [sic], Kesselly's [UP]..." (Dennis and Dennis 2008, p. 111) "With the selection of UP's vice standard bearer Emmanuel Nuquay, the 'Kpelle factor' seems to be working for Nuquay here." (Dodoo 2017)
Malawi	DPP	<i>Linked Ethnicities:</i> Lomwe <i>Sources:</i> "The four front-runners hail from three ethnic groups: Mutharika [of the DPP] is Lomwe..." (Yi Dionne and Dulani 2014)
	MCP	<i>Linked Ethnicities:</i> Chewa <i>Sources:</i> "Similarly, in the central region, voters have shown steadfast support for the MCP (former President of Malawi Banda's old party). During his reign, Banda (a Chewa) favored the Chewa..." (Ferree and Horowitz 2007, p. 2)
	UDF	<i>Linked Ethnicities:</i> Yao <i>Sources:</i> "...the [MCP] seemed always stronger and promissory in dynamic sense than its counterparts namely the [UDF], which though based in the populous Southern region, proved inter-temporally unstable owing to their strong association with particular areas which had competing ethnic groups themselves (the Yao...)" (Matchaya 2010, p. 221)
Mali	RPM	<i>Linked Ethnicities:</i> Malinke <i>Sources:</i> Basedau and Stroh (2011)
Mozambique	RENAMO	<i>Linked Ethnicities:</i> Cindau <i>Sources:</i> "...a Cindau, just like Afonso Dhlakama, the president of Renamo..." (Cahen 2015, p. 4)
Nigeria	ACN	<i>Linked Ethnicities:</i> Yoruba <i>Sources:</i> "...the attempt by Gov. Obi and Dora Akunyili to paint his party, the [ACN] as a Yoruba party." (Nigeria Newsday 2011)
	ANPP	<i>Linked Ethnicities:</i> Hausa, Fulani <i>Sources:</i> "With ANPP considered as a party predominantly occupied by the Hausa/Fulani..." (Azeez 2009, p. 4)
	APC	<i>Linked Ethnicities:</i> Nupe, Tangale, Waja, Igbira <i>Sources:</i> "...Nupe women...have endorsed President Muhammadu Buhari [of the APC]..." (Sardauna 2019) "...the majority of the Tangale/Waja had decided to turn in their votes for the APC..." (ThisDay 2018) "...whose strategic influence in the election delivered a bloc vote of Igbira people to Fayemi and the newly elected Chairman of the APC..." (Udeh 2018)
	APGA	<i>Linked Ethnicities:</i> Ibo <i>Sources:</i> "Emeka Odumegwu-Ojukwu, who declared a separatist Republic of Biafra for his south-eastern ethnic Ibo people in 1967, will run under the colours of the newly-registered [APGA]." (BBC 2002)
	PDP	<i>Linked Ethnicities:</i> Jukun, Tarok, Igala, Igede, Esan/Ishan, Berom, Ikwere, Jarawa, Ijaw

		<p><i>Sources:</i> “The apex Jukun socio-political organisation Jukuns Development Association of Nigeria (JDAN) and other groups. . . have resolved to support the presidential candidate of the Peoples Democratic Party (PDP) Alhaji Atiku Abubakar. . . ” (Aderibigbe 2018)</p> <p>“The senator. . . Gen. Jeremiah Useni (rtd), has received the endorsement of his people, Tarok nation, to contest for the 2019 governorship election. . . on the platform of the Peoples Democratic Party (PDP).” (Adinoyi 2018)</p> <p>“The Igalala and Igede communities in Oyo state have expressed their support for the candidate of the [PDP]. . . ” (Alimi 2019)</p> <p>“Esan North East/Esan South East Fed Constituency is. . . regarded as the strong hold of PDP in Edo state. . . ” (Austin 2018)</p> <p>“. . . you will see that the Berom today are predominantly APC and no longer PDP as it was in 2015.” (Idegu 2018)</p> <p>“A former Acting National Chairman of the [PDP], Prince Uche Secondus, yesterday emerged as the National Chairman of the party. . . The Ikwere-born politician made his acceptance speech. . . ” (Nwachukwu 2017)</p> <p>“[Senator Mohammed’s of the PDP] background will give him an added advantage as the minority ethnic groups such as Jarawa. . . ” (Omokuvie 2019)</p> <p>“Pan Ijaw Socio-cultural organization, Ijaw National Congress, (INC) on Wednesday, endorsed the Presidential candidate of the PDP. . . ” (Vanguard 2019)</p>
Sierra Leone	APC	<p><i>Linked Ethnicities:</i> Temne, Limba, Koranko, Loko</p> <p><i>Sources:</i> “Most Temnes. . . viewed the APC as their own party. . . ” (Kandeh 1992, p. 91)</p> <p>“. . . the APC has been Temne or a loose coalition of northern ethnic groups (for example, Temne, Limba, . . . , Koranko. . .)” (Kandeh 1992, p. 92)</p> <p>“Siaka Stevens, a Northern Limba and trade unionist, formed an opposition party, the [APC] that appealed to Northern interests. His party had been joined by. . . Loko. . . ” (Refworld 2004)</p>
	NDA	<p><i>Linked Ethnicities:</i> Fullah</p> <p><i>Sources:</i> “The NDA leader, Mr. Mohamed C. Bah hailed from the Fullah tribe. . . ” (Bah 2015)</p>
	SLPP	<p><i>Linked Ethnicities:</i> Creole, Mende, Mandingo</p> <p><i>Sources:</i> “. . . the SLPP was formed by the merger of three different semi political organizations. . . the Sierra Leone Organization Society (SOS), the Protectorate Educational and Progressive Union (PEPU) and the People’s Party (PP) which was formed by Lamina Sankoh. . . the People’s Party. . . was an association of Creole intellectuals.” (Gorvie 2010)</p> <p>“. . . the Mende-dominated [SLPP].” (Kandeh 1992, p. 81)</p> <p>“The Mandingo have played an important role in the politics of Sierra Leone. They have traditionally supported the [SLPP]. . . ” (Taylor 2014, p. 128)</p>
Uganda	FDC	<p><i>Linked Ethnicities:</i> Acholi</p> <p><i>Sources:</i> “. . . the main opposition [FDC] performed well among Acholi. . . ” (Conroy-Krutz and Moehler 2014, p. 12)</p>
	NRM	<p><i>Linked Ethnicities:</i> Karamoja, Nyankole</p>

		<p><i>Sources:</i> "...the people of Karamoja who overwhelmingly voted for the ruling [NRM]." (Ariong 2016)</p> <p>"...save the Baganda-Nyankole alliance at the heart of the [NRM]..." (Roessler 2016)</p>
	UPC	<p><i>Linked Ethnicities:</i> Langi</p> <p><i>Sources:</i> "...while the UPC...remain[s] largely Langi..." (Conroy-Krutz and Moehler 2014, p. 12)</p>
Zambia	MMD	<p><i>Linked Ethnicities:</i> Ngoni</p> <p><i>Sources:</i> "Dr. Nevers Mumba, President of the opposition [MMD] and himself a Ngoni" (Hewitt and Phiri 2014, p. 4)</p>
	PF	<p><i>Linked Ethnicities:</i> Bemba, Lamba, Ngumbo, Bisa, Lala</p> <p><i>Sources:</i> "Support base [of Edgar Lungu of the PF]: Home area of Eastern province, the capital, Lusaka, and the Copperbelt plus the Bemba-speaking regions" (BBC 2016)</p> <p>"Senior Chief Kalilele of the Lamba speaking people...has praised President Edgar Lungu [of the PF] on the major projects government has embarked in his district." (Kombe 2018)</p> <p>"Chief Mulongwe of the Ngumbo speaking people said only under the leadership of President Edgar Chagwa Lungu [of the PF] has Zambia seen this kind of development being rolled out to all parts of the country..." (Mwebantu 2018)</p> <p>"He said Bembas should not be dragged into the Patriotic Front succession wrangles which he says are purely a Bisa internal struggle." (Zambian Eye 2013)</p> <p>"Chief Kabamba of the Lala Speaking people in Serenje District in Central Province has observed that the [PF] government has performed well in the first twenty two months of being in power." (Zambian Watchdog 2013)</p>
	UPND	<p><i>Linked Ethnicities:</i> Tonga, Lozi, Kaonde, Luvale, Lumda, Chokwe, Mbunda, Tokoleya</p> <p><i>Sources:</i> "Pro-UPND groups: Tonga, Lozi, Kaonde, Luvale, Lumda, Chokwe, Mbunda...and Tokoleya." (Seekings (2018), table 4, p. 11)</p>

B.4.1 Party Name Abbreviations

Benin

- FARD-A: Front d'action pour le renouveau et le développement, FARD-Alafia (Action Front for Renewal and Development)
- PSD: Parti Social-Démocrate (Social Democratic Party)
- RB: Parti de la Renaissance du Bénin (Benin Rebirth Party)

Burkina Faso

- ADF-RDA: Alliance pour la Démocratie et la Fédération–Rassemblement Démocratique Africain (Alliance for Democracy and Federation–African Democratic Rally)
- CDP: Congrès pour la Démocratie et le Progrès (Congress for Democracy and Progress)

- UPC: Union pour le Progrès et le Changement (Union for Progress and Reform)

Cameroon

- CPDM: Cameroon People's Democratic Movement (Rassemblement démocratique du Peuple Camerounais)
- NUDP: National Union for Democracy and Progress (Union Nationale pour la Démocratie et le Progrès)
- SDF: Social Democratic Front (Front Social-Démocratique)

Côte d'Ivoire

- RDR: Rassemblement des Républicains (Rally of the Republicans)
- UPDCI: Union pour la démocratie et la paix en Côte d'Ivoire (Union for Democracy and Peace in Côte d'Ivoire)

Ghana

- NDC: National Democratic Congress
- NPP: New Patriotic Party

Kenya

- NARC: National Rainbow Coalition–Kenya
- ODM: Orange Democratic Movement (Party)
- TNA: The National Alliance
- URP: United Republican Party

Liberia

- CDC: Congress for Democratic Change
- NUDP: National Union for Democratic Progress
- UP: Unity Party

Malawi

- DPP: Democratic Progress Party
- MCP: Malawi Congress Party
- UDF: United Democratic Front

Mali

- RPM: Rassemblement pour le Mali (Rally for Mali)

Mozambique

- RENAMO: Resistência Nacional Moçambicana (Mozambican National Resistance)

Nigeria

- ACN: Action Congress of Nigeria
- ANPP: All Nigeria Peoples Party
- APC: All Progressives Congress
- APGA: All Progressives Grand Alliance
- PDP: People's Democratic Party

Sierra Leone

- APC: All People's Congress
- NDA: National Democratic Alliance
- SLPP: Sierra Leone People's Party

Uganda

- FDC: Forum for Democratic Change
- NRM: National Resistance Movement
- UPC: Uganda People's Congress

Zambia

- MMD: Movement for Multi-Party Democracy
- PF: Patriotic Front
- UPND: United Party for National Development

References

- Aderibigbe, A. (2018, November 20). Jukun back Atiku. *The Nation*. Retrieved from <https://thenationonlineng.net>
- Adinoyi, S. (2018, September 4). Tarok endorse Useni in Plateau guber race. *This Day*. Retrieved from <https://www.thisdaylive.com>
- Alimi, N. (2019, January 20). Igala/Igede communities in Oyo drum up support for Olajide. *Nigerian Tribune*. Retrieved from <https://tribuneonlineng.com>
- Ariong, S. (2016, June 8). Karimojong elders protest Janet Museveni's appointment. *Daily Monitor*. Retrieved from <https://www.monitor.co.ug/691150-691150-7icavj/index.html>
- Austin, O. (2018, September 16). APC: The battle for Esan North East/Esan South East. *Vanguard*. Retrieved from <https://www.vanguardngr.com>
- Awalson, M. (2013). Adamaoua: L'opposition en campagne sénatoriale. Retrieved from <https://www.journalducameroun.com/adamaoua-lopposition-en-campagne-senatoriale/>
- Azeez, A. (2009). Ethnicity, party politics and democracy in Nigeria: Peoples Democratic Party (PDP) as agent of consolidation? *Studies of Tribes and Tribals*, 7(1), 1-9.
- Bah, M. S. (2015, November 30). Sierra Leone news: for 2018 elections... NDA presents Presidential candidate. *Awoko*. Retrieved from <https://awokonewspaper.com>
- Basedaum, M. & Stroh, A. (2011). How ethnic are African parties really? Evidence from four Francophone countries. *International Political Science Review*, 33(1), 5-24.
- Bassett, T. J. (2011). Briefing: Winning coalition, sore loser: Côte d'Ivoire's 2010 presidential elections. *African Affairs*, 110(440), 469-479.
- Battle, M. & Seely, J. C. (2010). It's all relative: Modeling candidate support in Benin. *Nationalism and Ethnic Politics*, 16(1), 42-66.
- BBC (2002). Ex-rebel contests Nigerian poll. Retrieved from <http://news.bbc.co.uk/2/hi/africa/2609113.stm>
- Cahen, M. (2015). [Review of the book *The Battle for Mozambique: The Frelimo-Renamo Struggle (1977-1992)*, by S. A. Emerson]. H-Net Reviews, Retrieved from <http://www.h-net.org/reviews/showrev.php?id=42879>
- Cameroun. Réélection contestée du président Paul Biya. (1992). Retrieved from <https://www.universalis.fr/evenement/11-23-octobre-1992-reelection-contestee-du-president-paul-biya/>
- Chi, A. (1998). Sustainable self-help development efforts in the Cameroon Grassfields. *Development in Practice*, 8(3), 366-371
- Dennis, B. G. & Dennis, A. K. (2008). *Slaves to racism: An unbroken chain from America to Liberia*. New York: Algora Publishing.
- Doodoo, L. (2017). How five leading political parties in Bong County faring across central Liberia? Retrieved from <https://frontpageafricaonline.com/politics/how->

are-the-five-leading-political-parties-in-bong-faring-across-central-liberia/

- Elischer, S. (2013). *Political parties in Africa: Ethnicity and party formation*. Cambridge: Cambridge University Press.
- Ferree, K. & Horowitz, J. (2007). *Identity voting and the regional census in Malawi* (Afrobarometer Working Paper No. 72). Retrieved from <https://www.afrobarometer.org>
- Fridy, K. (2009). The elephant, umbrella, and quarrelling Cocks: Disaggregating partisanship in Ghana's Fourth Republic. *African Affairs*, 106(423), 1-25.
- Ga-dangme youths declare support for Joshua Alabi. (2019). Retrieved from <https://www.adomonline.com/ga-dangme-youths-declare-support-for-joshua-alabi/>
- Give PF time to develop the country- Chief Kabamba. (2013). Retrieved from <https://www.zambiatwatchdog.com/give-pf-time-to-develop-the-country-chief-kabamba/>
- Gombe residents express support for Gwamna. (2018, October 23). *This Day*. Retrieved from <https://www.thisdaylive.com>
- Gorvie, R. (2010). Sierra Leone - One country, One people. Retrieved from <http://www.thepatrioticvanguard.com/sierra-leone-one-country-one-people-5500>
- Hewitt, R. L. & Phiri, S. (2014). Nc'wala: The first fruits ceremony of the Ngoni. Retrieved from <http://kapmalawi.org/pdf/ncwala.pdf>
- Hoffman, B. D. & Long, J. D. (2013). Parties, ethnicity, and voting in African elections. *Comparative Politics*, 45(2), 127-146.
- Idegu, Y. (2018, December 27). Why PDP can't retain power in Plateau, by ex-SSG. *The Nation*. Retrieved from <https://thenationonlineng.net>
- Ijaw group endorses Atiku for president. (2019, February 13). *Vanguard*. Retrieved from <https://www.vanguardngr.com>
- Kabimba has survived the axe, he is Sata's tribesmate – Edward Mumbi. (2013). Retrieved from <https://zambianeye.com/kabimba-has-survived-the-axe-he-is-satas-tribesmate-edward-mumbi/>
- Kandeh, J. D. (1992). Politicization of ethnic identities in Sierra Leone. *African Studies Review*, 35(1), 81-99.
- Kombe, S. (2018). The opposition in Zambia has nothing to offer – Lamba chief. Retrieved from <https://zambiareports.com/2018/01/24/opposition-zambia-nothing-offer-lamba-chief/>
- Konings, P. (2004). Opposition and social-democratic change in Africa: The Social Democratic Front in Cameroon. *Commonwealth & Comparative Politics*, 42(3), 1-23.
- Liberia election: Winston Tubman profile. (2011). Retrieved from <https://www.bbc.com/news/world-africa-15590715>
- Luapula traditional leaders hail the Patriotic Front developmental trail. (2018). Retrieved from

- <https://mwebantu.news/luapula-traditional-leaders-hail-the-patriotic-front-developmental-trail/>
- Macdougall, C. (2011, October 25). Prince as kingmaker: Can an ex-warlord deliver Johnson-Sirleaf the Liberian presidency? *The Christian Science Monitor*. Retrieved from <https://www.csmonitor.com>
- Maina, D. M. (2017). Explainer: a breakdown of what you need to know about the Kenyan elections. Retrieved from <http://theconversation.com/explainer-a-breakdown-of-what-you-need-to-know-about-the-kenyan-elections-81837>
- Matchaya, C. G. (2010). The performance of the Malawi Congress Party in general elections: The role of sectionalism of a regional and ethnic nature. *African Journal of Political Science and International Relations*, 4(6), 221-230.
- Minorities at Risk Project. (2004). Chronology for Temne in Sierra Leone. Retrieved from <https://www.refworld.org/docid/469f38d9c.html>
- Moehler, D. & Conroy-Krutz, J. (2016). Eyes on the ballot: Priming effects and ethnic voting in the developing world. *Electoral Studies*, 42, 99-113.
- Morse, Y. L. (2017, June 2). Cameroon has been in crisis for six months. Here's what you need to know. *The Washington Post*. Retrieved from <https://www.washingtonpost.com>
- Nicaise, M. (2004). Les partis politiques au Benin essay d'approche fonctionnaliste. *l'African Journal of Political Science*, 9(2).
- Nigeria Newsday exclusive: Dora Akunyili Gov. Peter Obi are hypocrites—Dr. Chris Ngige. (2011). Retrieved from <http://saharareporters.com/2011/05/09/nigeria-newsday-exclusive-dora-akunyili-gov-peter-obi-are-hypocrites-dr-chris-ngige>
- Nwachukwu, J. O. (2017, December 10). PDP convention: Full text of Uche Secondus? acceptance speech. *Daily Post*. Retrieved from <https://dailypost.ng>
- Nyambura, Z. (2017). In Kenya, politics split on ethnic divide. Retrieved from <https://www.dw.com/en/in-kenya-politics-split-on-ethnic-divide/a-37442394>
- Okoampa-Ahoofe, Jr., K. (2018). Nyaho Tamakloe is worth less than a nuisance to NPP. Retrieved from <https://www.ghanaweb.com/GhanaHomePage/features/Nyaho-Tamakloe-is-worth-less-than-a-nuisance-to-NPP-696356>
- Omokuvie, P. (2019, January 17). Bauchi: A wide and open race. *The Sun Nigeria*. Retrieved from <https://www.sunnewsonline.com>
- Présidentielle au Burkina: Zéphirin Diabré se présente comme le “soldat de l’alternance.” (2015). Retrieved from <https://www.voaafrique.com/a/presidentielle-burkina-faso-zephirin-diabre-soldat-alternance-favori/3074711.html>
- Raila Odinga - the man Kenyans either love or loathe. (2017). Retrieved from <https://www.bbc.com/news/world-africa-21587054>

- Roessler, P. (2016). *Ethnic politics and state power in Africa: The logic of the coup-civil war trap*. Cambridge: Cambridge University Press.
- Sardauna, F. (2019). Igbo, Yoruba, Nupe women endorse Buhari, Masari for re-election. Retrieved from <https://theworldnews.net/ng-news/igbo-yoruba-nupe-women-endorse-buhari-masari-for-re-election>
- Seekings, J. (2018). *Voters, parties, and elections in Zambia* (CSSR Working Paper No. 422). Retrieved from http://www.sll.uct.ac.za/sites/default/files/image_tool/images/256/Publications/WP%20422%20-Seekings.pdf
- Somé, Constantin. (2009). *Pluralisme socio-ethnique et démocratie: cas du Bénin* (Unpublished masters dissertation). Université du Québec à Montréal, Montreal.
- Taylor, B. K. (2011). *Sierra Leone: The land, its people and history*. Dar es Salaam: New Africa Press.
- Udeh, K. (2018, July 27). Ekiti 2018: APC's Victory and the Political Nicety of Dr. Orji Kalu. *The Sun Nigeria*. Retrieved from <https://www.sunnewsonline.com>
- Vœux de nouvel an: Le président de l'ADF-RDA communique avec «ses troupes». (2019). Retrieved from <https://lefaso.net/spip.php?article87484>
- The World Bank. (2008). *Local governance and national crisis in Côte d'Ivoire: Opportunities and challenges for community-based development and peace-building*. Washington, DC.
- Yi Dionne, K. & Dulani, B. (2014, May 19). On the eve of Malawi's election. *The Washington Post*. Retrieved from <https://www.washingtonpost.com>
- Zambia's President Edgar Lungu declared election winner. (2016). Retrieved from <https://www.bbc.com/news/world-africa-37086365>

B.5 Village Characteristics

B.5.1 Variable definitions

Distance from Cities is computed using GIS software as the geodesic distance from the DHS cluster coordinates to the closest city classified as such in [Natural Earth \(2018\)](#).

Distance from Improved Roads is the geodesic distance from the DHS cluster coordinates to the closest point on a colonial road (in existence in 1960). Computed using GIS software. Data on colonial roads come from [Jedwab and Storeygard \(2021\)](#).

Elevation is computed using GIS software from cell-level data on elevation at a 30 arc-second resolution from [U.S. Geological Survey \(1996\)](#) as the elevation value at the DHS cluster coordinates.

Terrain Ruggedness is computed using GIS software from cell-level data on elevation at a 30 arc-second resolution from [U.S. Geological Survey \(1996\)](#). Given the grid cell data, picture a 3×3 block of 9 cells and let $e_{r,c}$ be the elevation of the cell in row r , column c of the grid. Following [Nunn and Puga \(2012\)](#), we compute ruggedness as $\sqrt{\sum_{i=r-1}^{r+1} \sum_{j=c-1}^{c+1} (e_{i,j} - e_{r,c})^2}$, that is, the square root of the sum of all the squared differences in elevation between the middle cell and the surrounding 8 cells.

Agricultural Suitability is mean agricultural suitability in the constituency. The latter variable is computed, using GIS software, as the within-constituency zonal statistic of a raster provided by [Ramankutty, Foley, Norman, and McSweeney \(2002\)](#), which we resample to a resolution of 30 arc-seconds prior to computing the statistic.

Malaria Suitability is mean stability of malaria transmission in the constituency. The latter variable is computed, using GIS software, as the within-constituency zonal statistic of a raster provided by [Kiszewski, Mellinger, Spielman, Malaney, Ehrlich, and Sachs \(2004\)](#), which we resample to a resolution of 30 arc-seconds prior to computing the statistic.

Table B.16: Summary Statistics of DHS Cluster/Village Characteristics

Variable	Mean	Std. Dev.	Min	Max	Observations
Latitude	0.727	10.986	-26.817	16.656	13526
Longitude	13.781	19.532	-17.498	41.877	13526
Distance from Improved Roads (km)	16.942	29.636	0	306.802	13526
Distance from Cities (km)	32.02	27.429	0.077	161.295	13526
Elevation (m)	584.648	572.963	-0.556	3193.583	13526
Terrain Ruggedness	56.756	73.387	1.024	1165.691	13526
Agricultural Suitability	0.407	0.214	0	0.987	12954
Malaria Suitability	14.492	10.396	0	37.609	13526

Notes. The table reports the summary statistics of all village-level variables used in the empirical analysis. Unit of observation is a DHS cluster/village.