Materials for Online Appendix

1 Survey methodology

Survey design

Each instrument was comprised of six separate surveys, one each for the chief executive (elected head of the district assembly), coordinating director (head civil servant), presiding member of the assembly (elected), finance officer, human resources/personnel officer, and revenue officer (tax collector). The survey instruments were developed after six months of fieldwork, and in concert with informants in six case study districts and the Ministry of Local Government. The survey was pre-tested and then implemented countrywide between May and August 2012.

The Deputy Minister of Local Government informed each district in a letter that the government did not seek any personal or identifying information from respondents, and that it hoped the exercise would increase understanding of the functioning of districts. Districts then received hand-delivered survey instruments that contained no identifying information such as the district's name. Districts were instructed to identify only their region in their response. They were assured that the government would not be able to connect completed surveys with individual districts. Completed surveys were then returned to me, not the Ministry.

Survey response rates

The response rate was 88% (n=149), meaning 88% of districts submitted some completed part of the questionnaire.[24] Summary statistics show no reason to believe that non-responses bias the data. Response rates for individual respondents are shown in Table 4 and are visualized in Figure A1. Cells show responses as percentages of all districts (not percentages of respondents). The response rate is lowest for the presiding member because they do not work full time at the district building.

[24] 85% of respondents (not all districts) completed at least five of the six sections.

Figure A1: Map of survey responses

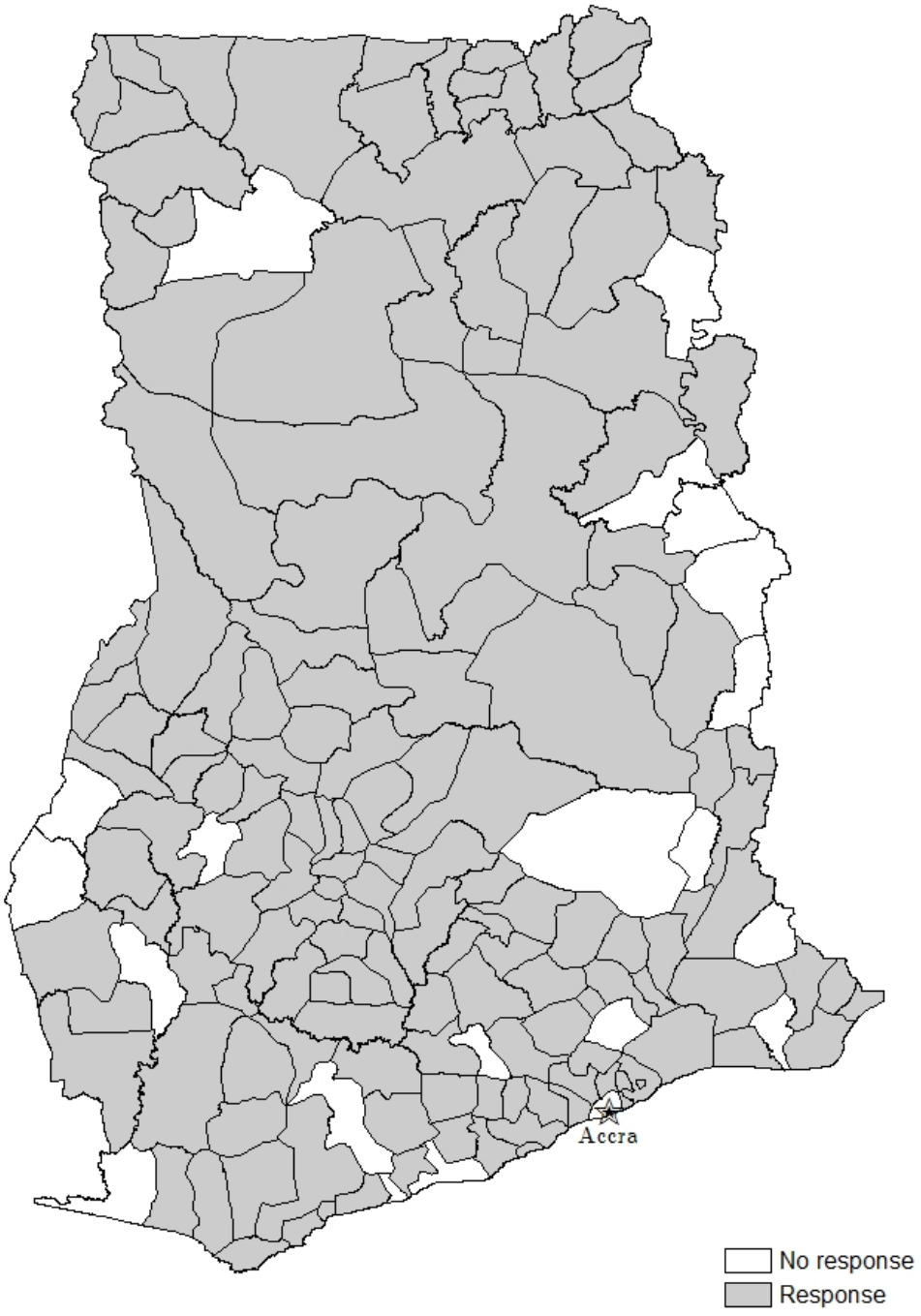


Table 4: Response rates for district assembly survey

Response rates by respondent	%
All districts	88
Chief executives	78
Presiding members	71
Coordinating directors	84
personnel officers	78
revenue officers	82
Finance officers	81

Non-responses were quasi-random owing to the sudden death of Ghana's president on July 24, 2012. Districts that had not completed their survey by that point ultimately returned no survey at all because government business ground to a halt for the late president's funeral. Table 5 shows no systematic relationship between the predicted probability that a district returned at least some part of the survey (model 1) or the total number of sections returned (model 2).

Demographic data

Population, urban/rural, and ethnicity data come from the 2010 Census. To calculate the degree of ethnic heterogeneity at the district level I use the conventional formula for *Ethnic Fractionalization*. This is a simple Herfindahl concentration index where $ELF = 1 - \sum_{i=1}^{n} s_i^2$ where s_i is the share of group $i(i{=}1,\ldots,n)$. Measuring the size of key ethnic groups at the district level is not only important for a study of African politics, since a concern with the effects of ethno-regionalism remains important to Africanist political science, but also because the exact size of key ethnic groups is important to my own argument. In particular, it is important to accurately estimate the size of Akan sub-groups, since four major sub-groups

Table 5: Regression for bias in survey responses

	(1 Marginal effects from probit) Any survey section complete	(2 OLS) Num. sections complete (max 6)
Population (2010, log)	-0.043	-0.628
	(0.071)	(0.485)
Urban (2010)	0.003*	0.012
	(0.002)	(0.010)
District Age	0.000	-0.002
	(0.003)	(0.024)
Distance to Mine (log)	-0.009	-0.143
	(0.031)	(0.216)
Crop value	-0.038	-0.017
	(0.032)	(0.223)
Illiterate (2010)	-0.000	-0.012
	(0.004)	(0.028)
No Toilet (2010)	0.002	0.008
	(0.002)	(0.015)
Distance to Kumasi (log)	-0.126***	-0.404*
	(0.045)	(0.223)
Distance to Accra (log)	0.020	-0.033
	(0.047)	(0.312)
Area (Km. Sq., Land)	-0.002	-0.021
	(0.042)	(0.259)
Ethnic Fractionalization (2010)	0.001	0.004
	(0.001)	(0.010)
Electricity (2010)	-0.002	-0.015
	(0.002)	(0.016)
Cocoa Dummy	-0.022	-0.364
	(0.072)	(0.458)
Margin in Pres. Elections (00-12)	0.000	-0.002
	(0.002)	(0.011)
Ethnic bloc difference	-0.000	-0.002
	(0.001)	(0.007)
Constant		15.818***
		(6.039)
Observations	168	168
R^2		0.051

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of the Akan are loyal to the NPP while other Akan groups are not.[25]

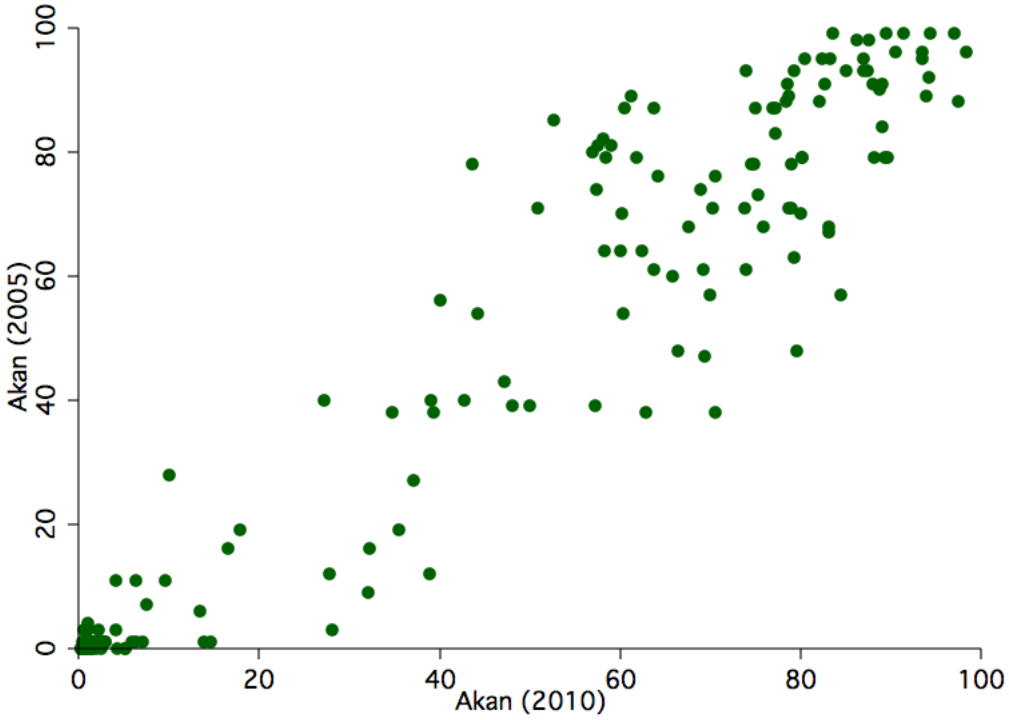
Statistical inquiries into Ghanaian politics frequently include measures of the size of Akan populations, but they do this in error since the Akan are heterogeneous in political behavior (see for example Bossuroy, 2011). The problem is often one of data availability, however, since census data records only if a person is Akan rather than any of its important sub-groups such as Asante, Akyem or Kwahu. Contacts at the Ghana Statistical Service tell me such fine-grained ethnicity data is actually recorded during the census but that all Akan groups are pooled when the data are entered electronically.

To overcome this problem I estimate the district proportions of Akan sub-groups using the Ghana Living Standards Survey 5 (2005). This is a nationally representative survey of almost 40,000 people in almost 9,000 households. Figure A2 presents my estimates of the size of Akan populations by district and allows for a comparison of the accuracy of my Akan sub-group data with 'true' values from the census. Each dot is a district. The horizontal axis shows the percentage of the district population that is Akan using census data from 2010. The vertical axis presents the same using household survey data from 2005. We see that although there is variance along the line, the 2005 estimates do correspond broadly to the 2010 data. The variables are highly correlated ($r^2 = .96$).

There are four main ways in which the Akan sub-group data are measured with error. First, and most obvious, the 2005 survey data are measured with greater error since they are not a census. Second, there is a time lag between the two data collection periods. Third, while the census data is generated at the level of the 170 districts, the 2005 data was generated for the 138 districts in existence from 2004-08. For some districts that were later split I had to use the same ethnicity values as the mother district. The Akan sub-group data

[25]Akans make up 47.5 percent of all Ghanaians according to the 2010 Census. Of the Akans, Asantes, Akyems, Akuapem and Kwahus generally vote NPP, while Fantes, Nzemas and Sefwis generally vote NDC. Akan sub-groups fitting into neither category include Brongs, Ahantas, Ahafos, Agonas, Assins, Denkyiras, Awutus, Aowins, and Wasa.

Figure A2: Akan population using GLSS 5 2005 and Census 2010 data



are thus clearly imperfect, but they are all we have. In the analysis that follows I use the Akan sub-group data advisedly given the weakness I have outlined, but knowing also that it is important to not treat the Akan as homogeneous. Last, it is important to note that I only use GLSS 5 2005 data for Akan sub-groups. For all other major ethnic groups I use 2010 census data.

Geospatial data

To calculate distances I used a GIS shapefile and Google Earth to obtain geo-coded locations of every district capital and I then used a form of the Haversine equation for calculating distances between two points in kilometers.[26] Care is required in calculating the physical size of districts since some border Lake Volta, the largest reservoir by surface area in the

[26]In Excel I used the formula =ACOS(COS(RADIANS(90-A2)) *COS(RADIANS(90-A3)) +SIN(RADIANS(90-A2)) *SIN(RADIANS(90-A3)) *COS(RADIANS(B2-B3))) *6371 (earth's radius)

world. To avoid including water in my calculations I overlaid two GIS shapefiles, one with boundaries and one with Lake Volta, and manually clipped out from each district bordering Lake Volta the water portion and then used ArcMap to re-estimate land area in kilometers squared.

To get an accurate list of major mines I combined data from the Annual Reports of the Ghana Chamber of Mines and Minerals Yearbooks from the US Geological Service with Google Earth to first establish a list of all mines currently active and operated by a legally recognized company. This coding decision ensures my data cover all districts with mines of significant size. There is no concern that using a list of legally operating firms under-counts actual mining operations. Mines of significant size are operated by multi-national mining companies such as Anglo-Gold Ashanti and their operations are so large as to be easily observable by satellite imagery. Small-scale operations, by contrast, are labor intensive, technology-lite, of low yield and do not contribute to central or local revenue. Using the list of major mining operations I then located each mine using Google Earth to obtain its coordinates, and used my geospatial formula to calculate the distance of each district capital to the nearest mine. The variable is then calculated as a continuous rather than dummy variable to allow for spillover effects into neighboring districts. Note that Ghana's coastal oil had not started flowing during the period under study.

To these measures of household welfare I add data on district endowments more generally. First, the dataset includes a measure on district proximity to major mines, as explained already. Second, the dataset includes multiple estimates of the size of district cocoa economies. Cocoa data at the district level are surprisingly hard to collect. The Ministry of Agriculture collects data through its local offices on the yields of multiple major agricultural commodities such as yams and plantains, but it does not collect data on cocoa. The task of collecting data on cocoa falls to the Ghana Cocoa Board (COCOBOD), which does not report district level cocoa data. To measure cocoa yields, therefore, I use data from a nationally representative household survey in 2005, which asked farming households about

past and future farming activities. I confirmed the accuracy of this data by overlaying it with maps of exposure to direct sunlight and of soil types.

The dataset also includes Ministry of Agriculture data on the cultivated area, production, yield and value of major cash crops for 2010. To calculate the value of each crop, and the value of all crop production for each district, I used bi-monthly commodity wholesale price data from 2010. I took the median price of each crop across all available towns to get median prices across the country for one month. I then calculated median prices for the year using the monthly median price for each commodity. For each district this generates one value, which is the total estimated value of all crop production in Ghana Cedis.

2 Case selection methodology

Establishing variation among cases

I used variation in district tax performance to generate a rank ordered list of districts from which to choose cases. Estimating the quality of district tax collection took the form of regression analysis to compare what a district actually collects with what it 'could' collect. The methodology is drawn from the tax capacity literature, in which actual income from taxation can be compared to a hypothetical yield given a tax base, such as GDP in the case of a country or agricultural or industrial output for a subnational unit (Pessino & Fenochietto, 2010; Stotsky & WoldeMariam, 1997; Le *et al.* , 2008). The difference between what a polity 'could' collect and what it actually collects is a measurement of capacity to tax. At the Ministry of Local Government in Accra, I gathered data on district revenue, which included locally collected taxes as well as transfers from central government

Analysis took the form of a simple OLS regression in which the dependent variable (*Internally Generated Funds*, or IGF) was locally collected taxes and the explanatory variables were covariates that should plausibly affect tax collection using available data, such as agricultural output, urbanization, population density and distance to major markets. Results are shown in Table 6. Estimated models do not include tax rates, since these data were not centrally housed at the Ministry of Local Government. Although tax rates would affect actual collection, I was not overly concerned with how this might affect my case selection, because if variation in tax collection was driven by differential rate settings, understanding the reasons for these differential rates would simply open a line of inquiry for my interviews. From this regression I obtain a residual, which is the difference between a district's actual tax collection and a district's predicted tax collection using the regression estimates. This residual becomes a rough measure of the quality of a district's tax collection, where residuals above zero indicate better than expected tax collection.

Table 6: OLS models for tax collection

	IGF per cap. (median, deflated)
Population (2010)	-0.00000365***
	(-3.78)
Urban (2010)	0.0167***
	(3.49)
Illiterate (2010)	-0.0355*
	(-2.44)
No Toilet (2010)	0.0156
	(1.73)
Electricity (2010)	-0.00320
	(-0.49)
Land Area (Km. Sq.)	-0.00000236
	(-0.07)
Population Density	0.000537**
	(2.72)
District Age	-0.0370***
	(-5.07)
Distance to Accra (log)	-0.0459
	(-0.26)
Distance to Region Capital (log)	0.0107
	(0.15)
Distance to Kumasi (log)	0.00594
	(0.06)
Distance to Mine (log)	-0.285*
	(-2.46)
Cocoa Dummy	-0.251
	(-1.24)
Crop value per capita (log, Ghana Cedi)	0.0353
	(0.31)
Constant	3.703**
	(2.73)
Observations	165
R^2	0.478

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Selecting cases

My measures of tax capacity allowed me to sort all districts from best to worst. I used five criteria to narrow the list from 170 districts to a few cases for immersive study. First, cases needed to vary along my explanatory variable of political competition. It was not possible to simultaneously study a competitive district, an NPP stronghold and an NDC stronghold, since nowhere in the country can one find such districts near one another. My strongholds are thus NPP strongholds, since I was based in the Ashanti Region. Second, I narrowed my sample to Akan districts to take advantage of my training in the Akan/Twi language and to control for potential effects of differences in ethnic social structure or the nature of traditional authority. Third, I excluded districts with very large populations since I did not feel I could adequately get to know key actors in huge districts with my limited resources. Since two-thirds of Ghanaians live in districts with populations under 125,000, this criterion did not severely restrict the sample. Fourth, I chose districts that varied in age, since the number of districts has increased over time and institutional continuity and learning should affect the quality of government.

These considerations reduced the list of possible cases from 170 to 19. Five of the 19 met my definition of a party stronghold with 75 percent vote share for one party in the previous election. Three of the 19 met my criteria for a competitive district, which was an average margin between the NDC and NPP presidential candidates from 2000-2008 of no more than 10 percent. From these remaining districts I chose cases where performance in tax collection above or below expectation, defined as at least one quarter of one standard deviation above or below zero for tax collection residuals.

I used a GIS shapefile and Google Earth to obtain geo-coded locations of every district capital and then used a form of the Haversine equation for calculating distances.[27] Care was required in calculating the physical size of districts since some border Lake Volta, the

[27]In Excel the formula is=ACOS(COS(RADIANS(90-A2)) *COS(RADIANS(90-A3)) +SIN(RADIANS(90-A2)) *SIN(RADIANS(90-A3)) *COS(RADIANS(B2-B3))) *6371 (earth's radius)

largest reservoir by surface area in the world. I overlaid two GIS shapefiles, one with district boundaries and one with Lake Volta, and manually clipped out from each district bordering Lake Volta the water portion and then used ArcMap to re-estimate land area in kilometers squared.

3 Main models without Metropolitan Assemblies

Table 7: Main models without Metroplitan Assemblies (1)

	(1) Staff size	(2) Staff size (log)	(3) Staff size	(4) Staff size
Margin (00-12 Avg., Presidential)	-0.563**	-0.005***		
	(0.231)	(0.002)		
Margin (2008, Presidential)			-0.450**	
			(0.223)	
Margin (2008, Parliamentary)				-0.452
				(0.283)
Full controls	Yes	Yes	Yes	Yes
Constant	-1139.133***	-2.801**	-1156.903***	-1145.578***
	(355.253)	(1.330)	(354.673)	(349.224)
Observations	128	128	128	128
R^2	0.547	0.612	0.543	0.541

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Main models without Metroplitan Assemblies (2)

	(1) Low-level staff	(2) Low-level staff (log)	(3) Low-level staff	(4) Low-level staff
Margin (00-12 Avg., Presidential)	-0.119	-0.003		
	(0.124)	(0.003)		
Margin (2008, Presidential)			-0.115	
			(0.116)	
Margin (2008, Parliamentary)				-0.076
				(0.144)
Full controls	Yes	Yes	Yes	Yes
Constant	-440.573***	-5.480**	-442.248***	-444.787***
	(160.133)	(2.090)	(159.741)	(157.891)
Observations	125	125	125	125
R^2	0.434	0.438	0.434	0.431

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Main models without Akan regions* (1)

	(1) Staff size	(2) Staff size (log)	(3) Staff size	(4) Staff size
Margin (00-12 Avg., Presidential)	-1.737**	-0.005*		
	(0.707)	(0.003)		
Margin (2008, Presidential)			-1.831**	
			(0.730)	
Margin (2008, Parliamentary)				-0.927
				(0.708)
Full controls	Yes	Yes	Yes	Yes
Constant	-2736.335***	-6.010***	-2702.259***	-2949.811***
	(601.324)	(2.048)	(587.931)	(625.896)
Observations	84	84	84	84
R^2	0.810	0.723	0.813	0.798

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
*Excludes Ashanti, Brong-Ahafo, Eastern, Central and Western Regions.

Table 10: Main models without Akan regions* (2)

	(1) Low-level staff	(2) Low-level staff (log)	(3) Low-level staff	(4) Low-level staff
Margin (00-12 Avg., Presidential)	-0.440*	-0.004		
	(0.256)	(0.005)		
Margin (2008, Presidential)			-0.498*	
			(0.253)	
Margin (2008, Parliamentary)				-0.221
				(0.248)
Full controls	Yes	Yes	Yes	Yes
Constant	-909.781***	-9.872***	-892.009***	-970.583***
	(234.878)	(3.033)	(231.810)	(238.207)
Observations	82	82	82	82
R^2	0.748	0.604	0.752	0.739

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
*Excludes Ashanti, Brong-Ahafo, Eastern, Central and Western Regions.

4 Main models without majority Akan regions

5 Incumbent security of tenure

Table 11: Summary statistics for incumbent tenure: No. of district chief executives (DCEs)

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
No. DCEs	1.26	0.56	1	4
N		120		

Table 12: Regression models for incumbent tenure

	(1)	(2)	(3)
	No. DCEs	No. DCEs	No. DCEs
Margin (00-12 Avg., Presidential)	-0.006*		
	(0.003)		
Margin (2008, Presidential)		-0.006**	
		(0.003)	
Margin (2008, Parliamentary)			-0.005*
			(0.003)
Full controls	Yes	Yes	Yes
Constant	-0.268	-0.305	-0.217
	(1.734)	(1.741)	(1.779)
Observations	119	119	119
R^2	0.172	0.175	0.163

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 Intra-party competitiveness

Table 13: Summary statistics for measures of intra-party competition and turnout

Variable	Mean	Std. Dev.	Min.	Max.	N
MP is NDC (2008)	0.5	0.5	0	1	170
St. Dev. in Average Margin (2000-2012)	6.76	5	0.22	29.01	167
Candidates per EA 2006	2.97	0.66	1.19	6.27	133
Pcnt of EAs contested 2006	92.08	8.81	65.63	100	134
Turnout MMDA Election 2006	48.71	10.54	13.82	73.73	132
Turnout (2008)	70.55	5.11	53.84	80.60	170

Table 14: Regressions for staff size using measures of intra-party competition and turnout (1)

	(1) Staff size	(2) Staff size	(3) Staff size	(4) Staff size	(5) Staff size	(6) Staff size
MP is NDC (2008)	-8.915					
	(20.959)					
St. Dev. in Average Margin (2000-2012)		2.859				
		(1.883)				
Candidates per EA 2006			-3.672			
			(11.086)			
Pcnt of EAs contested 2006				0.487		
				(0.850)		
Turnout MMDA Election 2006					-0.632	
					(0.940)	
Turnout (2008)						-3.391
						(2.535)
Full controls	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-1621.477***	-1660.201***	-1232.555**	-1325.630**	-1258.459*	-1360.903***
	(378.283)	(381.542)	(498.425)	(591.739)	(690.361)	(357.127)
Observations	131	129	73	73	72	131
R^2	0.687	0.696	0.590	0.592	0.549	0.695

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Regressions for low-skill staff size using measures of intra-party competition and turnout (2)

	(1) Low-level staff	(2) Low-level staff	(3) Low-level staff	(4) Low-level staff	(5) Low-level staff	(6) Low-level staff
MP is NDC (2008)	-1.869					
	(8.847)					
St. Dev. in Average Margin (2000-2012)		1.436				
		(1.216)				
Candidates per EA 2006			-7.883			
			(5.350)			
Pcnt of EAs contested 2006				-0.168		
				(0.358)		
Turnout MMDA Election 2006					-0.571	
					(0.372)	
Turnout (2008)						0.252
						(1.172)
Full controls	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-580.256***	-590.882***	-427.666	-454.390	-476.582	-602.537***
	(167.359)	(169.359)	(293.930)	(335.923)	(307.534)	(194.792)
Observations	127	125	70	70	70	127
R^2	0.521	0.535	0.463	0.447	0.459	0.522

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7 Bureaucratic quality

Table 16: Summary statistics for bureaucratic quality

Variable	Mean	Std. Dev.	Min.	Max.	N
Coord. Dir. Yrs in System	13.56	8.66	1	40	137
Coord. Dir. Masters Degree	0.51	0.5	0	1	141
Age (Coord. Dir.)	47.77	6.92	30	60	139

Table 17: Regression models for staff size using bureaucratic quality

	(1) Staff size	(2) Staff size (log)	(3) Staff size	(4) Staff size	(5) Low-level staff	(6) Low-level staff (log)	(7) Low-level staff	(8) Low-level staff
Margin (00-12 Avg., Presidential)	-1.486***	-0.006***			-0.438**	-0.004		
	(0.501)	(0.002)			(0.208)	(0.003)		
Margin (2008, Presidential)			-1.495***				-0.375**	
			(0.530)				(0.186)	
Margin (2008, Parliamentary)				-1.610***				-0.495*
				(0.602)				(0.257)
Coord. Dir. Yrs in System	-1.718	-0.000	-1.815	-2.031*	-0.936	-0.001	-0.952	-1.037
	(1.148)	(0.005)	(1.151)	(1.173)	(0.766)	(0.009)	(0.765)	(0.790)
Coord. Dir. Masters Degree	-5.113	-0.024	-5.559	-1.853	-10.837	-0.129	-10.947	-9.906
	(17.284)	(0.066)	(17.278)	(17.380)	(9.649)	(0.108)	(9.720)	(9.424)
Age (Coord. Dir.)	1.781	-0.003	1.952	1.772	0.860	0.005	0.878	0.874
	(1.536)	(0.005)	(1.540)	(1.532)	(0.670)	(0.011)	(0.657)	(0.681)
Full controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-1507.561***	-3.258***	-1505.915***	-1455.334***	-508.645***	-6.062***	-514.694***	-485.536***
	(382.823)	(1.195)	(377.756)	(378.770)	(176.777)	(2.155)	(174.498)	(179.397)
Observations	120	120	120	120	116	116	116	116
R^2	0.730	0.713	0.731	0.729	0.554	0.526	0.552	0.555

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8 Ethnic diversity

Table 18: Summary statistics for ethnic diversity

Variable	Mean	Std. Dev.	Min.	Max.
Ethnic fractionalization (2010)	41.15	21.27	2	82
N		170		

Table 19: Regression models for the size of district staff with ethnic diversity

	(1) Staff size	(2) Staff size (log)	(3) Staff size	(4) Staff size
Margin (00-12 Avg., Presidential)	-1.286***	-0.005***		
	(0.437)	(0.002)		
Margin (2008, Presidential)			-1.268***	
			(0.472)	
Margin (2008, Parliamentary)				-1.330**
				(0.519)
Ethnic fractionalization (2010)	0.314	-0.002	0.294	0.295
	(0.401)	(0.002)	(0.414)	(0.386)
Full controls	Yes	Yes	Yes	Yes
Constant	-1561.641***	-2.973***	-1554.602***	-1510.333***
	(337.515)	(1.053)	(332.018)	(328.858)
Observations	131	131	131	131
R^2	0.704	0.701	0.704	0.702

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 20: Regression models for the size of low-level district staff with ethnic diversity

	(1) Low-level staff	(2) Low-level staff (log)	(3) Low-level staff	(4) Low-level staff
Margin (00-12 Avg., Presidential)	-0.342**	-0.004		
	(0.164)	(0.003)		
Margin (2008, Presidential)			-0.299*	
			(0.154)	
Margin (2008, Parliamentary)				-0.373*
				(0.211)
Ethnic fractionalization (2010)	0.240	-0.001	0.248	0.230
	(0.239)	(0.003)	(0.257)	(0.225)
Full controls	Yes	Yes	Yes	Yes
Constant	-564.393***	-5.638***	-568.228***	-546.590***
	(161.282)	(2.009)	(163.662)	(158.479)
Observations	127	127	127	127
R^2	0.529	0.512	0.527	0.529

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

9 Social capital

Table 21: Operationalizing and measuring social capital

Concept	Survey question	Variable	N	Mean	Std. Dev.	Min	Max
General trust	Would you say that most of the time people try to be helpful to your household or that they are just looking out for themselves?	Others look out for themselves	110	43.33	18.89	2.5	83.1
Communal relations	Sometimes different groups living in the same area live together peacefully. Other times there is tension and disagreement among different groups. How would your describe your community/neighbourhood these days?	Peaceful groups	110	86.12	11.01	41.1	100
	Do members of your household feel safe walking down your street at night?	Safe to walk ('very' and 'somewhat')	110	93.39	6.99	61.5	100
Civic engagement	How often, if at all, do members of your household read a daily newspaper or have one read to them or listen to the radio?	Listens to radio ('regularly' and 'occasionally')	110	83.05	12.29	46.5	98.3
		Reads newspaper ('regularly' and 'occasionally')	110	18.05	11.42	1.2	51.4
	Turnout in 2000 presidential election	Turnout (2000)	104	62.56	6.22	40.9	73.9
'Bridging' social capital	How common is it for people in this area to marry outside their religion/ethnic group?	Inter-marriage (somewhat common)	110	22.62	13.26	.2	52.9
Collective action	Is there a policing or neighbourhood watch system in your area?	Neighborhood watch	110	18.12	13.61	0.3	68.5

Table 22: Regression models for staff size using measures of social capital

	(1) Staff size	(2) Staff size	(3) Staff size	(4) Staff size	(5) Staff size	(6) Staff size	(7) Staff size	(8) Staff size
Others look out for themselves	2.589**							
	(1.006)							
Peaceful groups		-0.648						
		(2.193)						
Safe to walk			-5.248					
			(3.599)					
Conflict is political				9.570***				
				(2.993)				
Listens to radio					0.491			
					(2.100)			
Reads newspaper						-0.601		
						(2.923)		
Inter-marriage							1.653	
							(1.532)	
Neighborhood watch								1.360
								(1.152)
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-351.768	-274.452	-372.595	-1736.996	-275.121	-199.481	-1120.438	37.908
	(1571.187)	(1533.571)	(1316.109)	(1063.843)	(1545.877)	(1563.440)	(1958.663)	(1288.027)
Observations	34	34	34	33	34	34	34	34
R^2	0.949	0.930	0.941	0.952	0.930	0.930	0.934	0.934

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 23: Regression models for low-skill staff size using measures of social capital

	(1) Low-level staff	(2) Low-level staff	(3) Low-level staff	(4) Low-level staff	(5) Low-level staff	(6) Low-level staff	(7) Low-level staff	(8) Low-level s
Others look out for themselves	0.357							
	(0.232)							
Peaceful groups		-0.587						
		(0.458)						
Safe to walk			-0.092					
			(0.852)					
Conflict is political				1.560**				
				(0.685)				
Listens to radio					0.109			
					(0.475)			
Reads newspaper						0.293		
						(0.524)		
Inter-marriage							0.281	
							(0.308)	
Neighborhood watch								0.311
								(0.324)
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-19.449	-38.761	-18.310	-354.572	-17.218	-44.157	-168.497	51.415
	(390.135)	(364.549)	(395.476)	(298.039)	(393.786)	(409.041)	(462.636)	(354.108
Observations	32	32	32	31	32	32	32	32
R^2	0.969	0.968	0.965	0.976	0.965	0.965	0.966	0.968

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 24: Summary statistics for social capital models using 2000 and 2003 data

Variable	Mean	Std. Dev.	Min.	Max.	N
Others look out for themselves	43.33	18.89	2.5	83.10	110
Peaceful groups	86.12	11.01	41.1	100	110
Safe to walk	93.39	6.99	61.5	100	110
Listens to radio	83.05	12.29	46.5	98.3	110
Reads newspaper	18.05	11.42	1.2	51.4	110
Inter-marriage	22.62	13.26	0.2	52.9	110
Neighborhood watch	18.12	13.61	0.3	68.5	110
Turnout (2000)	62.56	6.22	40.9	73.90	104
Ethnic fractionalization	0.41	0.2	0.1	0.8	110
Population (2003)	187857.65	237319.51	49253	1818050	110
% Urban (2000)	30.02	21.75	0	100	109
Distance region capital	66.48	52.15	1	293	110
Distance Accra	249.23	167.84	1	638	110
Distance Kumasi	198.87	127.52	1	509	110
Distance mine	174.39	136.21	1	505	110
Owns radio (2003)	66.42	10.02	41.5	90.8	110
Illiterate (2000)	43.66	17.42	12	87	110
Good wall (2003)	30.22	23.42	0.70	92.60	110
Poor roof (2003)	24.27	24.58	0	90.3	110
Good water (2003)	30.61	26.01	0	99.5	110

Table 25: Regression models for electoral competition and social capital using 2000 and 2003

	(1) Others look out for themselves	(2) Peaceful groups	(3) Safe to walk	(4) Listens to radio	(5) Reads newspaper	(6) Inter-marriage	(7) Neighborhood watch	(8) Turnout (2000)
Margin (2000, Presidential)	-0.198**	0.042	0.040	-0.082**	0.014	-0.001	-0.042	0.040
	(0.081)	(0.047)	(0.036)	(0.033)	(0.050)	(0.056)	(0.061)	(0.035)
Ethnic fractionalization	-2.106	9.575*	-0.297	2.163	2.543	-1.718	-16.009**	1.603
	(9.073)	(5.627)	(3.373)	(3.660)	(5.527)	(7.804)	(7.588)	(3.590)
Population (2003, log)	-1.353	-7.955***	-1.582	-3.983**	0.985	1.392	-0.378	-4.169***
	(4.398)	(2.489)	(1.546)	(1.558)	(1.935)	(3.076)	(3.411)	(1.432)
% Urban (2000)	0.025	0.109	0.112*	0.058	-0.062	0.009	0.172	0.125***
	(0.184)	(0.092)	(0.066)	(0.068)	(0.073)	(0.139)	(0.125)	(0.044)
District age (2003)	1.586	-2.476	-0.167	0.369	-1.178	3.630	2.628	1.865
	(3.049)	(1.926)	(1.273)	(1.672)	(1.667)	(2.656)	(2.972)	(1.531)
Distance regional cap. (log)	2.305	1.064	0.229	-0.690	-1.589**	1.330	-0.383	-0.592
	(1.686)	(1.098)	(0.762)	(0.494)	(0.780)	(1.289)	(1.217)	(0.410)
Distance Accra (log)	-4.723	-1.802	-0.893	-2.621***	-0.890	-0.536	1.795	-1.247
	(2.878)	(1.957)	(1.343)	(0.915)	(1.130)	(2.045)	(2.222)	(0.902)
Distance Kumasi (log)	-1.272	-0.986	-1.357*	-4.807***	0.009	-0.997	-1.461	-1.725**
	(3.042)	(1.586)	(0.815)	(1.300)	(1.019)	(1.709)	(2.064)	(0.843)
Distance mine (log)	-3.560	-3.629***	-0.295	-0.676	-1.120	0.226	2.488	-0.711
	(2.159)	(1.077)	(0.994)	(0.688)	(0.714)	(1.550)	(1.835)	(0.581)
Illiterate (2000)	0.296***	0.022	-0.065	0.072	-0.121**	0.406***	0.143	-0.089*
	(0.106)	(0.073)	(0.053)	(0.052)	(0.057)	(0.081)	(0.123)	(0.048)
Good wall (2003)	0.305**	0.090	-0.152**	-0.058	0.129*	0.042	-0.123	-0.110**
	(0.131)	(0.085)	(0.060)	(0.054)	(0.069)	(0.119)	(0.100)	(0.053)
Poor roof (2003)	-0.033	0.064	-0.093*	0.044	-0.082	0.159	-0.128	0.011
	(0.105)	(0.065)	(0.053)	(0.047)	(0.056)	(0.098)	(0.101)	(0.050)
Good water (2003)	-0.066	-0.110	-0.040	0.051	0.114**	0.069	0.021	0.001
	(0.113)	(0.071)	(0.055)	(0.042)	(0.056)	(0.099)	(0.083)	(0.038)
Owns radio (2003)				0.602***				
				(0.092)				
Constant	63.156	234.167***	132.601***	124.942***	39.550	-69.276	-24.962	106.242***
	(68.380)	(51.586)	(37.090)	(38.779)	(39.373)	(64.958)	(71.575)	(33.061)
Observations	103	103	103	103	103	103	103	103
R^2	0.405	0.288	0.196	0.718	0.606	0.231	0.173	0.257

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

10 Public services

Table 26: Summary statistics for public services

Variable	Mean	Std. Dev.	Min.	Max.	N
Revenue per capita (2010)	21.49	9.32	5.29	56.83	167
No toilet	30.73	29.57	2.84	93.27	170
Bad garbage	75.63	16.57	6.62	96.77	170
Capital Exp. (%)	79.45	13.2	35.68	98.40	161

Table 27: Correlations between public services and electoral competition

	Margin (00-12 Avg., Presidential)	Revenue per capita (2010)	No toilet	Bad garbage	Capital Exp. (%)
Margin (00-12 Avg., Presidential)	1				
Revenue per capita (2010)	0.0163	1			
No toilet	-0.0763	0.259***	1		
Bad garbage	0.103	0.0616	0.309***	1	
Capital Exp. (%)	0.000486	0.198**	0.535***	0.575***	1

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 28: Regression models for public services and electoral competition

	(1) Capital Exp. (%)	(2) Salaries as % of expenses	(3) Revenue per capita (2010)	(4) Bad garbage	(5) No toilet
Margin (00-12 Avg., Presidential)	-0.013	0.029	-0.007	0.018	0.063
	(0.047)	(0.033)	(0.042)	(0.042)	(0.045)
Full controls	Yes	Yes	Yes	Yes	Yes
Constant	131.535***	-68.128***	173.914***	97.792***	-53.536
	(37.463)	(22.170)	(22.868)	(27.406)	(33.977)
Observations	160	160	165	168	168
R^2	0.487	0.351	0.304	0.712	0.902

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 29: Summary statistics for multi-country models

Variable	Mean	Std. Dev.	Min.	Max.	N
Country level variables					
Party System Institutionalization	3.59	1.65	0.5	5.8	32,399
GDP per capita (PPP, log)	7.88	0.91	6.60	9.63	32,399
Country Size (1000s, log)	5.7	1.64	0.71	7.14	32,399
Population (m, log)	2.82	1.39	-0.69	5.2	32,399
Equator (100s miles, log)	2.5	0.81	0.26	3.48	32,399
Poverty headcount ratio	43.97	13.28	19.3	75.3	31,199
Parliamentary	0.11	0.31	0	1	32,399
Individual level variables					
Gift for Vote: Often	0.03	0.16	0	1	31,770
Gift for Vote: Never	0.86	0.35	0	1	31,770
Age (Centered)	0	14.85	-19.29	67.71	32,062
Age (Sqrd)	0	336.34	-220.41	4364.04	32,062
Economic Conditions	2.89	0.98	1	5	31,349
Gender (1 = Female)	0.5	0.5	0	1	32,399
Full-Time Employed	0.23	0.42	0	1	32,283
Some Secondary School	0.2	0.4	0	1	32,338

Note: Data for individual level variables come from Afrobarometer Round 5 (2010-2012). *Gift for vote* asked "During the last national election, how often, if ever did a candidate or someone from a political party offer you something, like food or a gift or money, in return for your vote?" *Economic conditions* asked "In general, how do you rate your living conditions compared to those of others?"
Party System Institutionalization comes from (Riedl, 2014), with higher numbers indicating greater institutionalization. Data on GDP, country size, population, and poverty come from the World Bank Development Indicators. *Equator* is the log distance in hundreds of miles from the capital city to the equator. *Parliamentary* is coded 1 if a parliamentary system and 0 if (semi)presidential, using (Weghorst & Bernhard, 2014).

Table 30: Multilevel models for vote buying across Africa

	(1) Gift for Vote: Often	(2) Gift for Vote: Never
Country level predictors		
Party System Institutionalization	-0.224***	0.249***
	(0.066)	(0.082)
Parliamentary	-1.022	0.367
	(1.026)	(0.904)
GDP per capita (PPP, log)	-0.222	0.210
	(0.162)	(0.149)
Country Size (1000s, log)	0.111	0.129
	(0.217)	(0.209)
Population (m, log)	-0.039	-0.289
	(0.285)	(0.267)
Equator (100s miles, log)	-0.482***	0.420***
	(0.087)	(0.139)
Poverty headcount ratio	-0.010	0.010
	(0.013)	(0.013)
Individual level predictors		
Age (Centered)	0.005	-0.003
	(0.005)	(0.002)
Age (Sqrd)	-0.000	0.000***
	(0.000)	(0.000)
Economic Conditions	0.011	0.057
	(0.108)	(0.042)
Gender (1 = Female)	-0.419***	0.226***
	(0.101)	(0.056)
Full-Time Employed	-0.192**	0.209***
	(0.082)	(0.077)
Some Secondary School	0.256***	-0.059
	(0.089)	(0.041)
Constant	-0.007	-2.346*
	(1.454)	(1.373)
Variance component		
Constant	0.234	0.321**
	(0.184)	(0.142)
Countries	19	19
Respondents	29,161	29,161

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Country weighting used

Figure A3: Predicted probability of reporting "Vote Buying: Never"
Greater predicted probability of reporting "Vote Buying: Never" in more highly institutionalized party systems