

# Attribution and Accountability: Voting for Roads in Ghana

## Supplementary Material

Robin Harding

University of Rochester

[robin.harding@rochester.edu](mailto:robin.harding@rochester.edu)

January 2015

# Appendix

## A Summary Statistics

Table 1: Summary Statistics

	Obs.	Mean	Std. Dev.	Min.	Max.
<b>Electoral Area Level Data</b>					
<i>Dependent Variable:</i>					
Incumbent vote share change	438	-4.006	6.298	-30.692	35.825
<i>Independent Variables:</i>					
Average road condition change (3km)	216	12.443	30.388	-59	95
Average road condition change (4km)	251	12.941	29.923	-59	95
Average road condition change (5km)	280	12.433	29.008	-59	95
Average road condition change (6km)	313	12.471	28.183	-59	95
Average road condition change (7km)	344	12.702	28.971	-59	95
Closest road condition change	438	12.011	33.3	-96	97
<b>Constituency Level Data</b>					
<i>Independent Variables:</i>					
Wealth change	256	-0.207	0.631	-1.834	1.002
Water change	256	-158.63	246.81	-18.03	988.11
Electricity change	256	0.040	0.284	-0.654	0.937
<b>District Level Data</b>					
<i>Dependent Variable:</i>					
Incumbent vote share change, by district	109	-2.705	6.575	-22.966	14.543
<i>Independent Variables:</i>					
Primary schools	109	12.376	17.747	-44	88
Classrooms per school	109	0.628	0.697	-0.811	5.029
Schools with toilets	109	-0.094	0.279	-0.651	0.777
School with drinking water	109	0.252	0.215	-0.216	0.779
Pupil-teacher ratio	109	-1.958	8.709	-48.813	7.887
Books per pupil	109	0.397	0.321	-0.699	1.818
Seats per pupil	109	-0.0004	0.219	-0.7	1.005
Average road condition change, by district	109	9.670	21.769	-96	77.75

## B Further Results

### B.1 Distribution of Road Maintenance

Table 2 reports results from constituency and district level estimates of the distribution of road maintenance in Ghana between 2004 and 2007. The dependent variable is the average change in road condition scores across all roads within a constituency or district between 2004 and 2007. The explanatory variables are the average condition scores in 2004, the share of the vote won by the incumbent party (the NPP) in 2004, the margin of victory in 2004 (calculated as the difference in vote shares between the top two vote-winning parties in each constituency), and the proportion of the population in each district belonging to the Akan group, as reported in the 2000 Ghanaian Population and Housing Census. Estimates including the share of the population belonging to the Akan group are run at the district level because the census data is only available by district, not by constituency. Standard errors are clustered by region.

Table 2: Estimates of Changes to Road Condition Scores

	(1)	(2)	(3)	(4)	(5)	(6)
Average road condition, 2004	-0.427** (0.161)	-0.436*** (0.130)	-0.441** (0.147)	-0.438*** (0.124)	-0.758*** (0.132)	-0.692*** (0.145)
Incumbent vote share, 2004	-0.012 (0.108)	-0.095 (0.153)				
Margin of victory, 2004			0.090 (0.087)	0.005 (0.109)		
Akan % of population					6.188 (5.195)	12.675 (7.214)
Region fixed effects	No	Yes	No	Yes	No	Yes
Constituencies	230	230	230	230		
Districts					110	110

*Note:* Dependent variable is change in average road conditions between 2004 and 2007. Standard errors in parentheses, clustered by region. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01.

Different theories of distributive politics suggest that patronage is more likely to be targeted at core or swing voters. To rule out any correlation between road improvements and patronage targeted

to core supporters, I use constituency level data to investigate whether changes to road conditions are correlated with the share of the vote won by the incumbent in 2004. If road improvements were related to patronage we would expect to see a positive correlation with the incumbents 2004 vote share, but the results show no statistically significant relationship between these two variables. In addition, since ethnicity is widely recognized as an important determinant of voter support in Ghana, this represents a possible alternative measure of core supporters. Therefore I use the proportion of the local population belonging to the Akan group, which is widely viewed as the incumbent party's ethnic stronghold, as another measure of core supporters. As the results in Table 2 show, there is no significant relationship between road improvements and core supporters when measured in this way either.

In addition, because alternative theories suggest that patronage is more likely to be targeted at swing voters, I also look at whether road improvements are correlated with the margin of victory in 2004. In this instance, if road improvements were related to patronage we should expect to see a negative correlation between changes to road conditions and the margin of victory, because more patronage should flow to areas where this margin is smaller. However, as with the measures of core supporters, I again find no significant relationship between these variables. Instead, in all of these estimates the only significant correlate of changes to road conditions is the prior condition of the roads, which provides further support for the claim that the distribution of resources for the maintenance of trunk roads in Ghana is programmatic. Therefore the evidence suggests that road quality improvements are not correlated with patronage, at least insofar as how we would expect patronage to be distributed, given theoretical work in this area.

## **B.2 Controlling for Outliers**

Table 3 reports results from first difference estimates of Incumbent vote share change, using robust and quantile regression techniques. As discussed in Section 5.2, these estimates are resistant to the effects of outliers. Robust regression uses iteratively reweighted least squares, which downweights outliers, thereby ensuring that they do not drive the results. Quantile regression estimates

coefficients by minimizing absolute deviations from the median, which as a measure of central tendency is more resistant to outliers than the mean. The explanatory variable is Average road condition change within a 5km radius of the electoral area. Columns (1) and (2) contain the results for the robust regressions, and columns (3) and (4) contain the results for the quantile regressions. The coefficient on the Average road condition change variable is significant at standard levels across all specifications.

Table 3: Robust and Quantile Regression Estimates of Changes to Incumbent Vote Shares

	Robust Regression		Quantile Regression	
	(1)	(2)	(3)	(4)
Average road condition change (5km)	0.032*** (0.007)	0.016* (0.009)	0.036*** (0.008)	0.022** (0.011)
District-specific time trends	No	Yes	No	Yes
Electoral areas	280	280	280	280
Districts	51	51	51	51

*Note:* Dependent variable is Incumbent vote share change. Standard errors in parentheses. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01.

### B.3 Analysis of Missing Data

Table 4 shows balance statistics for the variables of interest across constituencies included in the sample where election results were and were not available at the polling station level. This analysis is undertaken at the constituency level so that we can compare changes in vote shares across areas where election results were and were not available at the polling station level. Importantly, the results in Table 9 show that for the key variables of interest – changes in road conditions and incumbent vote shares – there are no significant differences across constituencies where election results were and were not available at the polling station level. There is also no significant difference in the margin of victory in 2004 across constituencies where election results were and were not available.

The only observable characteristic for which there is a significant difference across constituencies where election results were and were not available is for the incumbent’s vote share in 2004, which

Table 4: Balance Statistics For Constituencies With and Without Missing Data

	No Missing	Missing	Difference	T statistic
Mean road condition score, 2004	55.80	56.23	-0.43	-0.09
Mean road condition change, 2004-07	7.37	11.72	-4.35	-0.99
Incumbent vote %, 2004	57.80	46.69	11.11	3.18
Incumbent vote %, 2004 (w/o Ashanti)	46.77	46.21	0.56	0.15
Margin of victory, 2004	36.79	32.88	3.91	1.02
Incumbent vote % change, 2004-08	-4.19	-2.75	-1.12	-1.28
Constituencies	65	71	136	136

*Note:* T statistics from two-sample t tests with equal variances.

was higher in areas for which there was no missing data. This difference is driven by the exceptionally high yield of data in the Ashanti Region, the stronghold of the incumbent NPP party, where election results at the polling station level were available for 24 of the 25 constituencies in the sample. Data availability averaged around 50% across the remaining regions, apart from in Upper West, where none of the 2004 election results were available at the polling station level. As Table 4 demonstrates, if Ashanti region is omitted there is no longer a significant difference in incumbent vote share across constituencies where election results were and were not available at the polling station level. This would be a concern if there were some reason to think that the propensity of voters to hold the incumbent accountable for road conditions were greater in areas where the incumbent's prior share of the vote were higher. Although I can think of no plausible reason why this should be the case, the main findings reported in the paper are substantively unaffected when the Ashanti region is omitted from the analysis (although doing so reduces statistical power).

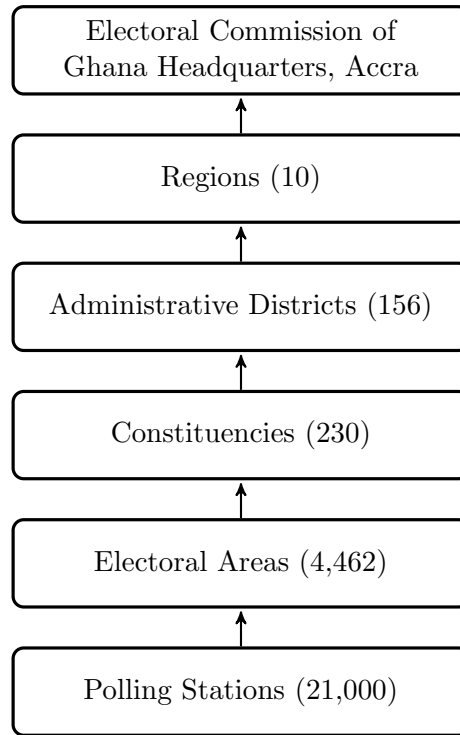
## C Data Appendix

### C.1 Election Results

As Figure 1 illustrates, elections in Ghana are administered via a nested structure. Ballots are cast and counted at Polling Stations, which are nested within electoral areas for administrative

purposes. The counted results are taken to Constituency Collation Centers, and once collated the constituency results are sent to the relevant District Electoral Officers for onward transmission to the Regional Offices of the Electoral Commission. The Regional Offices then communicate the results to the Commission's headquarters in Accra.

Figure 1: Organization of Electoral Administration in Ghana



*Note:* Elections in Ghana are administered via a nested structure, whereby each voter casts a ballot at one of the 21,000 polling stations, which are nested within 4,462 electoral areas (EAs), themselves nested within the 230 constituencies. The constituencies are gathered within 156 administrative districts, grouped within the country's ten regions. All of these are under the authority of the Electoral Commission's headquarters, based in Accra.

Results from the 2004 and 2008 Ghanaian Presidential elections were collected in person from various offices of the Electoral Commission throughout Ghana. The polling station level results for the 2008 elections are held centrally at the head office of the Electoral Commission in Accra, so it was possible to collect these from there. Since the 2004 results were collated at the district level, with just the aggregated constituency level results being sent to the head office, it was necessary to

travel to each of the districts in the sample, to collect the results in person. While a few District Officers had electronic versions of the polling station results, most of the results were kept only in paper format. Where this was the case, I took digital photographs of the results.

My analysis is based on polling station level data from a sample of electoral areas, because legal issues surrounding the release of the 2008 polling station results meant that the Electoral Commission was unwilling to release complete data for any constituency. Furthermore, collecting the 2004 returns from all 156 districts was not feasible. Therefore I drew up a pseudo-random sample of 700 electoral areas throughout the country, for which the Electoral Commission agreed that I could collect data. It is important to note that I drew up this sample myself, it was not dictated to me by the Electoral Commission. I am very grateful to them for agreeing to let me collect this data, and assisting me so generously in doing so. In an attempt to achieve a nationally representative sample, I made use of the sampling frame from the 2003 Demographic and Health Survey (DHS), which is itself a nationally representative household survey. To do this I used the GPS coordinates for the clusters sampled by the DHS to locate each cluster in an electoral area, which gave me a sample of 350 electoral areas. I then boosted this sample by randomly selecting additional electoral areas from each of the constituencies represented in the DHS sample, equal to double the number of electoral areas already selected from each constituency, giving me a final sample of 700 electoral areas.

In many districts the 2004 results have been lost or are incomplete, meaning that not all of the 2004 data was available (no data at all was available from Upper West Region). The polling station is the smallest unit, but there is no spatial data with which to geographically locate polling stations with sufficient precision, which is necessary in order to link them to data concerning local road conditions. The record of administrative units available from the Ghanaian Electoral Commission (EC) lists the localities that are contained in each electoral area. Using this list I located the electoral areas geographically, by matching them to a spatial database of 20,500 towns in Ghana (I am grateful to the Center for Remote Sensing and Geographic Information Services (CERSGIS) at the University of Ghana, Legon, for generously providing this database). After identifying all



of the towns that could be matched to a particular electoral area, I located the electoral area at their centroid. It was not possible to locate all of the electoral areas in the sample with confidence, meaning that some drop out of the sample. This left me with a dataset containing election results over time from 438 electoral areas (in 53 districts, across 9 regions).

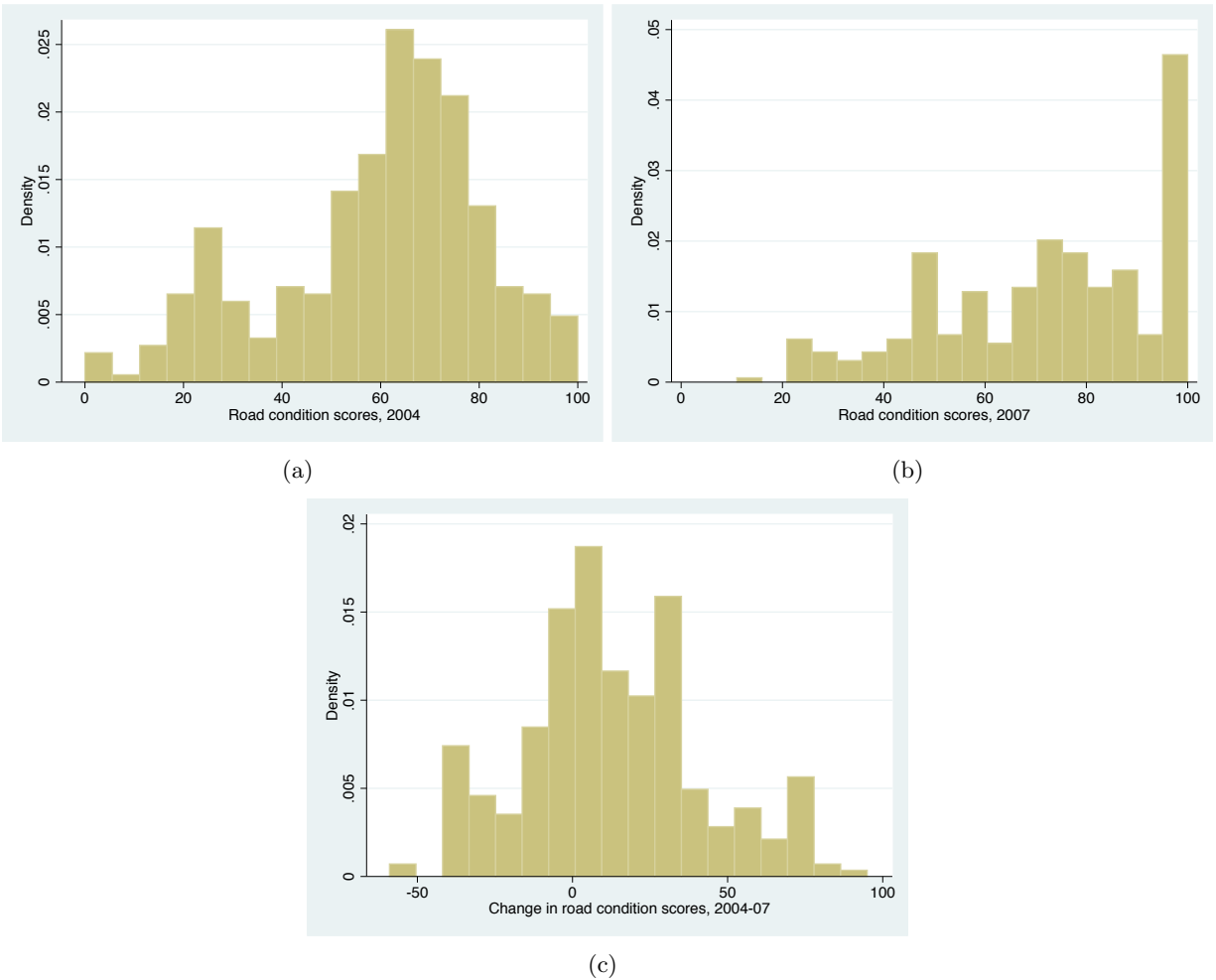
## C.2 Road Conditions

The maintenance and development of Ghana’s trunk road network is managed by the Ghana Highways Authority (GHA), which carries out an annual road condition survey of the entire trunk road network. Of the roads that are not part of the trunk road network, those within designated metropolitan areas are classified as urban roads, and those outside these designated areas are rural feeder roads. These are managed by the Department of Urban Roads and the Department of Rural Feeder Roads, respectively. The trunk road network consists of more than 2,400 segments, and between March and April of each year every segment is rated for distress and roughness. Distress is measured via a “Windshield” visual road condition survey. In this survey, raters travelling in a slow moving vehicle (at about 30km/h) observe the road surface for distresses, and determine their severity and extent based on guidelines set by the GHA. The raters also disembark from the vehicle and inspect distressed sections on foot in order to get more detailed measurements of slight distresses (e.g. cracks), which are not visible from the moving vehicle. Roughness is measured using a RIDEMATE device, installed in a vehicle, which measures the severity of bumpiness in a given road segment.

I geo-coded the annual data on road conditions into a digital map of the trunk road network, giving measures of changes to road conditions over time, throughout the country. I then linked this to election results for the sampled electoral areas by constructing 5km buffers around the centroid of each electoral area, and then taking the average of the condition scores across all segments intersecting each electoral area. The 2008 survey was carried out in May and June, during the rainy season, which creates problems of comparability with other years. Therefore the analysis focuses on changes in road conditions between 2004 and 2007. Figure 2 presents plots of the distribution of

road condition scores in 2004 and 2007, and of the change in condition scores between these years.

Figure 2: Distribution of Road Condition Scores



*Note:* Graphs show the distribution of road condition scores. Panel (a) shows the scores in 2004, panel (b) shows the scores in 2007, and panel (c) shows the change in condition scores between 2004 and 2007.

### C.3 Demographic and Health Survey Data

The estimates in Section 5.2 use household wealth data from the 2003 and 2008 DHS surveys. The DHS uses a standard Household Schedule survey instrument, which is available at [www.measuredhs.com/aboutsurveys/dhs/questionnaires.cfm](http://www.measuredhs.com/aboutsurveys/dhs/questionnaires.cfm). The DHS surveys contain an index of household wealth

that is constructed using factor analysis of responses to questions regarding ownership of a variety of household assets. Since the DHS uses a nationally representative sample, not all electoral areas are covered in the data, and even fewer are covered in both the 2003 and 2008 DHS datasets. Therefore in order to broaden the coverage of the data as far as possible, I constructed constituency level measures of wealth by averaging the DHS data across the sample clusters within each constituency.

#### **C.4 School Census Data**

The data on education inputs used in Section 5.3 comes from the annual schools census carried out by the Ministry of Education. District level data from 2001 to the present can be downloaded from [www.moess.gov.gh/emis.html](http://www.moess.gov.gh/emis.html). The census collects data and information on number of institutions, enrollment, teachers, classroom, textbooks and examination results, at the Basic, Secondary and Higher education levels, which is comparable over time.