Supporting Information for "Does Participation Reinforce Patronage? Policy Preferences, Turnout, and Class in Urban Ghana"
(Online Appendix)

Contents

1 Data

1.1 Sampling Procedure

Parliamentary constituencies in urban Greater Accra were stratified by wealth (top or bottom half of the distribution on the same wealth index in the paper), ethnic diversity (top or bottom half of the distribution on ethnic fractionalization), and 2012 presidential vote share (as NDC or NPP stronghold, or as competitive). 10 were randomly selected, with the number per stratum determined by the proportion of the total population in the stratum and the selection probabilities within each weighted by each constituency's population. The parliamentary constituencies are Ayawaso East, Ayawaso North, Ablekuma North, Ablekuma Central, and Okaikwei Central, within the official city of Accra, and Weija / Gbawe, Bortianor-Ngleshie Amanfro, La Dadekotopon, Ledzokuku, and Krowor, in the surrounding metropolitan area.

Within each constituency, 5 census enumeration areas were sampled with replacement, again stratifying by ethnic diversity and wealth, with the number chosen per stratum proportional to the stratum's share of the constituency population and weighting the selection probability of enumeration areas within strata by population. Only 4 enumeration areas were selected in two constituencies that are significantly smaller than the others (Awayaso East and Ayawaso North). Finally, selected enumeration areas were projected on the GIS census map and random geo-coded starting points for enumerators were chosen within each tract. These starting points were projected onto Google Maps for the enumerators, who began sampling respondents upon reaching each sampled point. This final step ensured that the random walks did not only begin in the more commercial part of each neighborhood, as would be the case if using polling stations or the community center as each starting point.

Interviews were conducted in four languages (English, Akan/Twi, Ga, or Ewe) by enumerators using smartphones. To select 21 respondents around each starting point, the smartphone gave each enumerator a new random direction and number of houses to count off to recruit each respondent. The first walk began at the start point and then the enumerators continued each new walk from the previous respondent's home. Within households, the phone assigned a gender to be selected (alternating by interview) and randomly selected a specific person of that gender after ranking household members by age. Enumerators conducted "call backs" for respondents who were initially unavailable and otherwise sampled replacements (from new households) via the same walk procedure. 40% of interviews were on weekends and holidays so employed and wealthy respondents were more likely to be available.

1.2 Coding Rules for Preferences

Responses to the question about preferences were blind-coded, with all respondent information removed except a ID number. Responses were coded on two dimensions: the category of preference and the specific topic or issue. In terms of the category, responses were coded as either being universalistic or particularistic and then as private or club goods within particularistic preferences. Universalistic preferences were those that could only be satisfied by a public policy that necessarily would affect many other people, not only the respondent or the respondent's immediate neighborhood. Club goods were preferences that could be satisfied by providing something to a specific neighborhood or community. Private goods were preferences that could be satisfied by providing something directly to an individual, especially the respondent.

The goal is to a provide a conservative lower bound on the extent of universalistic demands and upper bound on demands for particularistic goods. Responses were also coded for the substantive topics within these categories. These topics were adapted directly from the coding categories for a similar question on the 3rd Round of the Afrobarometer surveys (Question #63), also analyzed in Lieberman and McClendon (2013).[1]

There were several ambiguities in coding the broader categories of preferences. First, demands for club goods in which the respondent specifically specified that the beneficiaries be someone other than herself or community are coded as universalistic. Examples include "Government has to extend the electricity in a way that people in the rural areas could also have access to it" and "Put up education facilities in the rural areas." Coming from urban respondents, these are statements in support of pro-rural public policies and were coded as universalistic. But when the respondent did not specify a recipient, preferences for club goods were always coded as particularistic, consistent with this providing an upper bound on particularistic preferences. For example, "tar bad roads" would be coded as particularistic and a club good, in the same way that the response "tar bad roads in my area" would be coded. But "tar bad roads in rural areas" would be coded as universalistic, coming from an urban respondent.

Second, as discussed in the text, respondents asking for support with education expenses could have been making two different types of demands – those asking for direct assistance for themselves or families, and those demanding the NPP's national free secondary education policy, which was a central element of its 2012 campaign platform, as described in the main text. The first preference is private and particularistic and can be satisfied by patronage to a specific voter. Indeed, local politicians in Ghana describe support for school tuition as one of the main private goods they distribute to voters in clientelistic relationships. But the second preference is universalistic and regards a major and very salient national public policy proposal in Ghanaian political discourse. To separate these preferences, any statements that directly and specifically mentioned making secondary education free (e.g., "free shs education" or "make shs free", where SHS is "senior high school") were coded as universalistic, as these likely refer directly to the NPP's programmatic

[1] The full list of topics that respondents raised are: 1. rates and taxes; 2. education; 3. corruption; 4. petroleum; 5. unemployment; 6. health; 7. water supply; 8. economy; 9. infrastructure and roads; 10. housing supply; 11. sanitation; 12. support to local business; 13. governance; 14. poverty; 15. loans; 16. electricity; 17. crime; 18. social welfare; 19. farming; 20. political divisions / conflict ; 21. social problems; 22. public sector wages; 23. food shortages; 24. land; 25. flooding; 26. transportation; 27. orphans.

policy proposal during the 2012 election to make SHS free nationwide. But statements that made vaguer demands for support for education expenses were coded as particularistic (e.g., “reduce cost of education” or “make school fees affordable”), as these preferences *could* still be satisfied by targeted assistance to that particular respondent. I show below that the results are robust to this decision.

Third, many respondents said that they wanted the government to improve employment. These preferences were coded as universalistic only if the respondent specifically said that they wanted jobs created for everyone (e.g., “create jobs for all” or “create more jobs in Ghana so that living standard will improve”). Any other statement that was vague about whether or not the respondent herself wanted the new job was coded as particularistic (e.g., “create more jobs” or “provision of jobs”). This preference could still potentially be satisfied by a patronage job for that respondent, her family, or those in her immediate community.

1.3 Decomposing the Education and Employment Index

The main index used to measure socio-economic class is based on three characteristics: whether respondents have at least some secondary education (or higher), whether respondents are literate in English, and whether respondents are employed in the formal sector. 408 of the 995 respondents (41%) have none of these three characteristics. 117 have all three, 296 have two of three, and 174 have just one characteristic. Overall, 512 (51%) respondents have at least some secondary education, of which 119 have tertiary education. 448 (45%) respondents are literate in English. 157 (16%) explicitly report being employed in the formal (including public) sector, while 271 (27%) say they have jobs that pay regular, pre-determined salaries, which is another indicator of potential formal sector employment. The first, more restrictive, indicator is used to classify the education/employment index. Table SI.1 lists the most common occupations reported by the respondents in the formal (including public) sector after coding occupations into the 18 categories listed below. The table shows that, in general, these occupational categories line up closely with what would be expected for formal sector employment.

In addition, the correlation between the education/employment index and the household assets index is 0.42. Among people with the highest score on the education/employment index (who have all three characteristics above), the mean value of the assets index is 0.968 with a standard deviation of 1.1, ranging from -1.095 to 2.3 (the maximum value). The overall mean on the assets index is 0, with standard deviation 1. This indicates that while there are some people with few assets being marked as middle class by the education/employment index, most people who score highly on the education/employment index also score highly on the assets index. This is consistent with the education/employment index picking up real differences in socio-economic status.

Table SI.1: Most Common Occupations within Formal Sector Employment

Category:	Example 1	Example 2	Example 3
Office worker (25%)	"office worker"	"accountant"	"sales officer"
Civil service (18%)	"civil servant"	"government worker"	"military officer"
Healthcare professional (10%)	"nurse"	"pharmacist"	"lab technician"
Teacher (10%)	"teacher"		
Factory worker / machine operator (8%)	"factory hand"	"machine operator"	"fork lift operator"
Small businessperson (4%)	"businessman"	"supplier"	"filling station owner"
Security person (4%)	"security man"		
Banker (3%)	"banker"	"credit analyst"	"bank manager"

Within each category I list the most common specific professions recorded. In order of frequency the full list of categories is: 1. office worker; 2. civil service; 3. healthcare professional; 4. teacher; 5. factory worker; 6. small businessperson; 7. security person; 8. banker; 9. engineer; 10. janitor; 11. retail worker; 12. driver; 13. food service worker; 14. religious leader (pastor); 15. professional athlete; 16. construction worker; 17. building superintendent; 18. private school proprietor.

1.4 Calculation of Neighborhood Characteristics

Neighborhood boundaries in any city are nebulous social concepts. People living in the same place often disagree about the boundaries of their neighborhood (Wong et al. 2012). Universally agreed-upon definitions of neighborhood boundaries do not exist in Greater Accra. Moreover, census enumeration area (tract) boundaries provide an inappropriate means of defining neighborhoods; as elsewhere, enumeration areas in Ghana are politically and socially irrelevant creations of the census bureau. Taking enumeration areas as the unit of observation may also suffer from the "modifiable areal unit problem" (Openshaw 1983), in which arbitrary choices about how to impose discrete boundaries on continuous geographic data will significantly bias a study's results. I thus do not define neighborhoods as census enumeration areas. I also do not define neighborhoods as a political unit much larger than actual neighborhoods, such as a district or parliamentary constituency. This would define respondents' neighborhoods as areas far larger than any meaningful qualitative definition of what a neighborhood is.

My approach finds a middle ground between these unsuitable extremes. I attempt to mitigate the risk of bias by defining local neighborhood characteristics as weighted averages of census characteristics from all census tracts around each individual survey respondent, with information closer to the respondent weighted higher, and all data outside a given radius weighted as 0. This smooths the census data over the enumeration areas rather than imposing discrete, arbitrary neighborhood boundaries and follows best practice in the sociology methods literature to measure ethnic neighborhood compositions from tract-level census data (Reardon and O'Sullivan 2004, Lee et al. 2008). This also directly replicates the procedure used for similar data (in rural areas) in Ichino and Nathan (2013).

In this approach, the spatially weighted population share of group m around a respondent at point p is $\tilde{\pi}_{pm} = \frac{\int_{q \epsilon R} \tau_{qm} \phi(p,q) dq}{\int_{q \epsilon R} \tau_q \phi(p,q) dq}$, where τ_q is the population density at point q, τ_{qm} is the population density of group m at point q, $dist(p,q)$ is the distance in kilometers from the respondent at point p to the centroid of a surrounding census enumeration area (EA) at point q, and EAs are weighted by the function $\phi(p,q) = (dist(p,q) + 0.5)^{-1}$, as in Reardon and O'Sullivan (2004), up to a maximum distance, after which all EAs are weighted as 0. This is calculated via the `seg` package in `R`. I set the maximum radius to 500 meters for the reasons discussed in the text. This also accounts for population density, with the weights a function of both distance and population density around the respondent. The median census enumeration area is 0.03 sq. km in area, smaller than the 0.79 sq. km of each radius around respondents. Because the weights approach 0 further away from the respondent, minor changes in the size of the maximum radius (e.g., to 400m or 600m) do not significantly affect these measures. All results are robust to such changes.

1.5 Calculation of Neighborhood Wealth Index

To measure differences between poorer and wealthier neighborhoods a neighborhood wealth index was calculated for the 500 meter radius around each respondent based on the census data, weighting census characteristics in the manner described above. This is calculated as the first dimension in a factor analysis of census questions on assets, education, and employment. Variables that only measure service provision, not wealth, are excluded. The index includes: % with running water (privately provided by wealthier residents via tanker or borehole); % with a flush toilet, % with electricity (available to all who can afford it), % in a single-family home (excluding informal structures); % with a computer; % adults with more than a middle school education; and % adults employed in the formal or public sectors. The index is scaled in standard deviations from the city-wide mean of 0. It is calculated over all enumeration areas in urban Greater Accra. In the survey sample, the variable ranges from -1.4 to 3, with a mean of -0.10. Higher values indicate greater wealth.

1.6 Summary Statistics for Survey Data

Summary statistics for all variables from the survey data used in the analysis, except the survey experiment (see below) are presented here.

Table SI.2: Summary Statistics for Survey Data

	mean	min	max	sd	N
Universalistic preference (0,1)	0.56	0.00	1.00	0.50	987
Club good preference (0,1)	0.64	0.00	1.00	0.48	987
Private good preference (0,1)	0.53	0.00	1.00	0.50	987
Universalistic preferences (%)	0.31	0.00	1.00	0.33	987
Voted in 2012 (0,1)	0.84	0.00	1.00	0.36	994
Minimum participator (0,1)	0.36	0.00	1.00	0.48	927
Party Member (0,1)	0.16	0.00	1.00	0.37	928
Knows Local Party Agent (0,1)	0.22	0.00	1.00	0.41	995
Has Met Assembly-member (0,1)	0.64	0.00	1.00	0.48	995
Association that discusses politics (0,1)	0.28	0.00	1.00	0.45	995
Age	35.54	18.00	93.00	13.07	995
Muslim (0,1)	0.20	0.00	1.00	0.40	995
Male (0,1)	0.50	0.00	1.00	0.50	995
Akan (0,1)	0.42	0.00	1.00	0.49	995
Ewe (0,1)	0.18	0.00	1.00	0.38	995
Northerner (0,1)	0.18	0.00	1.00	0.38	995
Ga-Dangme (0,1)	0.30	0.00	1.00	0.46	995
Assets index	0.00	-1.10	2.30	1.00	995
Education/employment index	0.00	-1.04	1.74	1.00	995
Moved for club goods (0,1)	0.11	0.00	1.00	0.31	995
Prefers Good for Home Region (0,1)	0.13	0.00	1.00	0.34	995
Percent life in Greater Accra	0.69	0.00	1.00	0.33	995
Population Growth 2000-2010 in 500m	3.30	0.20	22.17	4.31	995
% Reporting Regular Water (by cluster)	0.43	0.00	1.00	0.35	48
Main road tarred (by cluster)	0.29	0.00	1.00	0.45	48
Neigh. Wealth Index in 500m	-0.10	-1.40	3.02	0.80	995
Pop. Density per sq. km (by cluster)	28710	1650	95120	27050	48
Saw door to door campaign (0,1)	0.67	0.00	1.00	0.47	930
Saw gift distribution (0,1)	0.69	0.00	1.00	0.46	930
% Reporting Door to Door Campaign (by cluster)	0.70	0.18	0.94	0.17	48
Abs. Difference NDC v. NPP 2008 in ELA	0.19	0.00	0.39	0.11	995
Eth. Fractionalization in 500m	0.69	0.39	0.86	0.11	995
Enumerator errors (0,1)	0.12	0.00	1.00	0.33	995
Cooperative respondent (0,1)	0.90	0.00	1.00	0.30	995

Note that indicators for ethnic group membership do not sum to 1 because some respondents are from multiple ethnic groups.

2 Empirical Results and Robustness Tests

2.1 Dropping Possible "Upper Class" Respondents

Given data limitations, there may be concern that I cannot differentiate the wealthy elite from the middle class. But indicators of housing quality in the survey data indicate that the number of wealthy elites in the sample is very small. In this robustness test I my main results for class are robust to dropping the respondents most likely to be wealthy elites, indicating that results from measures that distinguish the poor and non-poor can still be reasonably interpreted as distinguishing the poor and middle class.

Enumerators noted whether respondents lived in "luxury" or "upper class" housing. This is defined as well-maintained large single-family homes with walls and security gates, barbed wire or electric fencing above the walls, and household staff, all indicative of the truly wealthy. While some of the respondents living in these homes may not actually be wealthy – for example, if household staff were sampled – this variable should correlate with a higher probability of being a wealthy elite. I replicate Columns 1 and 3 of Table 2 in the main text after dropping all of these respondents. These are the main tests in the paper where the indicator for socio-economic status significantly predicts the outcome.

Table SI.3: Universalistic Preferences, by Socio-Economic Status, Dropping Luxury Housing Residents

	1	3
Outcome:	Binary	Binary
Educ/Emply. Index	0.178^{*}	0.182^{***}
	(0.085)	(0.085)
Assets Index	-0.126	-0.125
	(0.086)	(0.088)
Pop. Change 10 Years (500m)		-0.019
		(0.026)
Neighborhood Wealth (500m)		0.252
		(0.183)
Running Water (by cluster)		0.680^{*}
		(0.341)
Paved Road (by cluster)		$0.444^{\dagger}$
		(0.254)
Pop. Density (by cluster)		0.017^{**}
		(0.006)
Individual-level Controls	Y	Y
N	952	952

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.
Replicates Columns 1 and 3 from the main text.

2.2 Binary Indicator for Middle Class, as in Thurlow et al. (2015)

In addition to the continuous education/employment index used in the text, I show that my main results for socio-economic class (see Table 2) are robust to instead using a dichotomous classification of the poor and non-poor, which given the dearth of very wealthy respondents can be interpreted as an indicator for being in the middle class.

I code a binary indicator similar to that recommended by Thurlow et al. (2015), combining data on education, formal sector employment, and housing amenities. I also continue to include the measure of English literacy, as I believe this particularly important in the Ghanaian case for indicating that someone has the skills to compete in the formal economy, even if they are not actually employed in the formal sector. The measure for housing quality follows Thurlow in using a combined indicator for whether a respondent lives somewhere with electricity, piped water, and a flush toilet. I classify respondents as being "non-poor" if they have these three characteristics for housing amenities and have at least some secondary education, and are employed in the formal sector, literate in English (which indicates potential to be employed in the formal sector in the future), or both. 160 respondents of 995 fit this definition. I find that the main results for socio-economic class in Table 2 are identical using this binary classification, with non-poor respondents more likely to make universalistic demands.

Table SI.4: Universalistic Preferences, by Socio-Economic Status and Local Need

	1	2	3	4
Outcome:	Binary	Percentage	Binary	Percentage
Indicator for "Non-Poor"	0.506^{*}	0.082^{*}	$0.479^{\dagger}$	0.077^{*}
	(0.253)	(0.036)	(0.253)	(0.036)
Assets Index	−0.152	−0.014	−0.149	−0.015
	(0.095)	(0.014)	(0.096)	(0.014)
Pop. Change 10 Years (500m)			−0.012	−0.003
			(0.026)	(0.004)
Neighborhood Wealth (500m)			0.276	0.044
			(0.178)	(0.027)
Running Water (by cluster)			$0.650^{\dagger}$	0.106^{*}
			(0.345)	(0.053)
Paved Road (by cluster)			$0.449^{\dagger}$	0.068
			(0.257)	(0.040)
Pop. Density (by cluster)			0.017^{**}	0.002^{*}
			(0.006)	(0.001)
Individual-level Controls	Y	Y	Y	Y
N	987	987	987	987

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$. Replicates Table 2 in the main text.

Table SI.5: Universalistic Preferences (Replication of Table 2)

	1	2
Outcome:	Binary	Binary
Version:	Educ/Literacy Only	No Employment Demands
Educ/Literacy Index	0.184^{*}	
	(0.092)	
Educ/Emply. Index		$0.207^{\dagger}$
		(0.113)
Assets Index	-0.104	-0.234^{*}
	(0.083)	(0.106)
Individual-level Controls	Y	Y
N	987	619

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$. Logistic regression coefficients, with intercepts partially pooled by sampling cluster, following Gelman and Hill (2007). All models otherwise identical to column 1 of Table 2 in the main text.

2.3 Table 2 Results without Employment

To ensure that the correlation between being in the middle class and having universalistic preferences is not a mechanical outcome of the fact that the definition of middle class includes whether or not a respondent is employed in the formal sector and one of the most common particularistic preferences is a demand for employment, I repeat the main model from Table 2 after either: (a) only defining middle class status based on literacy and secondary education, not employment; (b) dropping all respondents who list employment as one of their preferences and using the original definition of middle class. Results in both alternative specifications are substantively identical to those in Table 2 in the paper.

2.4 Outcome as Demand for Club Goods

I also repeat the models for Table 2 with the outcome as a binary indicator for demanding any club good (gutters, roads, etc). I find that differences in wealth and socio-economic class do not predict whether respondents want club goods, such as new roads or water supply. But the variables measuring the quality of existing service provision in the neighborhood are all strong predictors of preferences for club goods, as described in the main text. In addition, having moved into your current neighborhood because of the public services there (controlled for in all models), predicts lower demand for club goods, consistent with the discussion of this in the text.

Table SI.6: Preferences for Club Goods

	1	2
Educ/Emply. Index	-0.057	-0.055
	(0.089)	(0.089)
Assets Index	0.092	0.101
	(0.091)	(0.093)
Pop. Change 10 Years (500m)		0.075^{*}
		0.035
Neighborhood Wealth (500m)		-0.211
		(0.200)
Running Water (by cluster)		-0.938^{*}
		(0.397)
Paved Road (by cluster)		$-0.566^{\dagger}$
		(0.294)
Pop. Density (by cluster)		-0.013^{*}
		(0.006)
Moved for Club Goods	-0.570^{*}	-0.557^{*}
	(0.246)	(0.247)
Individual-level Controls	Y	Y
N	987	987

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$. Logistic regression coefficients, with intercepts partially pooled by sampling cluster, following Gelman and Hill (2007). The outcome is either a binary indicator for listing at least one club good among the responses. Other than the outcome variable, the models are the same as columns 1 and 4 of Table 2 in the main text

2.5 Robustness to Measuring Preferences after the Election; Robustness to Education Coding Decisions

The tables below simultaneously account for two robustness tests referenced in the text. I address concerns about recording preferences after the election and address concerns about coding decisions related to demands for "free secondary education."

First, because turnout cannot be measured until after the election, and preferences and turnout must be measured in the same survey, preferences are observed after the decision to vote, yet are being used as an explanatory variable. Preferences could have been influenced by factors occurring after the election. The assumption I make is that the types of people who prefer universalistic versus particularistic goods are correlated over time, such that respondents prioritizing universalistic policies after the election were also those most likely to prioritize universalistic policies before the election.[2] Importantly, if the classification of preferences has been muddled by post-hoc measurement, with some respondents mis-assigned to the wrong category (universalistic vs. particularistic), the measurement error should bias against seeing significant differences in political behavior between the two groups.

But this does not mean that the specific issues that voters raise within these broader categories of universalistic or particularistic goods will be the same over time. The clearest example of this is that the government announced large increases in water and electricity prices immediately before the survey was fielded – this likely explains the prominence of preferences for lower electricity and water prices seen in Table 1 of the main text.[3] In the tables below, I drop all respondents who would only have been coded as having universalistic preferences because of complaints about electricity or water prices and re-estimate Tables 2 and 3 of the main text. I find identical results, suggesting that the findings are not explained by a topical issue that arose after the election.

Moreover, when those with universalistic preferences did turn out to vote, they were more likely to vote for the NPP than the NDC, even controlling for differences in ethnicity and individual-level wealth. This is consistent with the NPP putting somewhat more emphasis on policy-based appeals in its campaign and especially with the NPP being the only viable opposition party that voters dissatisfied with the status quo could turn to. There could be concern, however, that universalistic preferences measured after the election are outcomes of NPP support. This is especially for those who wanted free secondary education nationwide, which was mentioned by some who spoke about education (see Table 1 in the main text, 10% of total universalistic preferences were on education). This was by far the main universalistic policy issue emphasized in the 2012 NPP platform and campaign. In another robustness test in the second column of the three tables below, I drop all respondents who would only have been coded as having universalistic preferences because they wanted the NPP's secondary education policy and re-estimate the results in Tables 2 and 3 of the main text. I find no substantive differences with the results in the main text.

Second, this latter test also shows the robustness of my results to the decision to code demands for free

[2]Similarly, I make an assumption that measurement of who is in the middle class is correlated over time.

[3]These rates are set by the government, with a single price nationally, and thus involve a universalistic policy.

Table SI.7: Universalistic Preferences, Dropping Utility Price and Education Demands

	1	2
Outcome:	Binary	Binary
Dropped Respondents:	Utility Prices	Education
Educ/Emply. Index	0.191^{*}	0.182^{*}
	(0.091)	(0.084)
Assets Index	$-0.182^{\dagger}$	-0.131
	(0.097)	(0.088)
Pop. Change 10 Years (500m)	-0.039	-0.010
	(0.030)	(0.025)
Neighborhood Wealth (500m)	0.436^{*}	0.272
	(0.198)	(0.175)
Running Water (by cluster)	0.882^{*}	0.685^{*}
	(0.372)	(0.339)
Paved Road (by cluster)	0.588^{*}	$0.420^{\dagger}$
	(0.276)	(0.252)
Pop. Density (by cluster)	0.021^{***}	0.017^{***}
	(0.006)	(0.006)
Individual-level Controls	Y	Y
N	862	961

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$. Logistic regression coefficients with intercepts partially pooled by sampling cluster, following Gelman and Hill (2007). The outcome is the binary indicator for listing at least one universalistic policy. Models are the same as Table 2 in the main text.

secondary education as a universalistic rather than a particularistic preference (see discussion above and in the main text). By dropping all respondents who would only have been coded as demanding universalistic policies based only on this issue, I show that my main results are robust to this contextual decision regarding preference classifications.

Table SI.8: Turnout and Participation, Dropping Utility Price and Education Demands

	1	2	3	4
Outcome:	Turnout	Turnout	Withdrawal	Withdrawal
Dropped Respondents:	Utility Prices	Education	Utility Prices	Education
Universalistic Preferences (binary)	$-0.389^{\dagger}$	0.381^{*}	-0.544^{**}	0.365^{*}
	(0.219)	(0.168)	(0.201)	(0.158)
Educ/Emply. Index	-0.125	0.233^{*}	-0.071	0.195^{*}
	(0.125)	(0.096)	(0.112)	(0.090)
Assets Index	0.204	-0.164	0.164	-0.116
	(0.136)	(0.100)	(0.121)	(0.094)
Neighborhood Wealth (500m)	-0.094	0.178	-0.039	0.153
	(0.215)	(0.163)	(0.196)	(0.160)
2008 Competitiveness (by ward)	-3.153^{*}	0.622	$-2.438^{\dagger}$	-0.003
	(1.492)	(1.129)	(1.367)	(1.143)
Ethnic Fractionalization (500m)	1.236	-0.458	0.852	-0.354
	(1.804)	(1.342)	(1.638)	(1.348)
Pop. Density (by cluster)	0.007	-0.002	0.009	-0.002
	(0.010)	(0.007)	(0.009)	(0.007)
Individual-level Controls	Y	Y	Y	Y
Constituency FEs	Y	Y	Y	Y
N	861	799	960	895

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$. The outcome in columns 1-2 is self-reported turnout in the 2012 presidential and parliamentary elections. The outcome in columns 3-4 is an indicator for doing only 1 or 0 of the 5 forms of participation discussed in the text. All models are otherwise the same as Table 3 in the main text.

Table SI.9: Turnout, Adjusting for Registration Rates

	1
Outcome:	Turnout
Universalistic Preferences (binary)	-0.520^{*}
	(0.260)
Educ/Emply. Index	-0.109
	(0.143)
Assets Index	0.222
	(0.159)
Neighborhood Wealth (500m)	0.095
	(0.256)
2008 Competitiveness (by ward)	-3.782^{*}
	(1.857)
Ethnic Fractionalization (500m)	0.320
	(2.070)
Pop. Density (by cluster)	0.019
	(0.012)
Individual-level Controls	Y
Constituency FEs	Y
N	914

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$. The outcome is self-reported turnout in the 2012 presidential and parliamentary elections. Model is the same as Column 1 of Table 3 in the main text.

2.6 Turnout Results Adjusting for Registration Rates

There could be concern that the results for turnout are affected by possible class-based differences in underlying voter registration rates. To account for this, I drop all respondents who reported on the survey that they were not registered to vote and re-estimate column 1 of Table 3 in the main text. I find that the results for turnout are unaffected by accounting for registration rates.

2.7 Table for Interaction of Demands and Mobilization

This is the corresponding regression table for the model used to simulate differences in turnout between those with and without universalistic preferences at different rates of reported door to door mobilization.

Table SI.10: Turnout and Mobilization

	1
Outcome:	Turnout
Universalistic Preferences (percentage)	-2.407^{*}
	(1.105)
Door to Door Mobilization % (by cluster)	-0.508
	(1.124)
*Universalistic * Door to Door*	$2.852^{\dagger}$
	(1.584)
Educ/Emply. Index	-0.071
	(0.113)
Assets Index	0.156
	(0.119)
Neighborhood Wealth (500m)	0.027
	(0.220)
2008 Competitiveness (by ward)	-2.174
	(1.422)
Ethnic Fractionalization (500m)	1.130
	(1.641)
Pop. Density (by cluster)	0.009
	(0.009)
Individual-level Controls	Y
Constituency FEs	Y
N	986

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$. The outcome in columns 1-4 is self-reported turnout in the 2012 presidential and parliamentary elections. Logistic regression coefficients, with intercepts partially pooled by sampling cluster, following Gelman and Hill (2007). All other modeling details are identical to Table 3 in the main text.

Table SI.11: Treatment Values in the Survey Experiment

Treatment:	Value 1	Value 2	Value 3	Value 4
Ethnicity:				
AKAN	James Prempeh	Emmanuel Owusu Ansah	–	–
EWE	Joseph Dzorkpe	John Dodzi	–	–
GA	Alfred Nii Tawiah	Richard Laryea	–	–
NORTHERN	Isaac Yakubu	Amadu Muntari	–	–
Promise:				
PUBLIC	Water and fuel prices	National water production	–	–
CLUB	Roads in the constituency	Classrooms in the constituency	–	–
PRIVATE	Jobs for youth	Scholarships to families	–	–
Background text:	Doctor	Lecturer	Lawyer	Businessman

2.8 Experimental Vignettes

Respondents were read two pairs of vignettes about two hypothetical candidates each. After being asked to say which of the two candidates they would rather vote for (analyzed elsewhere), respondents were asked the follow-up question about credibility of one of the two candidates in each pair: "Do you think a politician like [NAME] will actually deliver on a promise like [PROMISE]?" (with the treatments inserted). The candidate that this follow-up question was asked about was randomly chosen from each pair. Each respondent thus answered this question on credibility twice, with two different conditions for NAME and PROMISE.

The vignettes about candidates varied on three dimensions: the name (and thus ethnicity) of the candidate, the good he promised to deliver, and his background. Table SI.11 lists the possible values of each treatment. Inserting example treatments, the vignettes took the form:[4]

> [AKAN 1, PUBLIC GOOD 1, LECTURER]: "Candidate A is named JAMES PREMPEH. He is a lecturer and teacher who graduated from KNUST.[5] He lives in the constituency here. If elected, he is promising TO LOBBY FOR KEEPING THE PRICE OF FUEL AND UTILITIES LOW so that everyone IN GHANA can continue to afford fuel and electricity. With your support, JAMES PREMPEH believes he can bring about a transformation in the development of this community."[6]

[4] All prompts were translated into three local languages (Akan, Ewe, Ga), or instead read aloud in English for respondents who preferred the interview in English. Enumerators could select which translation they wished to have appear on the smartphone before reading aloud the prompts.

[5] This is one of the three major national universities in Ghana – the Kwame Nkrumah University of Science and Technology.

[6] The "background" treatment also includes the "filler" text of the vignette. So for doctors, the final sentence was always "[NAME] wants you to wants you to vote for him so he can improve the lives of people in this community." For lecturers it was: "With your support, [NAME] believes he can bring about a transformation in the development of this community", etc. This ensured

> [EWE 1, PRIVATE 1, DOCTOR]: "Candidate B is named JOSEPH DZORKPE. He is a doctor who lives in this constituency and is running for parliament. He received his medical training at the University of Ghana Medical School.[7] In return for your support, he is promising TO FIND JOBS FOR SOME OF THE YOUTH in the constituency. JOSEPH DZORKPE wants you to vote for him so he can improve the lives of people in this community."

In each pair of candidates, respondents always received one candidate from their own ethnic group (based on their response to a question about their own ethnicity earlier on the survey) and the other candidate's ethnicity was then selected at random.[8] The specific names within these ethnicities were also selected at random. In the second pair of vignettes, respondents received the other name from their own ethnic group and a name from one of the two ethnic groups that they had not received in the first vignette, again selected at random. The four backgrounds were randomly allocated to the four candidates, without replacement, such that each background treatment occurred once. In all pairs the candidates promised the same category of good, with one promising the first example and the other promising the second example, again assigned at random. The category of the first pair was selected at random with equal probability from the three categories (PUBLIC/UNIVERSALISTIC, CLUB, PRIVATE) and then the category of second pair selected at random from the remaining two. The order of each specific example within these categories (PRIVATE GOOD 1 vs. PRIVATE GOOD 2) was also randomized with equal probability, as was the order of all the other treatments within the pairs (names, backgrounds).

For example, a single Akan respondent would receive treatments such as: PAIR 1: [AKAN NAME 1, PUBLIC GOOD 1, LAWYER] vs. [NORTHERN NAME 2, PUBLIC GOOD 2, BUSINESSMAN]; PAIR 2: [AKAN NAME 2, CLUB GOOD 1, DOCTOR] vs. [EWE NAME 2, CLUB GOOD 2, LECTURER]. And then the follow-up question about credibility could have been: [NORTHERN NAME 2, PUBLIC GOOD 2, BUSINESSMAN] for the first pair, and [AKAN NAME 2, CLUB GOOD 1, DOCTOR] for the second pair. The questions on credibility would then take the form:

> "Do you think a politician like AMADU MUNTARI will actually deliver on a promise like CONSTRUCTING NEW WATER PRODUCTION FACILITIES IN GHANA?"

> "Do you think a politician like EMMANUEL OWUSU ANSAH will actually deliver on a promise like CONSTRUCTING AND TARRING MORE OF THE ROADS IN THE CONSTITUENCY?"

---

that the exact wording of the vignette was not exactly the same for each of the four candidates and varied across the ethnicity and promise conditions.

[7] This is the main national medical school.

[8] For respondents who were not members of any of these four ethnic categories all names in the experiment were assigned at random, with equal probability.

Table SI.12: Responses to Credibility Question by Co-Ethnicity and Good Promised

Promise:	Public/Universalistic	Club	Private
Co-ethnic Name:	99 “Yes” (32%) N=313	96 “Yes” (34%) N= 285	88 “Yes” (30%) N=292
Non-co-ethnic Name:	69 “Yes” (23%) N=303	90 “Yes” (28%) N=327	84 “Yes” (26%) N=327

2.9 Summary Statistics for Credibility Question

In Table SI.12 I provide summary statistics for responses to the follow-up question about credibility. These are the counts of respondents in each treatment category who answer “yes,” that they expect a politician to follow through on his promise. Treatment conditions here are defined by whether respondents’ received a co-ethnic or non-co-ethnic name and by the category of good promised. More respondents overall received the non-co-ethnic name treatment because respondents from minor ethnic groups always received non-co-ethnic names.

2.10 Balance for Credibility Question

In Table SI.13 I provide balance statistics for differences in means for key covariates between the respondents receiving the credibility question in the experiment about a co-ethnic or about a non-co-ethnic. These statistics are listed separately by the type of good referenced in the question. Due to small sample sizes, balance remains imperfect after the randomization, especially for the club goods promises. Because of this, all analysis of the survey experiment the paper includes co-variates as controls to adjust for remaining imbalance across these conditions.

Table SI.13: Differences in Means, Co-Ethnic Name Treatment

Variable:	difference in means	p-value
Promise: Universalistic/Public		
Universalistic preference (0,1)	0.037	0.235
Male (0,1)	-0.009	0.778
Some Secondary Education (0,1)	0.003	0.926
Formal Sector Employment (0,1)	0.016	0.493
English LIteracy	-0.023	0.466
Education/employment index	-0.004	0.944
Assets index	-0.056	0.371
Moved for club goods (0,1)	0.014	0.467
Age	0.281	0.731
Neigh. Wealth index	0.033	0.527
Eth. Fractionalization	-0.017	0.016
Population Density	-1.773	0.318
Promise: Club Good		
Universalistic preference (0,1)	0.086	0.007
Male (0,1)	-0.009	0.780
Some Secondary Education (0,1)	0.086	0.007
Formal Sector Employment (0,1)	0.063	0.007
English LIteracy	0.094	0.003
Education/employment index	0.226	0.000
Assets index	0.046	0.475
Moved for club goods (0,1)	0.039	0.041
Age	-0.119	0.887
Neigh. Wealth index	0.101	0.042
Eth. Fractionalization	-0.002	0.747
Population Density	-1.211	0.485
Promise: Private Good		
Universalistic preference (0,1)	0.063	0.047
Male (0,1)	-0.087	0.006
Some Secondary Education (0,1)	-0.090	0.005
Formal Sector Employment (0,1)	-0.017	0.460
English LIteracy	-0.057	0.070
Education/employment index	-0.152	0.016
Assets index	0.024	0.708
Moved for club goods (0,1)	-0.009	0.629
Age	1.034	0.217
Neigh. Wealth index	-0.007	0.885
Eth. Fractionalization	-0.008	0.216
Population Density	-0.415	0.806

2.11 Experimental Results Pooled Across Goods

I confirm that universalistic preferences predict less trust in politicians' promises – and the respondents are less trusting of promises from non-co-ethnics – in a single model that pools across the types of goods being promised (universalistic/public, club, private). This model involves double counting individual respondents, however, as each respondent answered two versions of the credibility question, one each about two of the three types of promises.

Table SI.14: Survey Experiment: Credibility of MPs' Campaign Promises

	1
Promised Good:	All
Universalistic Preferences (binary)	-0.374^{**}
	(0.114)
Co-Ethnic Candidate	0.409^{***}
	(0.113)
Club Promise 2: New Classrooms	0.039
	(0.186)
Public Promise 1: New Classrooms	-0.197
	(0.186)
Public Promise 2: Low Utility Prices	-0.070
	(0.186)
Private Promise 1: Jobs	-0.071
	(0.186)
Private Promise 2: Scholarships	-0.080
	(0.185)
Individual-level Controls	Y
Neighborhood-level Controls	Y
Name and Background Controls	Y
N	1835

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$. The outcome is whether a respondent believes the MP in the vignette will actually deliver the cued good after the election. Logistic regression with intercepts partially pooled by sampling cluster, following Gelman and Hill (2007). Includes the same individual-level and neighborhood-level controls the main text, as well as controls for each additional treatment condition (name, background, etc).

References

Ichino, Nahomi and Noah L. Nathan. 2013. "Crossing the Line: Local Ethnic Geography and Voting in Ghana." *American Political Science Review* 107(2):344–361.

Lee, Barrett A., Glenn Firebaugh, Stephen A. Matthews, Sean F. Readon, Chad R. Farrell and David O'Sullivan. 2008. "Beyond the Census Tract: Patterns and Determinants of Racial Segregation at Multiple Scales." *American Sociological Review* 73(5):766–791.

Openshaw, Stan. 1983. The Modifiable Areal Unit Problem. In *Concepts and Techniques in Modern Geography*. Norwich, UK: Geo Books.

Reardon, Sean F. and David O'Sullivan. 2004. "Measures of Spatial Segregation." *Sociological Methodolgy* 34:121–162.

Wong, Cara, Jake Bowers, Tarah Williams and Katherine Drake Simmons. 2012. "Bringing the Person Back In: Boundaries, Perceptions, and the Measurement of Racial Context." *Journal of Politics* 74(4):1153–1170.