Online Appendix for "Vote-Buying, Anti-Corruption Campaigns, and Identity in African Elections"

Appendix A: Additional Information on the Experimental Outcomes

This section reports the full regression table associated with Figure 1 in the main manuscript as well as two other specifications to check for robustness. (See Table A1). It also provides distributions of our dependent variable alone and by our main independent variable: candidate appeal.

Table A1: AMCEs of Profile Features on Likelihood of Vote

	OLS (0-1 Scale)	Ordinal Logit (1-4 Scale)	OLS (0-1 Scale)
Appeal (Base: Pro-Immediate Goods)			
Pro-Community Goods	0.2793***	1.1836***	0.2747***
	(0.0607)	(0.2560)	(0.0589)
Pro-Future Goods	0.1836**	0.7564***	0.1751**
	(0.0508)	(0.1955)	(0.0505)
Anti-Future Goods	0.0952+	0.3988+	0.0984+
	(0.0538)	(0.2232)	(0.0536)
Anti-Community Goods	-0.0450	-0.2296	-0.0518
	(0.0441)	(0.1897)	(0.0467)
Anti-Immediate Goods	0.1121*	0.4539**	0.1062*
	(0.0398)	(0.1613)	(0.0399)
Co-Local (Base: Recent Returnee)			
Strong Local Ties	0.0737**	0.2957**	0.0673*
	(0.0232)	(0.1023)	(0.0247)
Ethnicity (Base: Non-Coethnic)			
Coethnic	0.0203	0.1090	0.0300
	(0.0236)	(0.0979)	(0.0266)
Gender (Base: Male respondent)			0.0000
Female respondent			-0.0726*
			(0.0292)
Age (Base: Respondent Age <25)			
Age≥ 25 ≤ 35			-0.0247
			(0.0367)
Age≥ 35 ≤ 45			-0.0157
			(0.0344)
Age >45			-0.0263
			(0.0381)
Respondent Ed (Base: No formal)			
Some primary schooling			-0.0662
			(0.0473)
Primary school completed			-0.1628**
			(0.0555)
Intermediate to postgraduate			-0.1369+
			(0.0697)
ELF (Base: Lowest ELF quartile)			
Second ELF quartile			-0.0239
			(0.0364)
Third ELF quartile			0.0290
			(0.0470)
Highest ELF quartile			-0.0070
			(0.0647)
Region (Base: Central Region)			
Northern			0.1906**
			(0.0605)
Southern			0.0460
			(0.0520)
Constant	0.3437***		0.4161***
	(0.0511)		(0.0908)
cut1		0.2128	
		(0.2452)	
cut2		0.5530*	
		(0.2258)	
cut3		1.1960***	
		(0.1702)	
R-squared	0.0657		0.1012
N	1166	1166	1150

Note: + 0.10 *** p<0.05, ** p<0.01, * p<0.001. Standard errors in parentheses clustered by TA.

Table A2: Distribution of Candidate Rating in the Experiment

	Frequency	**Percent**
Not at all likely to vote for the candidate	463	40
Not likely to vote for the candidate	92	8
Likely to vote for the candidate	174	15
Very likely to vote for the candidate	436	38
Total	1166	100

Table A3: Distribution of Candidate Ratings by Appeal

Appeal	**Not at all likely**	**Not likely**	**Likely**	**Very likely**
Pro-Immediate Goods	93(49%)	20(11%)	29(15%)	47(25%)
Pro-Future Goods	61(32%)	16(8%)	32(17%)	82(43%)
Pro-Community Goods	52(24%)	13(6%)	33(15%)	119(55%)
Anti-Immediate Targeted Goods	80(41%)	12(6%)	28(14%)	75(38%)
Anti-Future Targeted Goods	86(40%)	20(9%)	30(14%)	77(36%)
Anti-Community Goods	91(57%)	11(7%)	22(14%)	37(23%)

Appendix B. Distributions of Demographics and Additional Analyses

Balance Tests

We ran balance tests regressing standard respondent demographics such as gender, education, and wealth (described in more detail below) on our experimental arms. We do not find a significant relationship between these characteristics and the randomized components of our study suggesting that they are balanced. (See Table C1)

Table B1: Standard Respondent Demographics Regressed on Experimental Arms

	Gender	Education	Wealth
Appeal (Base: Pro-Immediate Goods)			
Pro-Community Goods	0.0073	0.0313	0.0944
	(0.0475)	(0.0918)	(0.0769)
Pro-Future Goods	-0.0013	0.0674	0.0875
	(0.0493)	(0.0927)	(0.0770)
Anti-Future Goods	0.0318	-0.0210	0.1063
	(0.0474)	(0.0886)	(0.0791)
Anti-Community Goods	-0.0071	-0.0007	0.0850
	(0.0519)	(0.0970)	(0.0816)
Anti-Immediate Goods	-0.0202	0.0024	0.0597
	(0.0490)	(0.0900)	(0.0760)
Local Ties (Base: Recent Returnee)			
Strong Local Ties	-0.0150	-0.0168	-0.0223
	(0.0279)	(0.0530)	(0.0444)
Ethnicity (Base: Non-Coethnic)			
Coethnic	0.0076	0.0682	0.0354
	(0.0280)	(0.0529)	(0.0446)
R-squared	0.0014	0.0025	0.0028
N	1190	1187	1183

Note: + 0.10 *** p<0.05, ** p<0.01, * p<0.001. Standard errors in parentheses clustered by TA.

Distributions of Demographics in the Main Survey and Experimental Sample

Wealth We measure wealth in two ways. First, we have three groups resulting from a wealth measure of self-reported needs: households in which the income covers needs without much difficulty, households in which the income does not cover needs and there are difficulties, and households in which the income does not cover the needs and there are great difficulties. Second, we created an asset index was by performing a multiple correspondence analysis on four assets a household could possess: motor vehicle, mobile telephone, radio, and bicycle. The higher the value, the more assets a household possesses. This measure was standardized to range between 0 and 1 and then cut into quartiles. The distributions of our sample compared to the larger survey sample follow below in Tables B2 and B3.

Gender Table B4 reports the distribution of gender in our experimental sample compared to the full survey sample.

Table B2: Self-Reported Needs: I will read out a few statements about your income. Please tell me, which of the following statement is closest to your situation?

Bracket	**Respondents in Dataset**	**Percent**	**Respondents in Experiment**	**Percent**
Our household income covers the needs alright, without much difficulty/we can save.	1358	18	222	19
Our household income does not cover the needs, there are difficulties.	2700	35	408	35
Our household income does not cover the needs, there are great difficulties.	3585	47	528	46

Table B3: Distribution of Asset Index

Bracket	**Respondents in Dataset**	**Percent**	**Respondents in Experiment**	**Percent**
Lowest Wealth	2390	31	368	32
Low Wealth	1840	24	277	24
High Wealth	1699	22	255	22
Highest Wealth	1738	23	262	23

Table B4: Gender

Gender	**Respondents in Dataset**	**Percent**	**Respondents in Experiment**	**Percent**
Male	2843	37	414	36
Female	4939	64	761	65

Education Table B5 lays out the distributions of educational groups in our experimental sample compared to those of the full sample.

Table B5: Education

Education	**Respondents in Dataset**	**Percent**	**Respondents in Experiment**	**Percent**
No Formal Education	1078	14	152	13
Some Primary School	4058	53	610	53
Finished Primary School	1353	18	215	19
Some Intermediate School +	1170	15	185	16

Ethnolinguistic Fractionalization (ELF) We also considered whether our respondents came from different types of locales than those in the main survey. We control for this in the third model in Table A1 above. Table B6 displays the distributions of our experimental sample across ELF quartiles copmpared to the full survey.

Table B6: Ethno-Linguistic Fractionalization by Thirds

ELF	**Respondents in Dataset**	**%**	**Respondents in Experiment**	**%**
Lowest Diversity (Lowest Quartile of ELF)	2181	27	314	27
Low Diversity (Second Lowest Quartile of ELF)	2234	28	324	28
High Diversity (Third Lowest of ELF)	2354	30	364	32
Highest Diversity (Highest Quartile of ELF)	1186	15	152	13

Regional Differences Table B7 displays distributions of our experimental sample compared to the full survey sample by region.

Table B7: Region

Region	**Respondents in Dataset**	**Percent**	**Respondents in Experiment**	**Percent**
Northern	2500	31	349	30
Central	2281	28	305	26
Southern	3283	41	510	44

Vote Buying Below in Table B8 we show those who report having received gifts, food, or money from a candidate before in an election. Our experimental sample is similar to the overall sample.

Table B8: In any election, have you received gifts, food, or money from a candidate?

	Respondents in Dataset	**Percent**	**Respondents in Experiment**	**Percent**
No	5752	75	850	73
Yes	1901	25	315	27
Total	7688	100	1191	100

Heterogeneous Effects

We worry that our sample is not large enough to trust the outcomes of heterogeneous effects. Nonetheless, in the spirit of transparency, Table B9 displays the outcomes of a fully interacted model of our experiment. Specifically, it shows the average component interaction effect (ACIE) for each of our experimental arms evaluating the casual effect of one candidate feature depending on what the value of another feature is held at.

Table B9: ACIEs of Profile Features on Likelihood of Vote

Appeal (Base: Pro-Immediate Goods)	
Pro-CommunityGoods	0.1999*
	(0.0861)
Pro-Future Goods	0.1567*
	(0.0656)
Anti-Future Goods	0.0807
	(0.0778)
Anti-CommunityGoods	-0.0631
	(0.0653)
Anti-Immediate Goods	0.1079
	(0.0906)
Ethnicity (Base: Non-Coethnic)	
Coethnic	0.0056
	(0.0566)
Co-Local (Base: Recent Returnee)	
Co-Local	0.0958
	(0.0642)
Interactions (Base: Pro-Immediate Goods Non-Coethnic vs. Coethnic)	
Community Goods × Coethnic	0.1453*
	(0.0695)
Pro-Future Goods × Coethnic	0.0367
	(0.0794)
Anti-Future Goods × Coethnic	0.0267
	(0.0801)
Anti-Community Goods × Coethnic	0.0653
	(0.0811)
Anti-Immediate Goods × Coethnic	-0.0063
	(0.0915)
Interactions (Base: Pro-Immediate Goods Recent Returnee vs. Strong Local)	
Pro-Community Goods × Strong Local	0.0260
	(0.0996)
Pro-Future Goods × Co-Local	0.0239
	(0.0767)
Anti-Future Goods × Co-Local	0.0031
	(0.0728)
Anti-Community Goods × Co-Local	-0.0295
	(0.0796)
Anti-Immediate Goods × Co-Local	0.0210
	(0.1097)
Interactions (Base: Strong Local Non-Coethnic vs. Coethnic)	
Coethnic × Co-Local	-0.0660
	(0.0531)
R-squared	0.0709
N	1166

Note: + 0.10 *** p<0.05, ** p<0.01, * p<0.001. Standard errors in parentheses clustered by TA.

Appendix C: Survey Methodology

We conducted two rounds of focus group discussions: One prior to implementing the survey in January 2016 and one after our initial analysis of the survey data in September 2017. The focus group discussion (FGD) participants at each site were screened by age and gender. Following standard FGD protocols, three groups were convened at each site, one made up entirely of women and one of men one of youth (gender mixed). The discussion checklist was designed to elicit in-depth discussion on the study objectives, while being mindful that the basic operative rule for FGDs requires that facilitation should, as much as possible, be kept to a minimum.

To ensure open and full discussions, the fieldworkers were advised to limit the number of FGD participants to a minimum of six and a maximum of ten. This was done in conformity with global FGD guidelines that suggest that large groups may often result in a higher number of participants who remain uncommunicative. The discussions were conducted in Chitumbuku and Chichewa, by two Malawian enumerators in each group.

Pre-Survey Focus Groups: January 2016

Before implementing the survey, we conducted 18 focus groups, for men, women and youth in the three regions of Malawi: Northern, Central and Southern. In each region we selected one urban and one rural district for the focus group discussions. We also used these focus groups to determine the types of goods candidates are most likely to promise during the campaign season. A total of 174 people participated in the focus group discussions that were carried out between January 4 and January 6, 2016. The main purpose of the focus group was to ensure a common understanding and use of terms and concepts. The focus group discussions were structured into four main themes: Key terms, the role and functions of various local authorities, service delivery, power, politics and voting.

Table C1: Pre-Survey Focus Group Discussions

	Village/area of discussion	Number of focus group participants (youth)	Number of focus group participants (male)	Number of focus group participants (female)
Northern region				
Urban	**Mzuzu City (Nkhata Bay)**	**10**	**11**	**13**
Rural	**Timbiri (Nkhata Bay)**	**10**	**11**	**8**
Southern region				
Urban	**Kachere**	**10**	**11**	**10**
Rural	**Chiraduzulu**	**7**	**14**	**10**
Central region				
Urban	**Lilongwe Area 25**	**11**	**8**	**13**
Rural	**Ntcheu District**	**9**	**10**	**8**
TOTAL		**57**	**65**	**52**

The Malawi Survey

The survey was conducted in Malawi during March and April 2016.[1] The survey sought to measure and better understand governance and service delivery at the local level. It is a highly clustered survey, which facilitates measurement and inference at the local (in this case, village) level. The survey covers the following topics: political participation, social norms and institutions, education, health, security, welfare, corruption, land, and dispute resolution.

The sample was stratified on region (North, Central, South), the presence of matrilineal and patrilineal ethnic groups, and the urban/rural divide. Because patrilineal groups are rare in Malawi and we wanted to maximize variation in matrilineal and patrilineal heritage, we oversampled Primary Sampling Units (PSUs) from the patrilineal stratum. We sampled 22 PSUs, namely 'Traditional Authorities' (TAs). These 22 sampled TAs are located in 15 of

[1]We implemented the survey using tablet computers. The experiment was seen by a random subsample of 1,191 of the full 8,100 survey respondents.

Malawi's 28 districts. [2] Within each TA (i.e., PSU), we selected randomly four enumeration areas (EAs) as Secondary Sampling Units (SSUs). EAs are comparable to census tracts. Both PSUs and SSUs were selected without replacement according to the principle of Probability of Selection Proportional to Measure of Size (PPMS). Within each EA, we sampled four villages, based on known geographical points provided on the maps of the EAs produced for Malawi's latest population census. Once in the village, enumerators followed a random walk pattern to select households. After they entered the household, the interviewer collected the necessary data about composition of the household. Both the contact questionnaire and the main questionnaire we programmed on digital tablets, including the selection of the final respondent in the household through a digital version of the "Kish grid". The target was to interview 22 respondents in each village. This process produced a sample of 8,100 respondents. See Table C2 for a list of the districts and TAs included in the sample.

While the sampling procedures were planned as presented, of course in practice this was not always the case. In total the research team had to draw 11 replacement EAs. One replacement EA was drawn because enumerators were chased out of a village and forced to withdraw from the EA. In the remaining 10 cases, EAs were not accessible (e.g. in one instance our team was unable to reach the designated EA because a bridge had washed away during heavy rains). In these instances, backup enumeration areas were randomly selected within the same EAs (excluding already selected and inaccessible zone) and were used as replacements. [3] In addition, given that multiple enumerators conducted surveys in the same village, the target number of 22 respondents per village (neighborhood in urban areas) was not always reached precisely. [4]

[2]Districts are the largest sub-national administrative units in Malawi.

[3]In total, only 11 of the 99 sampled EAs are replacement EAs.

[4]In some instances more were surveyed and in others slightly fewer than 22 households were surveyed. In addition, the boundaries between villages and neighborhoods were not always clear, which also caused our teams to deviate from the target of 22 per village/neighborhood.

Table C2: Survey Districts

Region/Stratum	District	Traditional Authority
Northern	Chitipa	Mwaulambya
	Rumphi	Mwankhunikira
	Mzimba	Chindi Kampingo Sibande Mtwalo
	Nkhata Bay	Kabunduli
	Mzuzu	Viphya ward
Central	Kasungu	Simlemba
	Lilongwe City	Area 25 ward Area 36 ward
	Dedza	Pemba Tambala
	Ntcheu	Kwataine
Southern	Balaka	Kalembo
	Blantyre	Kapeni
	Blantyre City	Namiyango ward
	Chikwawa	Chapananga Ngabu
	Mangochi	Jalasi
	Mulanje	Mabuka
	Nsanje	Mbenje
	Zomba	Mwambo

Interviewer training took place March 15–18, 2016 in which the authors and local partners trained 56 interviewers on site in Malawi. Interviewers were recruited based on experience, qualifications, and languages spoken. The survey was conducted in Chichewa, Chitumbuka, and English. Interviewers were divided into teams of five, each with a team leader. On average, each team spent 4 days in a TA: one day for each EA. This helped facilitate callbacks given that the teams were in a single area for multiple days. The survey was completed over the course of 33 days; March 26, 2016 – April 27, 2016.

Post-Survey Focus Groups

After our initial analysis of the survey, we conducted FGDs with women, men, and youth to further explore our findings and to gain a deeper understanding of voter attitudes toward vote-buying. Focus group discussions were conducted with a total of 99 participants in two villages each in the two districts of Dedza and Mzuzi. The discussions centered around the following main themes: 1) Local/National Elections, 2)vote-buying, 3) Local Origin Vs Local Residence Vs Ethnicity, 4) Ethnicization (Importance of Ethnic Group, Obligation to help: Ethnicity Vs Religion, Importance of working with co-ethnics, 5) diverse people working together, Roles in Ensuring quality education and health services), 6) Mixed Ethnicity.

A total of 12 FGD sessions —- three in each of the four sample sites – were held. The highest participation rate was ten respondents, while the minimum group had seven. Overall, a total of 34 men, 33 women, and 32 youth in the FGDs. Based on neighborhood, it is possible to divide our FGDs by income brackets. The three focus groups conducted for men, women and youth in Masasa (Mzuzu), may be characterized as middle class where typically, people that are able to pay bills but not save. Typically, the FGD members in Masasa were either engaged in business or they reported to be civil servants (men). The nine other focus groups conducted in rural Mzuzi (Malivenji village) and Dedza (Kabinda and Gunduze) can be characterised as low income, primarily rural people dependent on subsistence farming.