Is Vote-buying Effective?
Evidence from a Field Experiment in West Africa*

Pedro C. Vicente†

Forthcoming at the Economic Journal

Abstract:
Vote-buying, i.e., cash-for-votes, happens frequently in many parts of the world. However, in the presence of secret ballots, there is no obvious way to enforce vote transactions. To infer effects of vote-buying on electoral behaviour, we designed and conducted a randomized field experiment during an election in Sao Tome and Principe. We follow a voter education campaign against vote-buying, using panel-survey measurements as well as disaggregated electoral results. Results show that the campaign reduced the influence of money offered on voting, decreased voter turnout, and favoured the incumbent. This evidence suggests that vote-buying increases participation and counteracts the incumbency advantage.

JEL Codes: D72, O55, P16.
Keywords: Vote-buying, Electoral Politics, Political Economy, Randomized Experiment, Field Experiment, Sao Tome and Principe, West Africa.

* I wish to thank the editor David Myatt, four anonymous referees, Abigail Barr, Robert Bates, Cátia Batista, Tim Besley, Paul Collier, Esther Duflo, Marcel Fafchamps, John Gerring, Macartan Humphreys, Asim Khwaja, Rocco Macchiavello, Casey Mulligan, Roger Myerson, Rohini Pande, Susana Peralta, Torsten Persson, Gerhard Seibert, and Leonard Wantchekon for helpful suggestions. I am particularly grateful to Christian Helmers for excellent research assistance during the fieldwork. I thank participants at AEA/ASSA, NEUDC, BREAD/CEPR, ASA, APSA, EEA, and CSAE conferences, and at various seminar presentations. The National Electoral Commission of Sao Tome and Principe, through its president, José Carlos Barreiros, provided invaluable local support to this research, for which I am most indebted. I am also thankful to CPLP, namely Hélder Lucas and Leonardo Simão, for coordination with their electoral observation mission. A. Assunção, C. Castro, D. Coelho, A. Costa, A. Dias, M. Diogo, E. Lima, H. Mandinga, C. Moreno, A. Penhor, and E. Solé participated in the fieldwork team: their dedication was crucial for the success of this project. I wish to acknowledge financial support from the Open Society Institute. Any errors are my responsibility.
† Universidade Nova de Lisboa, and Bureau for Research and Economic Analysis of Development – BREAD. Postal address: Nova School of Business and Economics, Campus de Campolide, 1099-032 Lisboa, Portugal. Email: pedro.vicente@novasbe.pt.

> "We do like vote-buying. It is essential. That is the only way we have to see anything good coming from the politicians. Anyway, I can vote for whoever I want."
> - Anonymous (Sao Tome)

The idea that a few can capture the democratic electoral process of a country has been at the centre of political economics from its inception. Pressure groups and lobbying are recognized as essential ingredients for the understanding of modern democracies. In this context the profession tends to rest assured that competition (Becker, 1983) and informed voters (Grossman and Helpman, 1996; Besley, 2006) keep the political outcome close to efficiency and not too biased in favour of the elite. Privately-funded campaign spending is then the main visible component of an essentially-benign influencing process.

Interested observers are, however, becoming more alert to the weaknesses of democratic processes in developing countries. There, competition may be curtailed by credit constraints, and the demand for information may be undermined by voters' lower education. Moreover, weak political institutions may be inducing inefficient means of redistribution of resources (Coate and Morris, 1995; Acemoglu and Robinson, 2001). Indeed, clientelism and vote-buying have been prevalent in many recent African elections (Wantchekon, 2003; Collier and Vicente, 2012a). But other electoral problems such as electoral violence and ballot fraud have been prominent as well (Collier and Vicente, 2012b; Ichino and Schundeln, 2012). There is a widespread sense that many African elections have not been centred on policy accountability, with possible marked consequences for economic development. It calls for a deeper understanding of electoral campaigning in these elections.

Despite the prevalence of vote-buying, understood as the exchange of cash for votes before the elections,[1] our knowledge about the consequences of vote-buying is limited. The main puzzle has been that in the presence of secret ballots there is no clear enforcement mechanism for vote transactions.[2] Note that we distinguish vote-buying from clientelism, usually seen as the exchange of votes for favours (e.g., jobs in the public sector) conditional on winning the election. These transactions are easier to enforce and call for long-term relationships between politicians and their

[1] See Shaffer (2007), for a review of the main issues on the study of vote-buying.
[2] Secret ballots led to a substantial decrease in vote-buying at the end of the XIXth century in the US and England, as described by the historical studies of Cox and Kousser (1981) and Cox (1987). Converse (1972) relates the introduction of secret voting in the US to a decrease in voter participation, which is consistent with our experimental results.

clienteles (Robinson and Verdier, 2003). In this paper we seek to shed some light on whether vote-buying is effective in determining voting and its outcome.

For that purpose, we conducted a field experiment during the presidential elections of 2006 in Sao Tome and Principe (STP), a two-island country in West Africa where oil was discovered in the late 1990s. Earlier work (Vicente, 2010) found an important increase in corruption following the discovery announcements, with vote-buying featuring prominently. We followed a voter education campaign against vote-buying practices. This campaign was randomized across a representative set of locations in the country. We are thus able to quantify the impact of the campaign on the prevalence of vote-buying and on voting behaviour, as given by voter turnout and candidate selection at the polls. Those measured impacts correspond to the main quantitative results of the paper. In the event that the campaign reduced the effectiveness of vote-buying, and this constituted the campaign's main mechanism of change, we are able to infer the likely effect of vote-buying on the electoral outcomes.

The voter education campaign occurred in 40 enumeration areas of STP, with the sponsorship of the National Electoral Commission of STP. It was based on a leaflet distributed, read, and discussed door-to-door. The leaflet was mainly legalistic in that it stressed the illegal nature of vote-buying. Orally, the campaign underscored the idea that voting decisions should be conducted in good conscience, even if gifts were accepted. Measurement comes from a panel survey in treatment and control enumeration areas, and from the official electoral results per ballot station. The panel survey conducted before and after the elections included 1034 respondents (more than 1 percent of the electorate of STP). It gathered a wide range of reported perceptions and experiences about vote-buying, as well as reported voter behaviour. In connection to the survey, we designed a behavioural measure of the demand for political accountability on corruption in the public services, and we designed the timing of the treatment in order to quantify any conformity biases the treatment may have induced on survey respondents.

We find that the campaign decreased the reported perception that voting decisions were affected by the money offered by candidates, and increased the reported perception that the voting was conducted in good conscience (i.e., decreased the effectiveness of vote-buying transactions). These are large effects, reaching 0.32-0.48 standard deviation units. We also find effects on decreasing the frequency and price of vote transactions. These effects are shown to be robust to conformity biases, i.e., the possibility that respondents conformed to the treatment when reporting

the vote-buying outcomes. Regarding actual voter behaviour at the polls as measured by official results, we can document an effect of voter education on decreasing voter turnout (by 3-6 percentage points), on increasing the vote share of the incumbent, and on decreasing the vote share of the challenger (both by close to 4 percentage points). These effects are confirmed in our survey data, at the individual level. We do not find statistically significant changes in the demand for political accountability using our behavioural measure. Informed by the impact of the voter education campaign on decreasing the effectiveness of vote-buying, the results on voter behaviour point to a positive effect of vote-buying on participation and to greater reliance of the challenger (relative to the incumbent) on vote-buying practices. We interpret these implications of our results as being consistent with an incumbency advantage, possibly stemming from the higher ability of the incumbent to engage in clientelism. If vote-buying breaks voter inertia and raises participation, it may be used by the challenger to counteract the incumbency advantage close to the election.

Even though vote-buying can be seen as a substitute for public good provision, the purpose of this paper is not to make definite welfare considerations regarding the opening of markets for votes. Indeed, even the theoretical literature on vote-buying is ambiguous about the desirability of such markets on efficiency grounds. Disparate conclusions have been presented by a number of authors under slightly different assumptions. Philipson and Snyder (1996) argue that equilibria in markets for votes involve vote-selling only when it is Pareto-superior. Dal Bo (2007) finds that a principal can influence the decisions of a committee at no cost and induce inefficient outcomes. Dekel et al. (2008) find that efficiency is independent of the presence of vote-selling and of the specific forms that it may take (clientelism vs. enforceable vote-buying).

Our paper closely relates to empirical research undertaken on clientelism and vote-buying in developing countries. Wantchekon (2003) reports on a field experiment conducted in Benin where the contents of electoral campaigning by candidates were randomized across locations. A purely clientelistic message was contrasted with a message based on public policies. It is shown that clientelism works better for a wide range of candidates. Brusco et al. (2004), Stokes (2005), and Nichter (2008) use non-experimental survey data on vote-buying in Argentina to test different hypotheses. Brusco et al. (2004) look at individual correlates of vote-buying and reinforce the idea that vote-buying is effective when vote-buying transactions are enforceable. Stokes (2005) and Nichter (2008) test theories of party strategic behaviour considering the difficult enforcement of vote-buying transactions. While the first argues that parties target weakly

opposed voters, the second puts the target on own strong supporters for whom voting behaviour can be enforced by observing their turnout. Recently, Finan and Schechter (2012) provide evidence from Paraguay on the association between vote-buying (survey measured) and reciprocity (as measured through lab games) at the individual level, and find support for the idea of self-enforcement of vote-buying transactions.

Our work also relates to the extensive empirical literature on the effectiveness of campaign spending, even though this literature is centred on the US and our focus is on a very specific kind of campaign spending. This literature began with Jacobson (1978), who presented evidence that challengers have higher returns to campaign spending than incumbents. Later, Green and Krasno (1988), Levitt (1994), Gerber (1998), and Erikson and Palfrey (2000) challenged the initial result by showing similar voting returns for incumbents and challengers, typically close to zero. More recently, a series of experimental papers recovered the initial pattern of Jacobson (Gerber, 2004) and proved effectiveness of non-partisan get-out-the-vote campaigning (Gerber and Green, 2000). Our findings are closely aligned with this experimental literature, not only qualitatively, in terms of favouring the challenger and affecting turnout, but also quantitatively, since effects of campaigning are generally below 10 percentage points on both dimensions.

In Section 1 we present the historical context of the election we study. Section 2 offers the experimental design with details of the treatment, measurement, and estimation strategy. Section 3 describes the econometric results, namely on vote-buying and voting behaviour, with corresponding auxiliary findings. Section 4 provides our interpretation of the experimental results. Section 5 concludes.

1 Historical Context

STP is a two-island West African country with 155 thousand inhabitants. It is a low-income country with GDP per capita USD PPP 1500, ranking 168 out of 203 countries.[3] After almost five centuries of Portuguese colonization, it achieved independence in 1975. As was the common trend in Portuguese-speaking Africa, its first regime was socialist, with the MLSTP (Liberation Movement of Sao Tome and Principe), the independence movement, taking the role of ruling party. In the late 1980s, the fall of the Soviet Union, together with plummeting cocoa prices

[3] World Development Indicators, World Bank, 2006.

(STP's main export), led the ruling elite to start a democratization process that culminated with the first free elections in 1991.[4]

STP was constituted as a semi-presidentialist democratic regime, with most executive/legislative powers attributed to the National Assembly, from which the government emerges, but with important arbitrage and foreign affairs authority given to the president. Contrary to most first multi-party elections in Africa, the incumbent party MLSTP and president Pinto da Costa (1975-1991) were ousted in 1991. Post-democratization politics in STP have been dominated by MLSTP/Pinto da Costa and two other 'political families'. These have been centred on the figures of Miguel Trovoada (president from 1991 to 2001), who founded the ADI (Independent Democratic Alliance) in 1993, and Fradique de Menezes (president from 2001 to 2011), who founded the MDFM (Democratic Movement for Empowered Reform) in 2001.

Importantly, there was a significant oil discovery in STP in the late 1990s that brought considerable international attention to the country. However, the process of oil exploration has been tainted with problems, from unfavourable initial contracts for soundings/exploration to numerous allegations of corruption within the STP political elite. Auctions were opened and concession blocks allocated to oil companies starting in 2003 (STP already benefitted from significant contract signing bonuses), but the start of production activities has been postponed time and again. In this scenario, the STP government has faced strong pressures from the international oil sector. Frynas et al. (2003) and Vicente (2010) provide evidence linking the surge in oil-related interests to a steep increase in vote-buying, starting with the presidential and parliamentary elections of 2001/2002.

1.1 The 2006 Round of Elections

2006 was an important election year in STP, with both parliamentary and presidential elections. The first (parliamentary) took place in March and the second (presidential) on the 30th of July. As described above, due mainly to the expected oil boom, stakes were considered to be very high for both elections. MDFM-PCD gained control of the parliament with a victory over MLSTP (36.8 vs. 29.5 percent); ADI came third with 20 percent. International observers considered these elections to have respected international standards, but reported frequent vote-buying practices.

[4] See Seibert (2006) for a thorough historical coverage of this period.

The presidential elections featured Fradique de Menezes (FM), MDFM-PCD's driving force, running for re-election. It was clear by the last year of FM's mandate that he would be running for re-election. What was not clear at all, still by the beginning of June, was who would run against him, representing MLSTP and/or ADI. This was formally known later that month: Patrice Trovoada (PT), son of Miguel Trovoada, former minister with responsibilities in the oil-related deals and member of ADI. PT was a young and rising political figure, and managed to secure MLSTP support, apart from the natural ADI sponsorship - i.e., both major opposition parties. Since our analysis starts in the beginning of July, we are confident that our pre-election field activities precede or are contemporaneous with the whole campaign process by PT (including the bulk of his campaign's vote-buying activities). Moreover, this setting is one in which the incumbency advantage is clear. FM won comfortably with 60.6 percent against 38.8 percent by PT.[5] These results were well accepted by not only the candidates but also international observers, who nevertheless repeated concerns related to vote-buying.[6]

2 Experimental Design

Our experiment is based on the randomized allocation of a voter education campaign to neighbourhoods and villages of STP. This was a door-to-door information campaign against vote-buying practices sponsored by the Electoral Commission. The exogenous variation produced by this intervention allows us to identify the effect of the campaign on reported perceptions about vote-buying (arguably the most direct effect of the campaign), and the reduced-form effect on voting behaviour, as revealed by voter turnout and candidate selection. In this section, we begin by describing the intervention. We then explain the design of our measurement, which includes the use of survey-based and official electoral data. Finally, we present our main econometric specifications and the methods we employ for statistical inference.

[5] There was a third, independent candidate (Nilo Guimarães), who won 0.6 percent of votes. Given this outcome and the negligible political relevance of this candidacy, we do not mention this candidate in our analysis.

[6] The main contingent of international observers was deployed by CPLP (Portuguese-Speaking Community of Countries), including representatives from Angola, Brazil, Cape Verde, East-Timor, Guinea-Bissau, Mozambique, and Portugal. Full contents of their report can be found in CPLP, 'Relatório da Missão de Observação Eleitoral da Comunidade dos Países de Língua Portuguesa às Eleições Presidenciais de São Tomé e Príncipe de 30 de Julho de 2006'. Specifically, CPLP writes in its executive summary: 'Vote-buying was observed in various occasions. It apparently constitutes generalized practice among the candidates.'

2.1 The Voter Education Campaign Against Vote-buying

The intervention that is the focus of our analysis is a voter education campaign sponsored by the National Electoral Commission of STP. This endorsement was systematically mentioned at all moments of the campaign. It embedded a very specific focus on vote-buying, as by 2006, it had already become a widespread practice during election time in the country. The Electoral Commission was interested in improving the conduct of the electoral campaign and of the election in this regard.

The campaign consisted of the door-to-door distribution of a leaflet, which was fully read at the time of its delivery. Moreover, there was a subsequent discussion of the leaflet, meant to address questions raised by the campaign subjects and to verify that subjects understood the message of the campaign.

The leaflet was mainly legalistic, as it was centred on the fact that vote-buying is illegal under the Sao-Tomean law. Its main slogan was 'Do not let your conscience take a 'banho'', where 'banho' is the Portuguese word commonly used in STP (and only in STP to the best of our knowledge) to refer to vote-buying (meaning literally 'bath' or 'shower'). The leaflet also included an illustration, which enabled easier communication with the illiterate part of the population. The leaflet is shown in Figure 1.

< Figure 1 around here >

However, it was clear before the voter education campaign that vote-buying is a popular phenomenon for the impoverished STP population. In our baseline survey we asked how vote-buying was perceived (before the campaign). Although many associated vote-buying primarily with a bad practice on moral or legal grounds, there is a considerable portion who (explicitly) viewed vote-buying in a positive way (representing 22 percent of the answers) – perceiving vote-buying as 'good for the population' or even for democracy. As a consequence, we included in the campaign's verbal discussion an emphasis on voting according to one's judgment about the quality of the politicians, and not according to whoever had bought one's vote. Hence, the discussion of the leaflet emphasized voting in good conscience, even if still accepting cash from politicians. Despite the fact that the legalistic view was probably taken less strictly with this type of message, stressing the idea of voting in good conscience, even if accepting cash from

campaigners, is still fully consistent with the main research questions in our experiment, namely through diminishing the effectiveness of vote-buying.

The campaign delivered 10,000 leaflets in treated locations (roughly equivalent to one leaflet per household on average in those locations). The distribution of leaflets was conducted during the period of the pre-election survey and targeted our survey respondents primarily (who were chosen in a representative fashion – details are provided below). Neighbours of respondents (secondarily) were also targeted in the same way in treated areas. This design enabled a clear focus on our survey respondents while seeking to ensure real effects at the location level.

2.2 Measurement: Survey and Electoral Data

We use data from a representative panel survey we designed and conducted[7], and from official electoral data at the ballot station level made available by the National Electoral Commission for both the parliamentary and presidential elections of 2006. The first type of data was primarily aimed at assessing effects on reports about vote-buying perceptions. The second sought the identification of the effects of the campaign on electoral behaviour. The use of self-reported electoral behaviour from the surveys also serves the purpose of validating the results obtained at the ballot station level.

Our survey sample consisted of 50 enumeration areas. Within these, 40 were given the voter education campaign. The choice of the treated locations followed a simple randomization procedure, which was implemented by the author. The 50 enumeration areas are representative of STP. They were chosen randomly within all the census areas of the country, weighting for the number of households (using the 2001 census data from the National Statistics Office of STP). The sampled locations are shown in Figures 2 and 3.

< Figure 2 around here >

< Figure 3 around here >

[7] The field activities for the experiment (survey plus campaign) were conducted by a team of 11. This team was fully recruited and trained by the author and a research assistant. The author was present in the field working with the team at all times of the conduction of this experiment.

Within each enumeration area, sampled households were chosen randomly using standard techniques: namely by selecting the n^{th} house (depending on the number of households in the enumeration area), with second visits tried in the same day. The requisites for sampling a respondent within a household were: age 18 years or higher, and residence in the country. Despite the fact that this is a standard sampling procedure, we still faced imperfections to clustered random sampling. These stem from small differences in the number of interviews across enumeration areas and from non-respondents. To address these limitations, we use weighted data in the regressions shown further below.[8] This is uniquely for consistency with the sampling approach. Differences to un-weighted data were verified to be negligible.

We were able to gather 1275 observations in the pre-election survey and 1034 observations in the post-election survey. This outcome represented an 81 percent rate of re-surveying in the panel (differential attrition across comparison groups is tested below). The survey instruments were tailored for this experiment.[9] The pre-election questionnaire featured demographic questions, a module on perceived corruption in the public services (only marginally related to this paper), followed by questions on vote-buying (mainly about the parliamentary elections of March 2006), and individual political positioning (mainly intentions about the July presidential elections). The post-election survey focused on perceptions about vote-buying and electoral behaviour during the presidential elections.

In connection to the survey we designed a behavioural measure of the demand for political accountability – what we call the postcard variable. During the baseline survey, and after the module of questions on perceived corruption in the public services were asked, each respondent was given a pre-paid postcard. They were told that if enough postcards (50 percent) were sent to the sponsors of the survey the results of the module on perceived corruption would be made public through the media of the country.[10] The statement demanding the dissemination of results was written on the postcard. We were able to identify the respondents that sent the postcard back as each postcard was numbered beforehand. We interpret the sending of a postcard as a costly

[8] Data weighting accounts for non-respondents provided we collected information for those individuals: gender, and approximations for age, schooling, and income.
[9] Both pre- and post-election instruments are available upon request from the author.
[10] A number of media outlets were contacted in advance for this purpose. Several articles and reports were published/broadcast about our research project.
See http://www.csae.ox.ac.uk/resprogs/corruption/stp2006/stp%202006.htm for the details of fieldwork including media coverage.

action (i.e., mailing the postcard involved effort) undertaken to demand political accountability, specifically regarding corruption in a wide range of public services.

We note that because survey measurements of perceptions (about vote-buying) and self-reported behaviour (on voting) may have limitations, some attention should be devoted to robustness of results encountered for such outcomes. In particular, there is a possibility that the survey is inducing conformity biases. Experimental subjects in treated locations could in principle adapt their responses about politics to whatever they perceived to be the views of the sponsors of the experiment. We deliberately asked all pre-election questions about politics, namely about past events in the March parliamentary elections, after the treated subjects were offered the leaflets with corresponding discussion. We are thus able to measure whether there were undue reactions to the treatments in terms of reporting by contrasting treatment and control groups. Differences in reported perceptions about the past (e.g., about vote-buying in the March parliamentary elections), gathered at the baseline just after the campaign was delivered (in early July), constitute evidence of conformity. In addition, if one assumes that the same conformity bias exhibited by the respondents in the post-election survey is present in the pre-election survey, the difference between the two survey measurements can be taken as being closer to the true difference.[11] We therefore report results for difference-in-difference estimations whenever pre-election data are available.

In Figure 4 we show the sequence of the experiment including treatment and survey measurement.

< Figure 4 around here >

Crucially, we use data provided by the National Electoral Commission on voter registration and results for both the parliamentary and presidential elections of 2006. These are data for the 228 ballot stations of the country. The locations for the survey/campaign data were directly matched with these data. We are therefore able to test our main hypotheses on the impact of the campaign on the official records of electoral behaviour. We were also able to submit a questionnaire to all observers of the main international observation mission in the July elections. Specifically, we

[11] This assumes that any bias is additive and separable. To reinforce symmetry in the pre- and post-election measurements, respondents were asked in the post-election survey (in treatment locations) whether they remembered the leaflet campaign.

collected data on time spent at each ballot station. This variable has the potential to reveal a clear association with the electoral results at the ballot station level and, therefore, to perform as a valuable control in our analysis.

2.3 Estimation Strategy

In this paper our main econometric exercises regard the estimation of our campaign's average treatment effects. We are mainly interested in three types of outcome variables: vote-buying reported perceptions at the individual level, electoral results at the location level, and survey-based voter behaviour at the individual level. We turn now to the main econometric specifications employed.

For simplicity we show econometric specifications at the individual level, provided our regressions at the ballot station level follow a similar structure. We do not distinguish between vote-buying and voting behaviour outcomes, since we focus on reduced-form regressions. Our design allows us to estimate average treatment effects in different ways. Most simply, the effect of interest (f) could be estimated through the specification:

$$O_{il1} = a + bY_l + fT_l + \varepsilon_{il1}, \qquad (1)$$

where O is a political outcome (vote-buying or voting behaviour outcome), $i, l, t = 1$ are identifiers for individuals, locations, and time (specifically, 1 represents the post-election measurement), Y_l is a vector of district dummy variables, and T_l is a dummy variable with value 1 for treated locations.

In this setting, because of small sample size, we can also add individual-level control variables to compose one of our main specifications. This is in line with Duflo et al. (2007), who argued that even though controls do not generally change the estimate for the average treatment effect, they usually help to explain the dependent variable, and therefore typically lower the standard error of the coefficient of interest. We then have the following specification:

$$O_{il1} = a + bY_l + cX_i + fT_l + \varepsilon_{il1}, \qquad (2)$$

where X_i is a vector of individual demographic controls.

Specification (2) does not use the time dimension. In any event this may not be possible in some cases as we do not have repeated measurement for all our outcomes. However, when possible, it may be relevant to control for differing time trends across treatment and control groups. This may be achieved by including baseline values of the dependent variable in X_i while employing specification (2), or by using a time-differenced outcome variable, as in specification (3) below, which allows interpreting f as the difference-in-differences estimator. Note that just as for the regression in levels, it may be the case that individual-level control variables help to improve the precision of our estimate of interest.

$$\Delta O_{il} = d + fT_l + gY_l + hX_i + \Delta\varepsilon_{il}. \quad (3)$$

Since the data we use are clustered by enumeration area, we allow for within-group dependence in estimating standard errors of treatment effects by estimating cluster-robust standard errors through the use of the Huber-White variance estimator (see Moulton, 1990, for a defence of the use of corrected standard errors). Note, however, that a practical limitation of inference with cluster-robust standard errors is that the asymptotic justification assumes that the number of clusters goes to infinity. Bertrand et al. (2004) show that with a small number of clusters the cluster-robust standard errors are likely to be downward biased. Since we have 50 clusters in our survey-based data, we are close to having a small number of clusters, and therefore employ the wild bootstrap approach proposed by Cameron et al. (2008) as an additional inference method.

Cameron et al. (2008) recommend continuing to use the standard OLS estimator with the cluster-robust (Huber-White) variance estimator. However, they prescribe bootstrapping to obtain bootstrap critical values that provide an asymptotic refinement when there are few clusters. Bootstrap methods generate a number of pseudo-samples from the original sample; for each pseudo-sample they calculate the treatment effect; and use the distribution of the treatment effect across pseudo-samples to infer the distribution of the actual treatment effect. Wild bootstrap uses the fact that we are assuming additive errors and holds regressors constant across the pseudo-samples, while resampling the residuals at the level of the cluster, which are then used to construct new values of the dependent variable. Note that Cameron et al. (2008) advise that Rademacher weights (+1 with probability 0.5 and -1 with probability 0.5) are used when

resampling residuals, and that the null hypothesis of zero treatment effect is imposed. We follow both recommendations.

Because the wild bootstrap method applies only to OLS regressions, and for ease of interpretation of treatment effects, we run OLS regressions for all estimations in this paper.

3 Econometric Results

In this section we present our empirical results. We begin with standard balance tests, checking whether the randomization of the treatments was effective in identifying comparable treatment and control groups. We then focus on our main results: the effects of the voter education campaign on vote-buying outcomes, on electoral results per ballot station, and on political behaviour as reported by individuals in the survey. Next, we present a set of auxiliary results: a test for conformity biases by assessing effects of the treatment at the baseline, and a test for contamination of control locations.

3.1 Balance

We begin by assessing whether the randomized selection of treated locations was successful in identifying comparable treatment and control groups. We document differences across these groups in terms of a wide range of observable initial characteristics. In Table 1 we contrast our comparison groups in terms of their basic demographic profiles (age, household size, marital status, schooling, number of children, and health status), nationality, ethnicity, religion, occupation, expenditure and property, and in terms of baseline electoral outcomes, namely actual turnout and voting patterns in the March parliamentary elections. We contrast individual-level (survey-based) variables both at the baseline and at the post-election surveys, in order to account for panel attrition. Note that we cannot use political variables from the pre-election survey to test for balance because these data were gathered after the treatment was administered. We also test for differential panel attrition rates across treatment and control. Because demographic and baseline political variables are unaffected by the intervention, any differences between treatment and control groups should be understood as a product of chance.

< Table 1 around here >

We observe that there are generally no differences (at standard statistical significance levels) between treatment and control groups. The exceptions are that the treatment group has fewer respondents from Forro ethnic origin and more from Angolar origin (both at the baseline and the post-election surveys), and that there are more respondents working in the education sector in the treatment group (only at the baseline survey). Note that panel attrition in the post-election survey had no clear consequences in terms of comparability of treatment and controls groups. Consistently, panel attrition rates are found to be not statistically different across the two groups.[12] Overall, this is evidence that the randomization was generally effective in isolating similar groups of locations and respondents. The fact that observables are balanced across treatment and control makes it reasonable to expect unobservable dimensions to be balanced as well.

Table 1 provides complete descriptive demographics for our survey sample, in the process providing a comprehensive description of our experimental locations. As far as the control group is concerned, the average respondent was 38 years old, 58 percent of the sample was in an unmarried relationship, the main ethnicities represented were Forro and Angolar, 68 percent of the individuals in the sample self-identified as Catholic, agriculture and commerce were the most represented sectors of activity, the average expenditure per household/day was 6.5 USD, and 51/61 percent of the households owned land/house. In terms of voting patterns in control locations, we observe that on average 68 percent of registered voters actually voted in the March parliamentary elections, and that on average 37 percent voted for the (presidential) incumbent's party.

3.2 Vote-buying Survey Outcomes

We now turn to the analysis of the impact of our voter education campaign on vote-buying outcomes. We focus on reported perceptions and experiences of our survey respondents about vote-buying practices in their neighbourhoods or villages.

Vote-buying was referred to in our survey instruments using the word 'banho' (literally, bath or shower in Portuguese). 'Banho' is commonly used in STP to signify vote-buying. It refers to cash

[12] We also verified that the characteristics of the panel drops do not coincide with the few unbalanced characteristics. In order to account for panel attrition and a small number of missing observations, we re-ran our main empirical results using multiple imputation methods. We found similar results, which are available in the online Appendix.

or other gifts received from candidates or their representatives during electoral campaigning. Although we employ reported perceptions about vote-buying in the respondent's neighbourhood or village, we have also asked about vote-buying offers to personal acquaintances of surveyed subjects. While 66 percent of respondents mentioned having witnessed vote-buying in their location, only 34 percent mentioned that an acquaintance had been offered a gift for his/her vote (averages for the control group). This large difference is likely due to the immoral/illegal connotation vote-buying practices have in STP. We therefore resort to the former in the analysis that follows.[13] We have also asked control respondents who reported vote-buying offers to people in their social networks whether or not those offers were accepted: almost all (90 percent) reported acceptance. This is evidence consistent with clear availability of bribable votes: whoever is asked to sell his/her vote generally yields, even if that does not correspond to different voting behaviour.

With respect to the enforcement of vote-buying transactions, we asked about the prevalence of known techniques used by buyers of votes to ensure the agreed voting action actually happens. Namely, buying identity cards without which a voter cannot vote (i.e., precluding the supporters of the contender from voting), asking for a photograph of the filled in ballot paper at the ballot station, and substituting blank ballot paper (received by the voter at the ballot station) for pre-filled ballot paper provided by buyers of votes (this method requires access to official ballot paper from candidates).[14] However, we find that these techniques had limited use: as a percentage of the control respondents reporting offers to personal acquaintances, only 12 percent reported the use of these methods. This is consistent with the idea that self-enforcement may be the main mechanism by which vote-buying transactions have been enforced in STP.[15]

In our analysis we focus on five survey outcome variables relating to vote-buying. The first two are related to the perceived effectiveness of vote-buying in driving votes. First, we asked about the reported perception that vote-buying was driving voter behaviour, our main target in terms of

[13] Indeed, recent work by Gonzalez-Ocantos et al. (2012) show in a list experiment in Nicaragua that self-reported experienced vote-buying survey questions (close to our question on offers to personal acquaintances of surveyed subjects) underestimate actual vote-buying experienced. This list experiment is similar in spirit to our experimental measurement of conformity at the baseline, as both elicited experimental measures of response biases.

[14] Several reports on the 2006 parliamentary elections included references to the use of these techniques to ensure that vote-buying transactions were enforced. Namely observers reported that cell phones with camera devices had been supplied by vote-buyers outside the ballot stations in order for the voters to take photographs of the filled in ballot paper. Note that both photos and blank ballot paper were serving as proof that the agreed voting action was taken.

[15] Finan and Schechter (2012) corroborate the existence of self-enforcement of vote-buying in Paraguay.

voter education campaigning. The specific question asked was 'How frequent was it that 'banho' decided a voter's vote in your village/neighbourhood (in the March parliamentary or July presidential elections)?' Second, we asked a question closely aligned with the slogan of the voter education campaign: whether voting was happening in good conscience. This can be seen as analogous to the last question, but now indirectly referring to vote-buying, and asked in a positive manner. The specific question asked was 'How frequent was it that a voter in your village/neighbourhood voted following his/her conscience, i.e., voted for the candidate that offered better perspectives for the country (in the March parliamentary or July presidential elections)?' Note that for both of these questions, possible answers fitted a 7-point subjective scale. We label these measures as 'vote-buying impact on voting' and 'voting in conscience' and present them as z-scores for readability. Z-scores are calculated by subtracting the control group mean and dividing by the control group standard deviation. Thus, each component of the index has mean 0 and standard deviation 1 for the control group.

The remaining three vote-buying measures concern the occurrence of vote-buying. First, we asked about the frequency of witnessing vote-buying transactions in the respondent's neighbourhood or village. The specific question used was 'How frequently did you witness any 'banho' in your neighbourhood/village in the (March parliamentary or July presidential) elections?' Answers were coded according to an 8-point subjective scale. These are translated to z-scores to compose our measure labelled 'vote-buying frequency'. Second, we asked respondents about which candidacy was responsible for the witnessed vote-buying transactions. The specific question was 'Which candidates offered a 'banho'?' The answers to this question are coded in two binary variables, one concerning each of the main presidential candidates, and labelled 'vote-buying by incumbent/challenger'. Importantly, experimental subjects state that both candidates were active buying votes in their neighbourhood or village during the period of the electoral campaign: 91 percent of control respondents witnessing vote-buying in their neighbourhood or village report the presence of both candidates; the level-difference between the prevalence of vote-buying by the two candidates is not statistically different from zero.[16] Finally, we asked about the price of a vote (a 'banho') in the respondent's neighbourhood or village. The specific question used was 'On average, how much have parties/candidates spent for a vote when offering a 'banho'?' This outcome is labelled 'vote price' in the analysis that follows. We found median prices reported by control respondents who witnessed vote-buying in their neighbourhood

[16] This finding reinforces the neutrality of the campaign, seeing it not as an intervention set against any specific candidate.

or village to be 6.7 and 4.2 USD in the parliamentary and presidential elections, respectively. These are high prices for a country with monthly earnings of just over 100 USD per capita.

In the regressions reported in Tables 2a and 2b we begin by providing average treatment effects using post-election data only, without and with demographic controls[17], as per specifications (1) and (2) above. When pre-election data are available (not the case of vote-buying by specific candidates), we show estimates employing specification (2) with the baseline value of the dependent variable as additional control, and employing difference-in-differences specification (3) with demographic controls. In the regressions reported in Table 2c, on vote prices, we focus on our main specifications with controls - (2) and (3) - and add estimates using a Two-step Heckman selection model. This is because a non-trivial proportion of the sample is lost for this outcome variable due to no vote-buying witnessed in the neighbourhood or village, or to non-response. Our excluded variables that are added to the selection equation (on whether the respondent provides a price or a price difference) are individual measures of interest about politics[18].

< Table 2 around here >

We find clear effects of the voter education campaign against vote-buying on vote-buying outcomes. Namely, there is a large effect on the reported perceptions that vote-buying affected voting or that voting happened in good conscience: the effectiveness of vote-buying seems to have decreased as a result of the campaign. Those are the clearest effects of the campaign. These effects range from 0.42 to 0.46 standard deviation units for decreasing the vote-buying impact on voting, and from 0.32 to 0.48 standard deviation units for increasing voting in conscience. They are statistically significant at the 1 percent level using cluster-robust inference in all specifications shown. We find the same levels of statistical significance using wild bootstrap inference with the exception of the single-difference specification without controls for voting in conscience. We observe a negative effect on the frequency of vote-buying, ranging from 0.17 to 0.22 standard deviation units when employing single-difference specifications. This result is consistent with a

[17] Demographic controls follow Table 1 and include age, gender, household size, marital status dummies, education, number of children, nationality, ethnic group, and religion dummies, occupation dummies, expenditure, and property dummies.

[18] Specifically, the questions were on political activity in the last two years, the degree of interest about the 2006 presidential elections, and the degree of participation in the 2006 presidential elections. These measures of interest about politics are assumed to be related to whether one knows about the price of votes during the electoral campaign, but to have no direct relationship to the specific price.

decrease in the number of bribable voters. Statistical significance is achieved by both methods of inference. However, this effect is not statistically significant when using time-differences. We also document a negative impact on vote-buying activities by both candidates. This impact is more severe and more statistically significant for the challenger: the treatment implied a decrease of 8 to 9 percentage points in the likelihood of vote-buying by the challenger, which compares to 6 percentage points for the incumbent. Finally, we detect a negative effect of the campaign on the price of votes, ranging from around 12 to 18 USD, which is robustly significant when using time-differences.

3.3 Official Electoral Outcomes

We now present our main results, on the effects of the voter education campaign against vote-buying on actual voter behaviour, as determined by electoral results at the ballot station level. In Table 3 we look at voter turnout and candidate selection, as revealed by the score of the incumbent and the score of the challenger. We use OLS regressions following level specifications, with and without controls at the level of the location. Namely, we employ district dummies, number of registered voters, a dummy for urban locations, and time spent by international electoral observers.

< Table 3 around here >

We find a very clear pattern of effects, as all treatment effects estimated are significant at least at the 10 percent level (with location controls, all effects are significant at the 5 percent level). First, the campaign caused a decrease in voter turnout, on the order of 3 to 6 percentage points. This evidence is consistent with a positive effect of vote-buying on voter participation: as the campaign achieved a reduction in the effectiveness of vote-buying in driving voting behaviour, we observe the opposite effect, a reduction in voter turnout. Second, the campaign caused a shift of votes toward the incumbent and from the challenger, as revealed by the increase in the incumbent's score and reduction in the challenger's share of the vote, both estimated robustly at 4 percentage points. These findings are consistent with the notion that the challenger relies more strongly on vote-buying as a means of campaigning. They also add to the previous finding that the campaign had a greater effect on reducing vote-buying by the challenger than vote-buying by the incumbent.

3.4 Political Behaviour Individual Outcomes

We now turn to analysing political outcomes gathered in the context of the panel surveys we conducted. Table 4 depicts treatment effects of the voter education campaign on voter turnout as reported in the surveys, on the postcard outcome (i.e., the behavioural variable measuring whether each respondent sent the postcard demanding that the media reports on perceived corruption in the country), and on candidate selection by the respondents. We first employ specifications on levels, without and with demographic controls, i.e., specifications (1) and (2) above. These are all we have for the postcard outcome, for which we do not have a baseline measurement. When the baseline is available, namely in terms of intended turnout and candidate selection in the presidential elections of July 2006, we employ the level specification with demographic controls plus the baseline value of the dependent variable and the full difference-in-differences specification, i.e., specification (3) above. Note that all level variables are defined as binary.

< Table 4 around here >

The findings generally confirm those obtained for the electoral results at the ballot station level. We find that the campaign impacts negatively on voter participation: we observe an effect of 4-5 percentage points when employing level specifications, significant with both cluster-robust and wild-bootstrap statistical inference. This finding is consistent with the estimated effect when using ballot station results. Note, however, that statistical significance is lost when using time-differences. We also identify a positive impact on voting for the incumbent and a negative impact on voting for the challenger. These effects are robust to difference-in-differences and to wild-bootstrap inference, and are larger than the effects estimated with actual electoral results (both effects are around 6-7 percentage points in size), which is consistent with the fact that the campaign was centred on our sample of survey respondents. A final note goes to the postcard variable, the measure of political behaviour that we interpret as representing the demand for political accountability. We find no statistically significant results, even though a positive effect of the voter education campaign on this outcome variable is apparent. We cannot, however, claim a positive effect of voter education on political accountability on the basis of this evidence. It may be because the campaign was not powerful enough to go a step beyond vote-buying and communicate the need to base one's voting on a demand for political accountability, namely regarding corruption in the public services.

3.5 Conformity Effects

We now verify whether survey conformity biases existed, i.e., relating to the possibility that respondents in treatment locations biased their survey responses in order to adapt to whatever they perceived to be the views of the sponsors of the voter education campaign. Our experimental and survey design allowed us to gather a measure of survey conformity in the treatment effects of the campaign on vote-buying and voting outcomes. We purposely asked our pre-election questions about politics after the treated subjects were given the leaflets (and corresponding discussion). If we look at reported perceptions and experience (in terms of vote-buying and voting) pertaining to the March parliamentary elections, and compare treatment and control groups, we are able to measure whether there were undue reactions to the treatment, i.e. conformity biases. Since these outcomes relate to past events, there should be, on average and by design, no differences between the two comparison groups. This is despite the fact that the corresponding survey questions were asked after respondents in treatment locations were confronted with the voter education campaign.

In Table 5 we show all survey outcomes of interest as measured in the pre-election survey when asking about the March parliamentary elections. We ran regressions on treatment using these pre-election observations only, and generally found no statistically significant differences between treatment and control groups at the baseline. This is with the exception of voting in conscience, for which we see a negative impact of the treatment, which is significant at only the 10 percent level and is not robust to wild-bootstrap statistical inference. Note that this outcome is the one directly relating to the slogan of the campaign, and consequently the most likely to be affected by respondent biases. We can then conclude that we do not identify clear conformity biases for the vote-buying or voter behaviour survey outcomes, apart from the reported perception of voting in conscience, which stands as marginally significant. Note that this assertion relies on assuming that conformity biases were similar in the post-election survey.

< Table 5 around here >

3.6 Contamination

Finally, we test the robustness of our main results, on vote-buying and political behaviour outcomes, to treatment contamination of control locations (treatment spillovers) by running regressions of our outcomes on distance to closest treatment location employing control observations only. Table 6 reports our regressions of the main outcomes of the paper, i.e., survey perceptions of vote-buying and voter behaviour at the ballot station level, showing level specifications with controls. We observe clear evidence of contamination for candidate selection, as closest locations voted for the incumbent (relative to the challenger) in larger numbers. However, we do not identify significant effects on any other outcomes. We conclude for some evidence in favour of positive contamination of control locations in terms of impact on voting behaviour, specifically on candidate selection, which may be a sign of underestimation of our average treatment effects.

< Table 6 around here >

Note that generally we cannot rule out Stable Unit Treatment Value Assumption (SUTVA) violations, in that politicians could have reallocated vote-buying efforts toward control locations (in case they believed vote-buying became less effective in treatment locations). However, this is unlikely to have happened given the fact that we only encountered evidence in favour of positive contamination of the treatment to control locations. Moreover, the list of treatment locations was never public at the time of the experiment.

4 Interpretation

This paper analysed the impact of a voter education campaign against vote-buying in the 2006 presidential elections in STP. We found that the reported perception that voters voted according to money received was strongly diminished. The frequency and the price of vote transactions also appeared smaller as a result of the intervention. In terms of impact on voting behaviour, turnout was negatively affected (decreased by 3-6 percentage points), the incumbent was favoured (his score increased by 4 percentage points), and the challenger was harmed (his score decreased by 4 percentage points). We could not document a clear increase in the demand for political accountability. If we interpret the campaign simply, as an exogenous shift on vote-buying effectiveness, we can infer that vote-buying induces voters to participate, and that the challenger is relatively more reliant on that electoral strategy.

We turn now to the interpretation of these empirical results. We first explain the impact of the campaign on vote-buying outcomes.[19] We then proceed with the interpretation of the impact on the voting outcomes, i.e., on turnout and candidate selection. Vicente (2013) formalizes this discussion, while analysing a simple vote-buying game with an incumbency advantage.

The results on vote-buying outcomes are consistent with the notion that the effects of the voter education campaign on vote-buying happened primarily through convincing voters not to vote according to the money they received - in line with the verbal message of the voter education. This effect represents a clear decrease in the effectiveness of vote-buying. We now turn to the less clear effects on (i) the frequency of vote-buying and (ii) the price of the votes. The fact that, in poverty-ridden STP, the availability of voters for vote-buying is large (as measured by our survey questions on acceptance of gifts, mentioned above) leads us to think that any change in the frequency of vote-buying, i.e., on the acceptance of vote-buying offers, as a result of voter education is likely to be small. However, we do find some evidence of an effect of the campaign on decreasing the frequency of vote-buying. Let us call this a supply (of votes for transaction) shock. Then, if we accept this supply shock, likely induced by the decrease in the number of bribable voters, we should expect an increase in the price of votes. This is simply driven by having the same money available to fewer voters. In fact, what we find is a small reduction in price (our least significant result), which can only be explained by the demand (of votes) by politicians: the reduction in the effectiveness of vote-buying may have led to less money being directed by candidates for vote-buying, as they opt for more effective means of campaigning. The price of votes then decreases because less money is available to the same voters.

Our interpretation of the results on voter turnout is simple once we understand the impact on vote-buying. If we assume a standard cost of voting, swing (indifferent) voters abstain if not given money to vote for a candidate. Then, in the event that the effectiveness of vote-buying decreases (i.e., fewer voters taking into account the money they received from candidates when

[19] It is not the purpose of this paper to explore the correlates of vote-buying. However, the mention of some empirical patterns in our data may help to give a more complete picture of our vote-buying results. We find positive correlations of vote-buying frequency with competitive and rural locations (results available in the online Appendix). This indicates that bribable voters may be more easily found in competitive areas (also consistent with classical swing voter theories - see for instance Lindbeck and Weibull, 1987, and Dixit and Londregan, 1996), and that vote-buying effectiveness may be higher in rural areas (enforcement of vote-buying may be easier in less anonymous locations). We also observe that more schooled and richer households are less responsive to the voter education campaign (results available in the online Appendix). If we believe that reported perceptions about the neighborhood/village translate personal experiences to an extent, this observation is consistent with the robust association in the literature (see for instance Brusco et al., 2004) between less schooled/poorer voters and the acceptance of vote-buying offers.

deciding whether to turn out), or that the supply of bribable voters decreases (the possible effects of the campaign), more swing voters will abstain.

The effects on voting outcomes are more complex: our interpretation requires assuming an incumbency advantage. Collier and Hoeffler (2009) document that incumbents win most elections when there is electoral misbehaviour or when they are held in Sub-Saharan Africa – both conditions are satisfied in the context of this paper. Specifically, we assume that the incumbent was able to secure a large base before the start of the electoral campaign (and vote-buying happens). We recall that, in the context of the 2006 presidential elections in STP, PT's candidacy was announced just over a month before the presidential elections. The whole of his campaigning efforts occurred thereafter.[20] This is in contrast with FM's candidacy. FM had been the most powerful political personality in the country since 2001, having faced three elections before the one in July 2006, while leading MDFM. FM likely prepared his candidacy during his first mandate, namely through clientelism, offering public sector favours in exchange for political support. The presidential elections in STP 2006 may then have embedded a clear incumbency advantage. If this is true, we believe there were many voters supporting FM independently of what happened during the campaign. The swing-voter battle is all that matters for vote-buying. But if, as a consequence of the voter education campaign, the effectiveness of vote-buying or the supply of bribable voters decrease, fewer swing voters will vote for the challenger. If we credibly assume the voter education to be neutral (i.e., leading to proportional changes in the shares of swing voters voting for the incumbent and for the challenger), the constant base voters of the incumbent mechanically explain a higher vote share for the incumbent.

In summary, our empirical findings are consistent with an effective voter education campaign: one that was able to increase voting in good conscience, thereby disrupting the transactions of money for actual votes. The campaign also seems to have decreased the acceptance of money by impoverished voters, but not as clearly. In our interpretation, both shocks led to having more swing voters who were not driven by vote-buying, i.e., more swing voters remaining at the status quo (abstention). That explains the effect on increasing turnout. In the presence of an incumbency bias, having a lower number of swing voters casting ballots means the bias favouring the incumbent is more influential. That explains the increase in the incumbent's vote share.

[20] We find in our surveys that there is generally higher uncertainty associated with the challenger's candidacy by the population, both in terms of knowledge about the candidacy, and of the policies proposed.

5 Concluding Remarks

Our results provide good news for policy-makers interested in fighting against illicit electoral behaviour. We showed that voter education can work, i.e., that an information/sensitization campaign sponsored by an electoral commission was effective in diminishing the impact of electoral malfeasance.

In terms of welfare implications of these results, we should be cautious. On the one hand, we may argue that vote-buying leads to worse public policies. Vote-buying, as a form of redistribution, may be a substitute for public good provision. If vote-buying is the main means of accountability, and it is effective (as we infer) in driving votes, government action may forgo public goods, be tainted with corruption, and favour the political elite. On the other hand, we found that vote-buying understood as cash-for-votes is likely to benefit the challenger and increase political participation. Taken in isolation, these findings may be regarded as positive. Indeed, the incumbency advantage can then be counteracted, and greater electoral participation is not likely to harm democracy. However, improved electoral competition can only help development if there is real policy accountability, and it is unlikely that vote-buying helps policy accountability.

For future research, we would like to stress that despite the fact that we were only able to infer that vote-buying changes voting behaviour, we still know little about the mechanisms of enforcement of vote-buying transactions. Moreover, even though we measured conformity biases related to the voter education campaign, we need to improve the accuracy of our measures of vote-buying. We believe those, together with the empirical identification of the welfare implications of vote-buying, are the main challenges to take forward in the literature on vote-buying.

References

Acemoglu, D., and Robinson, J.A. (2001). 'Inefficient redistribution', *American Political Science Review*, vol. 95(3), pp. 649-661.

Becker, G.S. (1983). 'A theory of competition among pressure groups for political influence', *Quarterly Journal of Economics*, vol. 98(3), pp. 371-400.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). 'How much should we trust differences-in-differences estimates?', *Quarterly Journal of Economics*, vol. 119(1), pp. 249-275.

Besley, T. (2006). *Principled Agents? The Political Economy of Good Government*, Oxford: Oxford University Press.

Brusco, V., Nazareno, M., and Stokes, S.C. (2004). 'Vote-buying in Argentina', *Latin American Research Review*, vol. 39(2), pp. 66-88.

Cameron, A.C., Gelbach, J., and Miller, D. (2008). 'Bootstrap-based improvements for inference with clustered errors', *Review of Economics and Statistics*, vol. 90(3), pp. 414-427.

Coate, S., and Morris, S. (1995). 'On the form of transfers to special interests', *Journal of Political Economy*, vol. 103(6), pp. 1210-1235.

Collier, P., Hoeffler, A. (2009). 'Democracy's Achilles heel or, how to win an election without really trying', mimeograph.

Collier, P., and Vicente, P.C. (2012a). 'Violence, fraud, and bribery: the political economy of elections in Sub-Saharan Africa', *Public Choice*, vol. 153(1-2), pp. 117-147.

Collier, P., and Vicente, P.C. (2012b). 'Votes and violence: evidence from a field experiment in Nigeria', mimeograph.

Converse, P.E. (1972). 'Change in the American electorate', in (A. Campbell and P.E. Converse, eds.), *The Human Meaning of Social Change*, pp. 263-337, New York: Russell Sage Foundation.

Cox, G.W. (1987). *The Efficient Secret: The Cabinet and the Development of Political Parties in Victorian England*, Cambridge and New York: Cambridge University Press.

Cox, G.W., and Kousser, J.M. (1981). 'Turnout and rural corruption: New York as a test case', *American Journal of Political Science*, vol. 25(4), pp. 646-663.

Dal Bo, E. (2007). 'Bribing voters', *American Journal of Political Science*, vol. 51(4), pp. 789-803.

Dekel, E., Jackson, M.O., and Wolinsky, A. (2008). 'Vote-buying: general elections', *Journal of Political Economy*, vol. 116(2), pp. 351-380.

Dixit, A., and Londregan, J. (1996). 'The determinants of special interests in redistributive politics', *Journal of Politics*, vol. 58(4), pp. 1132-1155.

Duflo, E., Glennerster, R., and Kremer, M. (2007). 'Using randomization in development economics research: a toolkit', in (T.P. Schultz, and J. Strauss, eds.), *Handbook of Development Economics*, vol. 4, pp. 3895–962, Amsterdam: Elsevier.

Erikson, R.S., and T.R. Palfrey (2000). 'Equilibria in campaign spending games: theory and data', *American Political Science Review*, vol. 94(3), pp. 595-609.

Finan, F., and Schechter, L. (2012). 'Vote-buying and reciprocity', *Econometrica*, vol. 80(2), pp. 863-881.

Frynas, J.G., Wood, G., and Oliveira, R.S. (2003). 'Business and politics in São Tomé and Príncipe: from cocoa monoculture to petro-state', *African Affairs*, vol. 102(1), pp. 51-80.

Gerber, A.S. (1998). 'Estimating the effect of campaign spending on senate election outcomes using instrumental variables', *American Political Science Review*, vol. 92(2), pp. 401-411.

Gerber, A.S. (2004). 'Does campaign spending work? Field experiments provide evidence and suggest new theory', *American Behavioral Scientist*, vol. 47(5), pp. 541-574.

Gerber, A.S., and Green, D.P. (2000). 'The effects of canvassing, telephone calls, and direct mail on voter turnout: a field experiment', *American Political Science Review*, vol. 94(3), pp. 653-663.

Gonzalez-Ocantos, E., de Jonge, C.K., Melendez, C., Osorio, J., and Nickerson, D.W. (2012). 'Vote buying and social desirability bias: experimental evidence from Nicaragua', *American Journal of Political Science*, vol. 56(1), pp. 202–217.

Green, D.P., and Krasno, J.S. (1988). 'Salvation for the spendthrift incumbent: reestimating the effects of campaign spending in house elections', *American Journal of Political Science*, vol. 32(4), pp. 884-907.

Grossman, G.M., and Helpman, E. (1996). 'Electoral competition and special interest politics', *Review of Economic Studies*, vol. 63(2), pp. 265-286.

Ichino, N., and Schundeln, M. (2012). 'Deterring or displacing electoral irregularities? Spillover effects of observers in a randomized field experiment in Ghana', *Journal of Politics*, vol. 74(1), pp. 292-307.

Jacobson, G.C. (1978). 'The effects of campaign spending in congressional elections', *American Political Science Review*, vol. 72(2), pp. 469-491.

Levitt, S.D. (1994). 'Using repeat challengers to estimate the effect of campaign spending on election outcomes in U.S. house', *Journal of Political Economy*, vol. 102(4), pp. 777-798.

Lindbeck, A., and Weibull, J.W. (1987). 'Balanced-budget redistribution as the outcome of political competition', *Public Choice*, vol. 52, pp. 273-297.

Moulton, B.R. (1990). 'An illustration of a pitfall in estimating the effects of aggregate variables on micro units', *Review of Economics and Statistics*, vol. 72(2), pp. 334-38.

Nichter, S. (2008). 'Vote buying or turnout buying? Machine politics and the secret ballot', *American Political Science Review*, vol. 102(1), pp. 19-31.

Philipson, T.J., and Snyder, J.M. (1996). 'Equilibrium and efficiency in an organized voting market', *Public Choice*, vol. 89(3-4), pp. 245-265.

Robinson, J.A., and Verdier, T. (2003). 'The political economy of clientelism', mimeograph.

Seibert, G. (2006). *Comrades, Clients and Cousins: Colonialism, Socialism and Democratization in São Tomé and Príncipe*, 2nd Edition, Leiden and Boston: Brill Academic Publishers.

Shaffer, F., (2007). *Elections for Sale: the Causes and Consequences of Vote Buying*, Boulder: Lynne Rienner Publishers.

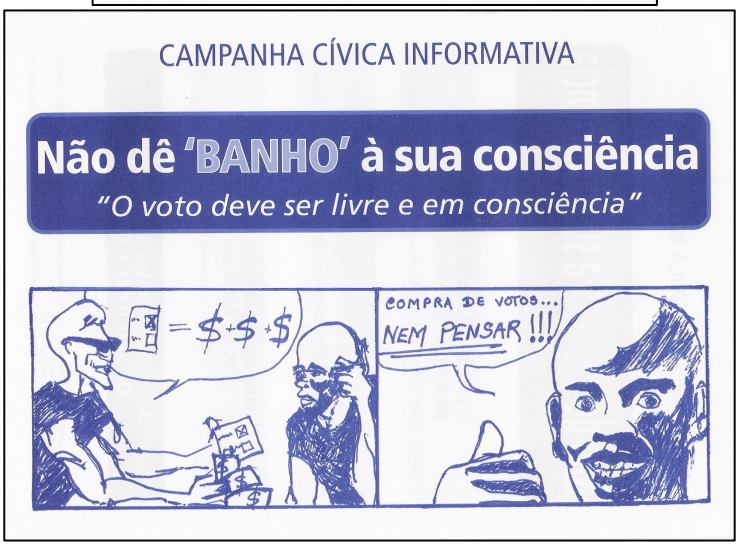
Stokes, S.C. (2005). 'Perverse accountability: a formal model of machine politics with evidence from Argentina', *American Political Science Review*, vol. 99(3), pp. 315-325.

Vicente, P.C. (2010). 'Does oil corrupt? Evidence from a natural experiment in West Africa', *Journal of Development Economics*, vol. 92(1), pp. 28-38.

Vicente, P.C. (2013). 'A model of vote-buying with an incumbency advantage', mimeograph.

Wantchekon, L. (2003). 'Clientelism and voting behaviour: evidence from a field experiment in Benin', *World Politics*, vol. 55(3), pp. 399-422.

Figure 1: The voter education campaign - leaflet (front and back)



Note: The main slogan is 'Do not let your conscience take a 'banho' - Your vote should be free and in good conscience'. The front page features three passages of the STP law (Constitution and Campaign Financing Law). The figure below presents a voter saying: 'Vote-buying… No way!!!'.

Figure 2: Map of Sao Tome Island – treatment and control areas

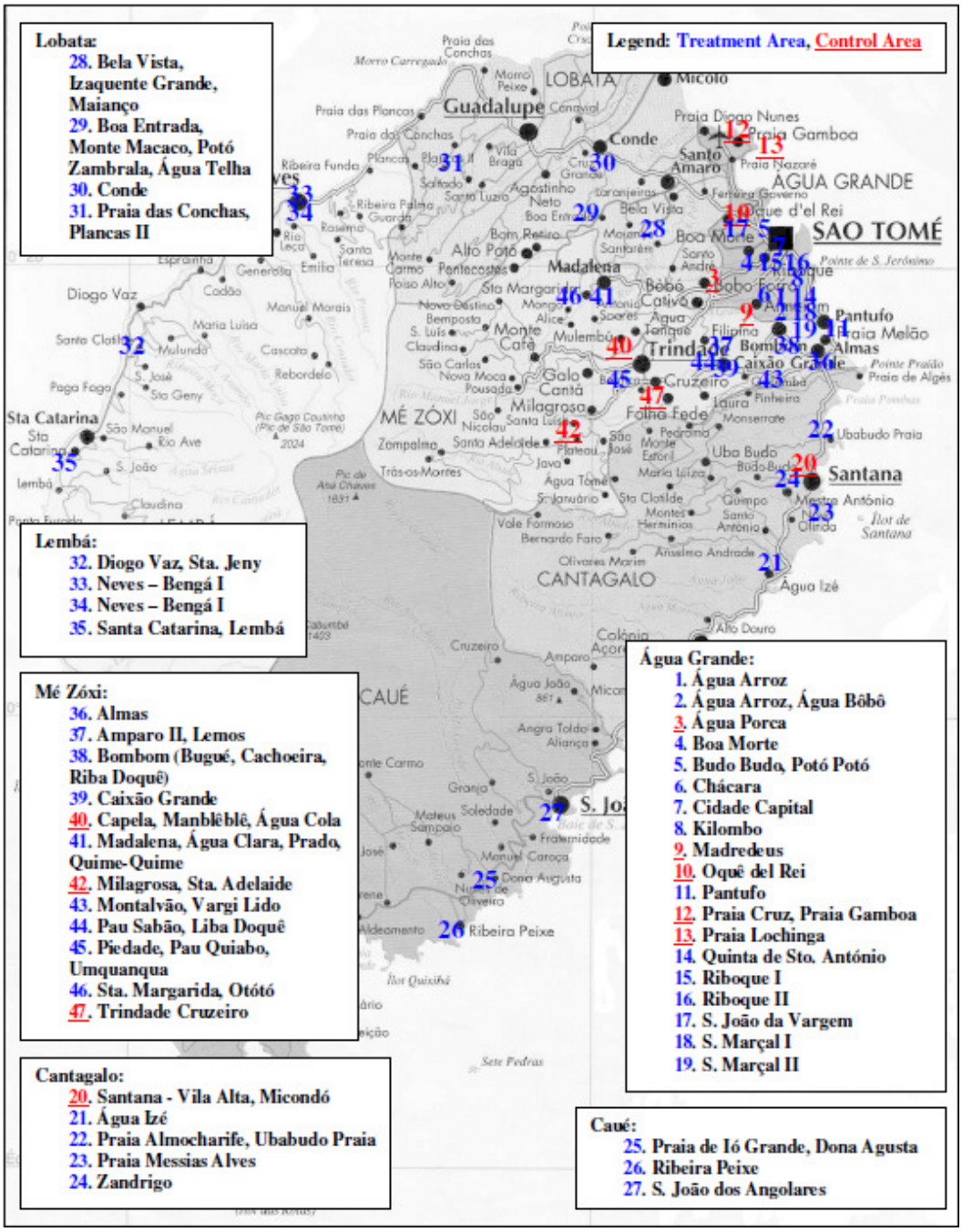


Figure 3: Map of Principe Island – treatment and control areas

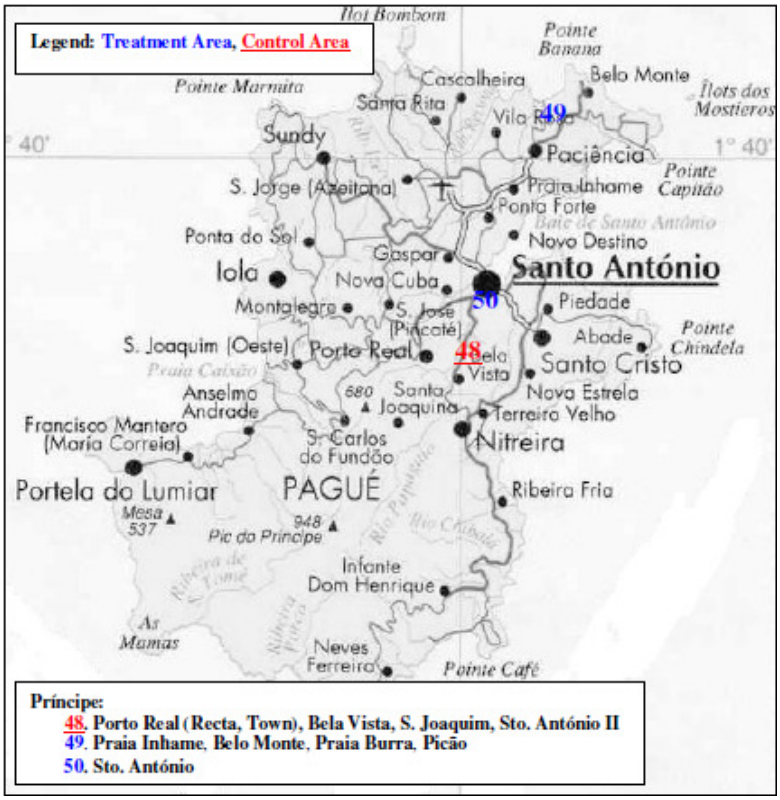


Figure 4: Timeline of the experiment in 2006

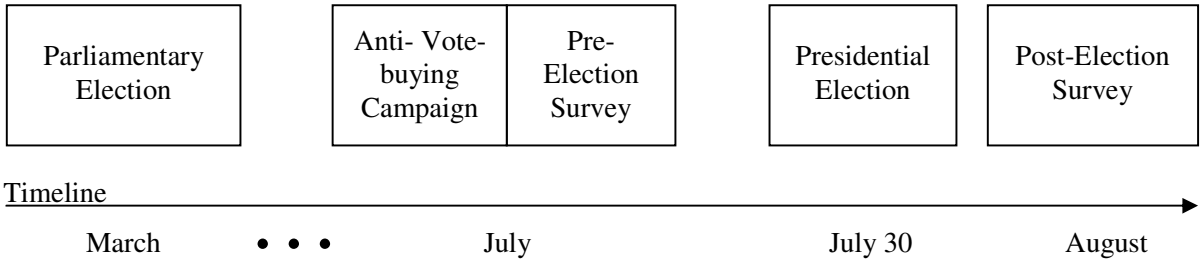


Table 1a: Differences across treatment and control locations (individual-level variables)

		baseline survey			post-election survey		
		control	treatment	difference	control	treatment	difference
basic demographics	age	37.746	37.448	-0.297 (1.634) 0.790	37.499	37.363	-0.136 (1.861) 0.906
	household size	5.187	5.272	0.085 (0.220) 0.696	5.376	5.336	-0.040 (0.196) 0.852
	single	0.290	0.280	-0.009 (0.048) 0.798	0.272	0.286	0.014 (0.047) 0.782
	unmarried couple	0.576	0.635	0.059 (0.049) 0.246	0.595	0.637	0.041 (0.056) 0.482
	widow	0.059	0.032	-0.027 (0.029) 0.544	0.059	0.031	-0.028 (0.037) 0.708
	married	0.036	0.029	-0.007 (0.014) 0.562	0.043	0.026	-0.017 (0.018) 0.354
	schooling over primary level	0.539	0.468	-0.071 (0.057) 0.266	0.564	0.481	-0.083 (0.052) 0.154
	number of children	4.043	4.082	0.039 (0.316) 0.992	3.878	3.998	0.121 (0.280) 0.666
	children in primary school	0.481	0.481	0.000 (0.044) 0.992	0.466	0.483	0.016 (0.046) 0.702
	children in secondary school	0.273	0.270	-0.002 (0.046) 0.984	0.277	0.262	-0.014 (0.049) 0.782
	malaria in the household	0.441	0.472	0.031 (0.048) 0.494	0.446	0.472	0.027 (0.048) 0.580
nationality, ethnic group, and religion	STP nationality	0.984	0.966	-0.018 (0.013) 0.162	0.980	0.964	-0.016 (0.015) 0.290
	CV nationality	0.019	0.027	0.008 (0.011) 0.462	0.023	0.030	0.006 (0.012) 0.594
	Forro	0.705	0.561	-0.145* (0.081) 0.098*	0.724	0.554	-0.171** (0.084) 0.070*
	Angolar	0.080	0.184	0.104** (0.042) 0.030**	0.065	0.188	0.123*** (0.045) 0.024**
	Contratado	0.042	0.045	0.003 (0.030) 0.938	0.047	0.046	-0.001 (0.032) 0.980
	Tonga	0.090	0.122	0.031 (0.022) 0.144	0.087	0.122	0.035 (0.023) 0.126
	Catholic	0.679	0.728	0.049 (0.050) 0.390	0.672	0.728	0.056 (0.051) 0.308
	non-religious	0.085	0.083	-0.002 (0.020) 0.940	0.097	0.080	-0.017 (0.021) 0.432

Note: Standard errors of the difference reported in parentheses; these are corrected by clustering at the location (census area) level. Wild bootstrap p-values shown below the standard errors. This method follows Cameron et al. (2008), with null hypothesis imposed, weights -1 and 1, and 1000 replications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 1b: Differences across treatment and control locations (individual-level variables)

		baseline survey			post-election survey		
		control	treatment	difference	control	treatment	difference
occupation	agriculture	0.091	0.104	0.013 (0.048) 0.822	0.096	0.112	0.016 (0.048) 0.772
	public administration	0.034	0.030	-0.004 (0.013) 0.818	0.029	0.032	0.003 (0.014) 0.836
	industry	0.035	0.043	0.008 (0.019) 0.712	0.036	0.039	0.003 (0.023) 0.978
	construction	0.034	0.050	0.016 (0.013) 0.234	0.041	0.047	0.005 (0.016) 0.708
	commerce	0.240	0.266	0.027 (0.037) 0.470	0.268	0.269	0.001 (0.044) 0.974
	transport	0.008	0.006	-0.002 (0.006) 0.740	0.004	0.005	0.001 (0.005) 0.876
	education	0.010	0.024	0.014* (0.008) 0.078*	0.012	0.025	0.013 (0.009) 0.160
	health	0.004	0.011	0.007 (0.005) 0.224	0.005	0.007	0.002 (0.005) 0.686
	household work	0.275	0.236	-0.038 (0.047) 0.502	0.286	0.245	-0.041 (0.049) 0.456
	unemployed	0.210	0.183	-0.027 (0.045) 0.522	0.173	0.177	0.004 (0.038) 0.918
expenditure and property	expenditure/day (USD)	6.492	6.463	-0.028 (0.569) 0.914	5.281	5.164	-0.117 (0.195) 0.576
	land	0.508	0.460	-0.048 (0.086) 0.586	0.517	0.480	-0.037 (0.086) 0.660
	house	0.611	0.535	-0.076 (0.075) 0.358	0.618	0.550	-0.068 (0.070) 0.386
	car	0.043	0.040	-0.004 (0.015) 0.852	0.051	0.042	-0.009 (0.019) 0.606
	cattle	0.391	0.425	0.034 (0.061) 0.552	0.372	0.435	0.063 (0.061) 0.300
	any property	0.727	0.780	0.052 (0.070) 0.470	0.730	0.796	0.065 (0.063) 0.306
	took a loan	0.161	0.133	-0.027 (0.042) 0.576	0.165	0.125	-0.040 (0.040) 0.402

Note: Standard errors of the difference reported in parentheses; these are corrected by clustering at the location (census area) level. Wild bootstrap p-values shown below the standard errors. This method follows Cameron et al. (2008), with null hypothesis imposed, weights -1 and 1, and 1000 replications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 1c: Differences across treatment and control locations (location-level variables)

		control	treatment	difference
official ballot records (parliamentary elections 2006)	turnout	0.680	0.658	-0.021
				(0.023)
	incumbent candidate's party (MDFM)	0.368	0.374	0.006
				(0.019)
	difference between incumbent and challenger candidates' parties (MDFM vs. ADI)	0.161	0.171	0.010
				(0.026)
attrition (panel re-surveying) (%)		0.833	0.811	-0.022
				(0.032)

Note: Standard errors reported; for ballot records, these are corrected by clustering at the ballot location level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2a: Regressions of vote-buying outcomes

dependent variable ------>		vote-buying impact on voting				voting in conscience			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatment effect	coefficient	**-0.496*** **	**-0.461*** **	**-0.457*** **	**-0.424*** **	**0.285*** **	**0.319*** **	**0.406*** **	**0.475*** **
	standard error	**(0.100)**	**(0.076)**	**(0.082)**	**(0.113)**	**(0.109)**	**(0.098)**	**(0.102)**	**(0.120)**
	p-value wild bootstrap	0.002***	0.002***	0.004***	0.006***	0.026**	0.004***	0.002***	0.002***
mean dep. variable (control)		0.000	0.017	0.024	0.009	0.000	0.004	0.001	0.010
r-squared adjusted		0.070	0.090	0.203	0.081	0.048	0.055	0.131	0.073
number of observations		993	966	951	951	1,008	957	950	950
basic specification		level	level	level cont. for baseline	time-difference	level	level	level cont. for baseline	time-difference
controls		no	yes	yes	yes	no	yes	yes	yes

Note: All regressions are OLS. Dependent variables are z-scores. All regressions include district dummies. Controls are individual demographic characteristics (see Table 1), which include age, gender, household size, marital status dummies, education, number of children, nationality, ethnic group, and religion dummies, occupation dummies, expenditure, and property dummies. Standard errors reported; these are corrected by clustering at the location (census area) level. Wild bootstrap method follows Cameron et al. (2008), with null hypothesis imposed, weights -1 and 1, and 1000 replications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2b: Regressions of vote-buying outcomes

dependent variable ------>		vote-buying frequency				vote-buying by incumbent		vote-buying by challenger	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatment effect	coefficient	-0.202**	-0.220***	-0.172**	-0.097	-0.047	-0.059*	-0.080**	-0.093***
	standard error	(0.087)	(0.078)	(0.087)	(0.109)	(0.039)	(0.033)	(0.041)	(0.035)
	p-value wild bootstrap	0.042**	0.014**	0.094*	0.444	0.254	0.078*	0.086*	0.022**
mean dep. variable (control)		0.000	0.002	0.017	-0.001	0.616	0.618	0.635	0.631
r-squared adjusted		0.024	0.035	0.085	0.030	0.028	0.043	0.021	0.042
number of observations		1,000	983	958	958	1,000	962	1,000	962
basic specification		level	level	level cont. for baseline	time-difference	level	level	level	level
controls		no	yes	yes	yes	no	yes	no	yes

Note: All regressions are OLS. Dependent variables are z-scores (1-4) and binary (5-8). All regressions include district dummies. Controls are individual demographic characteristics (see Table 1), which include age, gender, household size, marital status dummies, education, number of children, nationality, ethnic group, and religion dummies, occupation dummies, expenditure, and property dummies. Standard errors reported; these are corrected by clustering at the location (census area) level. Wild bootstrap method follows Cameron et al. (2008), with null hypothesis imposed, weights -1 and 1, and 1000 replications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2c: Regressions of vote-buying outcomes

dependent variable ------>		vote price			
		OLS		selection model	
		(1)	(2)	(3)	(4)
treatment effect	coefficient	-11.594	-144.762*	-13.605	-214.181**
	standard error	(13.757)	(85.508)	(14.232)	(102.623)
p-value wild bootstrap		0.426	0.088*		
inverse Mills ratio				-124.942**	269.802
				(48.984)	(259.980)
mean dep. variable (control)		82.323	-84.021	82.323	-84.021
r-squared adjusted		0.015	0.011		
number of observations		472	415	962	962
basic specification		level	time-difference	level	time-difference

Note: The dependent variable is monetary value. The selection model is a Heckman Two-step model. The selection equation is for having reported a price for the votes in the presidential elections (level) or the difference between prices in the parliamentary and presidential elections (time-difference) - excluded variables are individual measures of interest about politics. All regressions include district dummies and individual controls. Individual controls are demographic characteristics (see Table 1), which include age, gender, household size, marital status dummies, education, number of children, nationality, ethnic group, and religion dummies, occupation dummies, expenditure, and property dummies. Standard errors reported. For OLS, these are corrected by clustering at the location (census area) level. For OLS, wild bootstrap method follows Cameron et al. (2008), with null hypothesis imposed, weights -1 and 1, and 1000 replications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Regressions of electoral outcomes (official ballot records)

dependent variable ------>		turnout		voting for incumbent		voting for challenger	
		(1)	(2)	(3)	(4)	(5)	(6)
treatment effect	coefficient	-0.064***	-0.025***	0.035*	0.034**	-0.036*	-0.035**
	standard error	(0.017)	(0.009)	(0.021)	(0.017)	(0.020)	(0.017)
mean dep. variable (control)		0.732	0.732	0.581	0.581	0.414	0.414
r-squared adjusted		0.084	0.660	0.012	0.267	0.013	0.267
number of observations		228	228	228	228	228	228
controls		no	yes	no	yes	no	yes

Note: All regressions are OLS. Dependent variables are vote shares. Controls include district dummies, number of registered voters, urban dummy, and time spent by electoral observers. Standard errors reported; these are corrected by clustering at the ballot location level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4a: Regressions of political behavior (individual-level)

dependent variable ------>		turnout				postcard	
		(1)	(2)	(3)	(4)	(5)	(6)
treatment effect	coefficient	-0.041*	-0.054***	-0.042**	-0.020	0.005	0.028
	standard error	(0.023)	(0.014)	(0.018)	(0.034)	(0.056)	(0.050)
	p-value wild bootstrap	0.078*	0.002***	0.044**	0.622	0.922	0.624
mean dep. variable (control)		0.932	0.931	0.922	0.052	0.239	0.238
r-squared adjusted		0.015	0.092	0.132	0.050	0.060	0.060
number of observations		1,011	976	854	854	1,033	981
basic specification		level	level	level cont. for baseline	time-difference	level	level
controls		no	yes	yes	yes	no	yes

Note: All regressions are OLS. Dependent variables are binary or time-differenced binary. All regressions include district dummies. Controls are individual demographic characteristics (see Table 1), which include age, gender, household size, marital status dummies, education, number of children, nationality, ethnic group, and religion dummies, occupation dummies, expenditure, and property dummies. Standard errors reported; these are corrected by clustering at the location (census area) level. Wild bootstrap method follows Cameron et al. (2008), with null hypothesis imposed, weights -1 and 1, and 1000 replications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4b: Regressions of political behavior (individual-level)

dependent variable ------>		voting for incumbent				voting for challenger			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatment effect	coefficient	0.061	0.069*	0.061**	0.060*	-0.098**	-0.094**	-0.069**	-0.062**
	standard error	(0.044)	(0.038)	(0.029)	(0.032)	(0.047)	(0.040)	(0.028)	(0.027)
	p-value wild bootstrap	0.198	0.084*	0.054*	0.096*	0.056*	0.042**	0.034**	0.038**
mean dep. variable (control)		0.609	0.602	0.622	-0.041	0.318	0.324	0.295	0.088
r-squared adjusted		0.046	0.081	0.309	0.045	0.062	0.092	0.393	0.010
number of observations		1,011	976	854	854	1,011	976	854	854
basic specification		level	level	level cont. for baseline	time-difference	level	level	level cont. for baseline	time-difference
controls		no	yes	yes	yes	no	yes	yes	yes

Note: All regressions are OLS. Dependent variables are binary or time-differenced binary. All regressions include district dummies. Controls are individual demographic characteristics (see Table 1), which include age, gender, household size, marital status dummies, education, number of children, nationality, ethnic group, and religion dummies, occupation dummies, expenditure, and property dummies. Standard errors reported; these are corrected by clustering at the location (census area) level. Wild bootstrap method follows Cameron et al. (2008), with null hypothesis imposed, weights -1 and 1, and 1000 replications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Treatment conformity bias: regressions of baseline survey outcomes (parliamentary elections)

dependent variable ------>		vote-buying impact on voting	voting in conscience	vote-buying frequency	vote-buying by MDFM	vote-buying by ADI	vote price	turnout parliamentary elections	voting for MDFM	voting for ADI
		(1)	(2)	(3)	(4)	(5)	(6)	7)	(8)	(9)
treatment effect	coefficient	**-0.020**	**-0.275***	**-0.130**	**-0.072**	**-0.079**	**127.589**	**-0.009**	**0.004**	**0.011**
	standard error	**(0.144)**	**(0.143)**	**(0.103)**	**(0.050)**	**(0.060)**	**(131.855)**	**(0.032)**	**(0.033)**	**(0.016)**
	p-value wild bootstrap	0.932	0.126	0.220	0.170	0.212	0.370	0.798	0.826	0.546
mean dep. variable (control)		0.009	-0.008	0.005	0.821	0.767	241.407	0.873	0.571	0.085
r-squared adjusted		0.007	0.050	0.069	0.067	0.086	0.095	0.064	0.058	0.041
number of observations		959	951	1,209	1,184	1,184	902	1,187	1,187	1,187

Note: All regressions are OLS. Dependent variables are z-scores (1-3), binary (4-5 and 7-9), and monetary value (6). All regressions regard level specifications. They include district dummies and individual controls. Individual controls are demographic characteristics (see Table 1), which include age, gender, household size, marital status dummies, education, number of children, nationality, ethnic group, and religion dummies, occupation dummies, expenditure, and property dummies. Standard errors reported; these are corrected by clustering at the location (census area) level. Wild bootstrap method follows Cameron et al. (2008), with null hypothesis imposed, weights -1 and 1, and 1000 replications. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Regressions of main outcomes on distance to closest treatment area (control observations only)

dependent variable ------>		vote-buying impact on voting	voting in conscience	vote-buying frequency	vote-buying by incumbent	vote-buying by challenger	vote price	turnout	voting for incumbent	voting for challenger
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
distance to treatment effect	coefficient	-0.010	0.012	0.019	0.013	0.016	3.269	0.004	-0.018***	0.017***
	standard error	(0.039)	(0.028)	(0.029)	(0.012)	(0.015)	(9.456)	(0.003)	(0.004)	(0.004)
	p-value wild bootstrap	0.803	0.723	0.597	0.357	0.333	0.695			
mean dep. variable (control)		0.017	0.004	0.002	0.618	0.631	82.323	0.732	0.581	0.414
r-squared adjusted		0.027	0.039	0.012	0.006	0.007	-0.001	0.509	0.333	0.332
number of observations		211	216	218	217	217	99	121	121	121

Note: All regressions are OLS. Dependent variables are z-scores (1-3), binary (4-5), monetary value (6), and vote shares (7-9). Regressions on voter behavior employ the data from official voting records. All regressions regard level specifications. They include district dummies and individual or location controls. Individual controls are demographic characteristics (see Table 1), which include age, gender, household size, marital status dummies, education, number of children, nationality, ethnic group, and religion dummies, occupation dummies, expenditure, and property dummies. Standard errors reported; these are corrected by clustering at the location (census area) level. Wild bootstrap method follows Cameron et al. (2008), with null hypothesis imposed, weights -1 and 1, and 1000 replications. * significant at 10%; ** significant at 5%; *** significant at 1%.