

I sell my vote, and so what?

*Incidence, social bias and correlates of clientelism in Colombia**

Leopoldo Fergusson[†] Carlos Molina[‡] Juan Felipe Riaño[§]

June 30, 2018

Abstract

Exchanging one's vote for particularistic benefits – practices usually grouped under 'clientelism' – is often thought to weaken programmatic links between citizens and politicians and disincentivize public good provision, as well as undermine voter autonomy and the ideal role of elections. However, empirically analyzing this key phenomenon for the working of democracies entails formidable challenges. We conduct list experiments on a large sample of households to estimate the incidence of clientelistic vote buying, as well as the extent to which respondents refrain from openly recognizing this behavior. Nearly one out of every five respondents engage in clientelism and, surprisingly, they do not feel ashamed to admit it. Guided by the existing literature and systematically verifying the sensibility of the results to model specification, we examine the robust correlates of clientelism and discuss the implications of our key findings.

Keywords: Clientelism, vote buying, social desirability bias, list experiments.

JEL: C83, C93, D72, D73.

*We thank the entire team from the *Encuesta Longitudinal Colombiana de la Universidad de los Andes*, particularly its academic committee. We are especially indebted to Ximena Cadena, the project director during the design and implementation of the experiment described in this paper, for both successfully leading the survey and for numerous comments and suggestions on the survey design and our main findings. We also thank Kelley Friel and two anonymous referees for detailed comments.

[†]Universidad de los Andes, Department of Economics, Carrera 1 N° 18A - 12 Bogotá, Colombia. lfergusson@uniandes.edu.co, corresponding author.

[‡]Universidad de los Andes, Department of Economics, Carrera 1 N° 18A - 12 Bogotá, Colombia. ca.molina1@uniandes.edu.co.

[§]University of British Columbia, Vancouver School of Economics, 6000 Iona Drive Vancouver, Canada. jf.riano@alumni.ubc.ca.

1 Introduction

This paper examines the exchange of votes in elections for targeted transfers or benefits – practices usually grouped under ‘clientelism’. This phenomenon has fundamental economic and political implications. Clientelistic vote buying is key to understanding the relationship between citizens and the state in a democracy, influences economic policies and outcomes, and shapes incentives for corruption.

The prevalence of these practices is thought to be prejudicial for democracy, as it undermines more programmatic links between citizens and politicians. The standard view in political science and economics emphasizes that, with clientelism, politicians focus on providing particularistic benefits for powerful minorities rather than public goods for the general welfare (Bates, 1981; Kitschelt, 2000; Stokes, 2005, 2007). In addition, in a clientelistic exchange receiving a particularistic benefit is contingent on political support, and the benefit can be withdrawn if such support is lacking (Stokes, 2007; Hicken, 2011).¹ This contingency undermines elections as instruments of representation, and since immediate material benefits may be especially pressing for vulnerable voters, it also endangers equality of political rights (Stokes, Dunning, Nazareno, & Brusco, 2013) and traps vulnerable voters in these relationships (Bobonis, Gertler, Gonzalez-Navarro, & Nichter, 2017). Finally, by relying on public funds for the reproduction of the clientelistic network, clientelism can also incentivize arbitrary rules of redistribution and corruption in the public sector (Stokes et al., 2013; Maiz & Requejo, 2001; Singer, 2009; Camacho & Conover, 2011).

Despite the importance of this phenomenon, examining it empirically is challenging. Clientelism illustrates one major obstacle to empirical research in the social sciences: eliciting honest answers in surveys, especially on sensitive topics like discrimination, corruption, and illegality. Respondents may avoid truthful answers when asked, and the nature of these behaviors implies that there are often no other reliable records. Several survey techniques have been developed to address these difficulties. Among them, list experiments

¹As Stokes (2005) puts it, this is a “perverse accountability” system: it is not citizens who punish politicians when they fail to fulfill their promises and programs in office, but politicians who may punish citizens for not supporting them in the polls.

have received considerable attention and have been used in a variety of applications, usually as alternative to Randomized Response Techniques (RTT) (Imai, 2011).² This alternative, also known as the Item Count Technique (ICT) or the Unmatched Count, is based on the premise “that if a sensitive question is asked in an indirect fashion, respondents may be more willing to offer a truthful response even when social norms encourage them to answer the question in a certain way” (Blair & Imai, 2012).³

We apply list experiments to a large sample of households to explore the incidence and main correlates of clientelistic vote buying. Our data is drawn from the Politics Module of the *Encuesta Longitudinal Colombiana de la Universidad de los Andes (Elca)* (Bernal et al., 2014; Fergusson & Riaño, 2014). The Elca is the first large-scale household panel survey in Colombia, with roughly 10,000 households that are representative of urban Colombia and five rural macro regions. The baseline survey was conducted in 2010. In 2013, the first follow-up included the Politics Module with a number of questions on political participation and interest, sources of information, ideological positions, and the key sensitive items we focus on. The module was administered to one member of the household (the household head or his or her partner, with random assignment when both were available).

We pursue several objectives. First, to measure the incidence of both vote buying and the social stigma associated with it. Our list experiment, whose underlying assumptions we validate, quantifies vote buying while dealing with the potential biases in citizen responses. To evaluate the associated bias in citizen responses, we (randomly) divide our sample into respondents whose behavior is derived from the list experiments, and those that are directly asked about their experience with clientelism. The difference between these two

²These applications include discrimination against African Americans and other minority or marginalized groups, attitudes towards food, risky sexual behaviors, and other sensitive or illegal actions. List experiments have also been used to study electoral behavior (Gonzalez-Ocantos, de Jonge, Meléndez, Osorio, & Nickerson, 2012; Holbrook & Krosnick, 2010; Corstange, 2010, 2012), as we do in this paper. They have been recently used in Colombia to study support for certain groups, particularly military and rebel groups (Matanock & García-Sánchez, 2011a, 2011b; Steele & Shapiro, 2012), as well as clientelism Garcia and Pantoja (2015).

³A detailed explanation of the method and new developments can be found in Imai (2011) and Blair and Imai (2012). We discuss the details of our own implementation in Section 2.

sets of responses measures the extent of “social desirability bias”, and reflects how willing respondents are to admit to engaging in clientelism when asked indirectly as opposed to directly. In our setting, however, there is no difference: both methods suggest that almost one out of five respondents engage in clientelism. The lack of social desirability bias is consistent with vote buying carrying no social stigma in the population. This has important implications for the resulting obstacles to combat vote buying. Moreover, this lack of bias applies overall, as well as across a number of individual characteristics. This is important since, when present, social desirability bias is plausibly nonrandom, and analyses based on direct questions may produce different determinants than those based on list experiments (Gonzalez-Ocantos et al., 2012). On a more practical level, it is however good news for our empirical analysis, since we can confidently use answers to direct questions to examine clientelistic vote buying.⁴

Second, we use the resulting data to review the incidence and main correlates for clientelism, using the existing literature as a guide for our analysis. Third, we conduct a sensitivity analysis that helps us identify the variables that are robustly correlated with clientelism. This crucial step protects us against the ex-post rationalization and specification search. The results on robust (and non-robust) correlates inform our understanding of vote buying, have implications for the efforts to overcome clientelism, and the comparison of our findings with those reported in the literature also illustrates the richness of the dataset. Fourth, we also go beyond the robust determinants to discuss the broader implications on the potential detrimental implications of clientelistic vote buying for economic policy and performance by showing its correlation with municipal-level political features that might shape the quality of representation and policies.

Finally, for all these reasons our analysis can also be a first step for further analyses of these data. The data are freely available for download from the

⁴Multivariate methods based on list experiments are available (see Section 2.1), but it is more precise to rely on simple regressions of answers to the direct question. However, we also verified that the reported robust correlations with direct questions are in line with those using multivariate techniques for list experiments. Also, since the sample asked directly was randomly selected, and since we observe no differential attrition, we can use this sample to infer about general vote buying prevalence.

project’s website, at <https://encuestalongitudinal.uniandes.edu.co/en/>. The 2013 round contains the list experiment described in this paper, and baseline characteristics are also available for the 2010 panel.

The paper proceeds with a brief examination of the empirical methods. Section 2.1 describes the key questions and validates the underlying assumptions, and Section 2.2 explains the “extreme bounds methodology,” a sensitivity analysis approach we implement to assess the robustness of the correlations between clientelism and other variables. Section 3 reports the main results on the incidence of clientelism and documents the absence of social desirability bias in our survey responses. Section 4 presents the main findings, using the existing literature as a guide to study the most salient features of the data, and describing the robust correlates of clientelism. Section 5 reflects on the relevance of our findings on several dimensions. Section 6 concludes.

2 Empirical methods

2.1 List experiments to measure electoral clientelism

Our vote-buying list experiment seeks to capture the essential features of a clientelistic exchange where material benefits are traded for political support. Not all benefits can be considered clientelism, however. For example, a party may target a set of policies to benefit a particular ethnic minority in hopes of attracting greater support from this group. Stokes (2007) describes such targeted redistribution (or pork-barrel politics, where politicians redistribute benefits to a particular area, like a congressional district) as having “some public-good quality: they redistribute resources from classes of non-beneficiaries to classes of beneficiaries, but within a class of beneficiaries, particular people who qualify cannot be excluded.” Thus, our question emphasizes the *quid pro quo* nature of the exchange. Also, “support” may include voting for the patron or persuading others to do so, mobilizing acquaintances, working for the party’s campaign, etc. In our empirical measure, we focus on the specific action of exchanging benefits for one’s vote (vote buying) since it can be measured more accurately.

While citizens may interpret “giving their political support” differently, it is unlikely that this is the case when asked about “giving their vote.”

As [Nichter \(2014\)](#) notes, scholars mean different things when they refer to (clientelist) vote buying, thus confounding different phenomena with potentially different empirical relationships. Two key aspects of the differences in definitions of vote buying are the timing of the benefits (most notably, shortly before or during elections versus after elections) and the types of benefits (cash, or other goods and services). [Nichter \(2014\)](#) defines clientelist vote buying as “the distribution of rewards to individuals or small groups during elections in contingent exchange for vote choices” (p. 316). According to this definition, rewards are general in the sense that they could either be cash or other goods and services, but specific in that they must be provided during elections (excluding, for example, post-election benefits, jobs, or public programs). While we concur with Nichter’s broad definition of rewards, we disagree about the timing, as several benefits are delivered or available only after the election, even when support is narrowly defined as voting.⁵

Respondents are randomly assigned to various groups. Households in the *Treatment* group are told “I will read a list of five (5) things people have in mind when deciding who to vote for. I want you to tell me how many of these five things you have taken into account when voting for a candidate. Do not tell me WHICH, ONLY HOW MANY.”⁶ Then they are handed a card with the following options:

⁵There are other important distinctions. For example, the benefits exchanged may or may not come from public resources (like public jobs, access to government programs, or transfers of public funds). [Stokes \(2007\)](#) suggests referring to *patronage* as a type of clientelism that relies on government funds to deliver benefits. We cannot explore all these distinctions in a survey embedded in an already long and demanding questionnaire for households. We therefore focus on an important and concrete form of clientelistic vote buying, while amply defining the timing and sources of the funds to help capture its presence despite likely variations.

⁶In Spanish the script was: “Le leeré una lista de cinco (5) cosas que la gente tiene en cuenta para decidir por quién votar. Quiero que me diga cuántas de estas cinco ha tenido usted en cuenta para votar por un candidato. No me diga cuáles, únicamente cuántas.

1. La información sobre el candidato que usted oyó en la radio o televisión,
2. Lo que usted leyó en el programa de gobierno del candidato,
3. Los beneficios, regalos, o trabajos que el candidato le ofreció a usted a cambio de su voto,
4. Las conversaciones sobre el candidato que usted tuvo con amigos,
5. El partido del candidato.

1. The information about the candidate on the radio or television,
2. What you read about his or her government plan,
3. **The benefits, gifts, or jobs the candidate offered you in exchange for your vote,**
4. The conversations you had with your friends about the candidate,
5. The candidate's party.

In the first control group (*Control 1*), respondents are confronted with a similar prompt and list, except the sensitive item is not in the list (marked in bold above for emphasis, but not on the list used in the survey).⁷

The premise of the experiment is that when asking about vote buying indirectly by using a list, individuals are willing to answer truthfully even if social norms suggest that there is a “correct” answer. Since respondents in the *Treatment* condition only differ from those in *Control 1* in that they are presented with vote buying as an option in the list, the difference in the number of actions reported by the two groups estimates the proportion of individuals that has participated in vote buying.

Following the application of the list experiment, those in *Control 1* (who have not seen the sensitive item) are asked directly: “Could you tell me if, when deciding who to vote for, you have taken into account the benefits, gifts, or jobs that a candidate offered in exchange for your vote?” A third group, *Control 2* was not presented with the list experiment; respondents in this group were *only* asked this question directly.⁸

⁷The wording is similar to [Corstange \(2012\)](#) and [Carkoglu and Aytac \(2015\)](#), but implements a slightly different strategy. In the control group, these authors ask about each item individually. This has the advantage of more efficient multivariate estimation. However, the risk is that respondents answer differently when asked item by item than when asked about the full list, thus producing a design effect ([Corstange, 2009](#); [Blair & Imai, 2012](#)). We err on the safe side. Also, notice that since the respondent is asked on what factors shaped the decision to support (vote for) a particular candidate, with *exchange* of vote for benefits as one of the options, the question is unlikely to be interpreted differently by different respondents. Pilot testing followed by debriefing also showed that respondents interpreted the question as intended. Also, [Section 5](#) below shows our vote buying measure correlates intuitively with electoral practices akin to a very clientelistic pattern of political exchange, further validating our measure.

⁸Since we do not find significant differences in responses to the direct question between *Control 1* and *Control 2* in either experiment, we include both groups throughout when examining the direct questions.

The key assumption that individuals responding to the treatment and control questions are similar must hold in order for our exercises to be valid. While this should be the case by design since the groups were assigned randomly, we used the baseline survey from 2010 to verify balance on a number of observable baseline characteristics. Also, since randomization could fail in the field, we double checked balance on covariates in the follow-up 2013 survey when the list experiments were conducted. Online Appendix Tables A-2 to A-3 show that there are no systematic differences between the treatment and control groups, suggesting that their composition is unlikely to contaminate our results.⁹

Two additional assumptions underlie the ability of the list experiments to faithfully capture the behavior of interest. First, we require that there are no *design effects*¹⁰: adding the sensitive item must not affect the response on the *sum* of control items. Second, we require that there are no *sensitive liars*: respondents must answer truthfully to the sensitive item.

Formally, let there be $j = 1, \dots, J$ control items and one ($j = J + 1$) sensitive item, and let $Z_{ij}(t)$ indicate i 's implied answer for item j under treatment status $t = 0, 1$ (where $t = 0$ denotes the control list and $t = 1$ the treatment list).¹¹ Also, let $Y_i(0) = \sum_{j=1}^J Z_{ij}(0)$ and $Y_i(1) = \sum_{j=1}^{J+1} Z_{ij}(1)$ be the potential answer i would give under control or treatment list, respectively. Then there are no design effects if:

Assumption 1. (No design effects) For each $i = 1, \dots, N$,

$$\sum_{j=1}^J Z_{ij}(0) = \sum_{j=1}^J Z_{ij}(1) \text{ or, equivalently, } Y_i(1) = Y_i(0) + Z_{i,J+1}(1).$$

⁹In 2010, the ELCA surveyed a total of 9,830 households, and these were randomly allocated into the three treatment groups for the second wave in 2013: 3,535 to the *Treatment*, 3,146 to *Control 1*, and 3,149 to *Control 2*. When surveying households in 2013, the sample fell to 9,234, but attrition was balanced across groups and, consistent with this, Appendix Table A-2 shows the balance on covariates persisted. When further restricting the sample to keep households with complete information on the covariates, we have 7,940 households divided into three groups of 2,856, 2,564 and 2,520 households. Appendix Table A-3 shows these three groups are also balanced on covariates.

¹⁰To spell out the main assumptions for the validity of list experiments, we follow Blair and Imai (2012).

¹¹Notice also that $Z_{ij}(t) \in \{0, 1\}$

Notice that this does not require individuals to answer truthfully to control items, only that the answer *to the sum of control elements* does not change when confronted with the treated and control prompts.¹²

The no liar condition is:

Assumption 2. (No liar) For each $i = 1, \dots, N$,

$$Z_{i,J+1}(1) = Z_{i,J+1}^*$$

where $Z_{i,J+1}^*$ denotes the latent response to the sensitive item in the list.

While the validity of these assumptions is not directly testable, Blair and Imai (2012) propose hypotheses testing procedures that we perform to validate our experiment. The results of these tests and their details are available in online Appendix A.3. We fail to reject the null hypothesis of no design effects, as well as of the two most common sources of liars: ceiling and floor effects.

Estimating the incidence of electoral clientelism: When the described assumptions hold the difference in means between treatment and control groups is an unbiased estimator of the incidence of the sensitive item (Glynn, 2013). This estimator is:

$$\hat{\tau} = \frac{1}{N_1} \sum_{i=1}^N T_i Y_i - \frac{1}{N_0} \sum_{i=1}^N (1 - T_i) Y_i \quad (1)$$

Where, Y_i is the answer of individual i , N_1 and N_0 are the sizes of treatment and control groups, respectively; and T_i indicates the treatment status.

Table 1 takes a first broad view of the results for the treatment and control lists. The average number of factors people take into account when voting is 1.59 in the control list and 1.75 in the treatment list. This suggests that about 15.75% of people ($1.75 - 1.59$) exchange their vote for benefits, gifts, or jobs. This is our basic unconditional estimate of electoral clientelism.

¹²When this fails, the resulting bias may go in different directions depending on the underlying source of difference. One simple example arises with ceiling effects: suppose most people take into account all the things in the control list. When confronted with the control list, the respondent might have no problem of answering 'J' because all items are non-sensitive. But when confronted with the treatment list, the respondent (trivially if involved in the sensitive behavior, but plausibly even if not) might want to understate the number fearing it would expose her/him. In this case, we would observe an under-reporting of the sensitive behavior.

Table 1: Response frequency for the list experiment and unconditional estimator

	(1)	(2)	(3)	(4)
	Control group 1		Treatment group	
Response value	Frequency	Proportion (%)	Frequency	Proportion (%)
0	168	6.6	129	4.5
1	1,185	46.2	1,221	42.8
2	874	34.1	980	34.3
3	212	8.3	343	12.0
4	125	4.9	139	4.9
5			44	1.5
Average	1.587		1.746	
	(0.018)		(0.019)	
Observations	2,564		2,856	
Difference in means estimator			0.1588	
			(0.0261)	

Notes: The table lists the frequency in the reported number of factors people take into account when voting. The treatment list includes the same options as the corresponding control list, plus the following sensitive item: "The benefits, gifts, or jobs the candidate offered you in exchange for your vote." The difference in means estimator corresponds to the unpaired differences in mean estimator given by equation 1. Standard errors are in parenthesis.

This simple difference in means can also be computed for subsets of the population to study possible correlates of the sensitive behavior. However, this procedure is not statistically efficient.

Instead, we estimate the following model:

$$Y_i = f(X_i, \gamma) + T_i g(X_i, \delta) + \varepsilon_i. \quad (2)$$

For individual characteristics X_i , flexible functions f and g and parameters γ and δ , the resulting nonlinear least squares estimator is consistent but inefficient because it does not use all the available information about the joint distribution of $(Y_i(0), Z_{i,J+1}^*)$.¹³ To overcome this limitation, Imai (2011) proposes an additional maximum-likelihood estimator that models the joint distribution of

¹³This estimator includes both the difference-in-means estimator when $f(x, \gamma) = \gamma$ and $g(x, \delta) = \delta$, and a linear estimator with interaction terms if $f(x, \gamma) = x^T \gamma$ and $g(x, \delta) = x^T \delta$. The linear specification is easy to interpret, but it does not take into account that the response variables are bounded. Most of these models are estimated assuming logistic regression submodels: $f(x, \gamma) = \mathbf{E}(Y_i(0)|X_i = x) = J * \text{logit}^{-1}(x^T \gamma)$ and $g(x, \delta) = \Pr(Z_{i,J+1}^* = 1|X_i = x) = \text{logit}^{-1}(x^T \delta)$.

$(Y_i(0), Z_{i,J+1}^*)$ and treats $Z_{i,J+1}^*$ as (partially) missing data.¹⁴ We use this method in our analysis and, as Glynn (2013) shows, we can consistently estimate the probability that an individual with characteristic x will engage in a sensitive behavior (i.e $\Pr(Z_{i,J+1}^* = 1|X = x)$).

Estimating the extent of the social desirability bias: In addition to estimating incidence of vote buying and its correlates, investigating the extent of the social desirability bias is also possible when the experimental design includes the direct question about the sensitive item. As is the case in our application, individuals who are *not* shown the treatment list (since having observed the sensitive item on the list may influence them) are asked directly about the sensitive behavior. Comparing the incidence calculated using list experiments with the direct answer measures the average social desirability bias in the population, indicating the extent to which individuals want to conceal their behavior when asked directly. We can also move beyond this to estimate the social desirability bias for different types of individuals following Blair and Imai (2012). Let $Z_{i,J+1}(0)$ be i 's potential answer to the direct question in the control group. Then the social desirability bias for individuals with characteristics $X_i = x$ may be defined as:

$$S(x) = \Pr(Z_{i,J+1}^* = 1|X_i = x) - \Pr(Z_{i,J+1}(0) = 1|X_i = x), \quad (3)$$

for any $x \in \mathcal{X}$. As long as the no design and no liar assumptions hold, the first term can be estimated using the multivariate techniques discussed above. The second can be estimated directly with a regression of $Z_{i,J+1}(0)$ on observables (using logit, for instance).

We next explain the method we use to identify the most robust correlates of clientelism in our data.

2.2 Uncovering robust correlates of clientelism

To systematically establish which variables are most robustly correlated with clientelism and avoid the risk of data mining, we follow two parallel paths.

¹⁴Of course the partially missing data comes from the control list that does not disentangle $Z_{i,J+1}^*$ and $Z_{i,J}^*$

First, we select our list of covariates following the theoretical and empirical literature on clientelism and vote buying. Second, even with a judicious choice of covariates, there is scope for conscious or unconscious specification search leading to different, but actually not robust, conclusions. We use the extreme bounds methodology (Leamer, 1985; Levine & Renelt, 1992; Sala-i Martin, 1997).¹⁵ This exercise allows us to succinctly uncover the robust correlates of vote buying, without focusing attention in a subset of potential models that fall in line with our priors. However, since this method does not provide a basis for causal interpretation (Angrist & Pischke, 2010), we emphasize all our findings are mere correlations.

Consider the following regression for outcome y_i for individual i :

$$y_i = \alpha_j + \beta_j c_i + \gamma_j x_{ij} + \theta_j f_i + \epsilon_{ij},$$

where c_i is a correlate of interest and f_i is a set of controls that is always included in the regressions. We include in f_i a full set of region fixed effects (nine in total) that account for the potential heterogeneity in the presence of clientelism and beliefs about democracy across regions in Colombia (well recognized in the literature, see, for example Leal Buitrago and Dávila Ladrón de Guevara (2010)) and when examining interaction terms for covariates in c_i the lower direct terms.¹⁶ In our application, y_i is the answer to the *direct* question on clientelism (taking advantage of the lack of social desirability bias reported below). Thus y_i is a dummy variable (that indicates whether the respondent accepts benefits, gifts, or jobs in exchange for their vote). Finally, X is a pool of all correlates different to c_i and f_i , and x_{ij} is a vector of up to three variables in X .¹⁷

The method proceeds by estimating this regression for all possible x_{ij} (i.e.,

¹⁵This method has been widely used to study the correlates of economic growth (Reed, 2009) and other topics (Hafner-Burton, 2005; Sturm & De Haan, 2005; Wang, 2010; Gassebner, Lamla, & Vreeland, 2013).

¹⁶The nine regions are: Atlántica, Atlántica Media, Bogotá, Central, Centro-Oriente, Cundi-Boyacense, Eje Cafetero, Oriental y Pacífica.

¹⁷The number of variables x_{ij} to be included is up to the researcher. However, the limit of three variables has been a convention in the literature. See for instance Levine and Renelt (1992); Achen (2005)

for all j), and obtaining β_j and its standard error σ_j . In Leamer's (1985) original formulation, the lower extreme bound is simply the lowest value of $\beta_j - \tau\sigma_j$, and the upper extreme bound is the largest value of $\beta_j + \tau\sigma_j$, with τ the critical value for the confidence level. In this approach, c_i is considered a robust correlate of y_i when the lower and upper bounds have the same sign. This criteria may be overly conservative, potentially declaring a correlation fragile (as turns out to be the case for most correlates of cross-country growth, see Sala-i Martin (1997) and Levine and Renelt (1992)) on the basis of a single model.

Sala-i Martin (1997) instead proposes analyzing the entire distribution of β_j , and finding the cumulative density function to the left and right from zero. The largest of these two (hereafter $CDF(0)$) is the proportion of interest because it indicates where the coefficient is concentrated. Assuming normality, the densities can be recovered from the mean (β) and standard deviation (σ) of the distribution. These can be directly computed using the estimated β_j and σ_j as

$$\beta = \sum_j \phi_j \beta_j \text{ and } \sigma = \sum_j \phi_j \sigma_j,$$

where ϕ_j are weights proportional to some goodness-of-fit measure, like the adjusted R-squared or the integrated likelihood. However, with endogenous covariates the unweighted version may be preferable since endogenous regressions will have a better fit. The normality assumption can also be relaxed, computing $CDF(0)_j$ for each regression, and only then finding the (weighted) average $CDF(0)$.¹⁸ In this approach, variables that appear to be "significantly" correlated with the outcome are those with a (weighted) $CDF(0)$ larger than 0.95, or another benchmark confidence level.

In our exercise, we report Leamer's extreme bandwidth, the average (weighted and unweighted) parameters, and the cumulative density (both assuming and relaxing normality).¹⁹

¹⁸Some variations in the approach include weighting parameters with other measures of goodness of fit (Doppelhofer, Miller, & Sala-i Martin, 2004; Hegre & Sambanis, 2006) and using other types of regression models such as logit or ordered probits (Bjørnskov, Dreher, & Fischer, 2008; Moser & Sturm, 2011; Gassebner et al., 2013). Computational improvements are discussed in Hlavac (2015).

¹⁹Our results are similar when weighting with the likelihood or adjusted R-squared (we

3 Incidence of vote buying and (no) social desirability bias

Implementing the methods described in Section 2.1, Figure 1 plots the incidence of clientelism and provides evidence of the absence of social desirability bias. The point estimate for the prevalence of clientelism in the full sample is 18.5% (SE = 0.005). Notice that this estimate differs to the one obtained by the simple difference-in-means estimator of 15.75% because it employs a different, and superior, estimation procedure. In particular, here we are using the maximum likelihood estimator of equation (2) exploiting the joint distribution of $(Y_i(0), Z_{i,J+1}^*)$ and the full set of characteristics (X_i) listed in Figure 2 as controls.²⁰

Since rural and urban areas are very different, we show estimators separately as well. Estimated incidence is lower in urban areas (when asked directly, 14.9% with a standard error of 0.007) than in rural areas (22.1%, SE = 0.008). Using the list experiment, the conclusions and corresponding numbers are very similar (though the estimation precision is lower, as expected). Consequently, the point estimate for the social desirability bias is very small, less than 2 percentage points and not significant in either area.

Both the level of clientelism and the absence of social desirability bias contrasts with other findings in the literature. For example, in [Corstange \(2012\)](#) approximately one-fourth of Lebanese voters admitted to selling their vote when asked directly, but inferring from the list experiment suggests that nearly 50% had done so, which produced a large social desirability bias of 25 percentage points (twice the direct answer).²¹ Moreover, the willingness to admit to this behavior varies across religious beliefs.²² In contrast, we find that

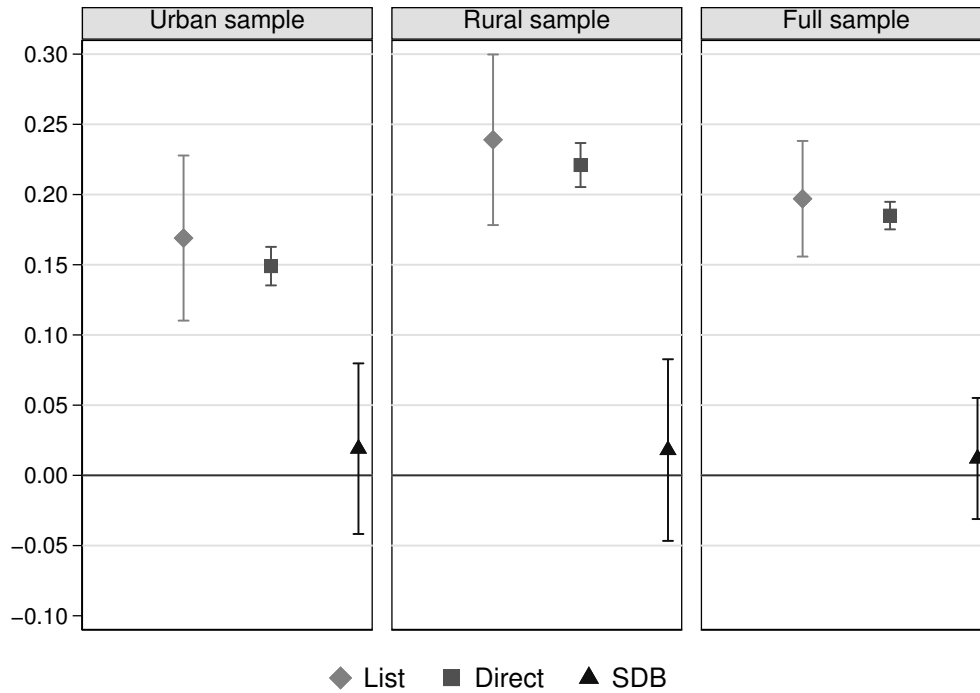
report the latter), or when running probit instead of linear models.

²⁰We experimented with alternative sets of controls, and the average incidence is not sensitive to these changes. Appendix Table A-1 describes all variables used in the analysis.

²¹In [Carkoglu and Aytac \(2015\)](#), asking the question directly suggested an estimated 16% prevalence of vote buying, which increased to 35% using the list. [Gonzalez-Ocantos et al. \(2012\)](#) conducted a similar experiment in Nicaragua, and found that 24% of registered voters (compared to 2% of those asked directly) were *offered* a chance to sell their vote (a wording that may fail to capture effective vote-buying transactions).

²²The Sunni are most willing to openly acknowledge engaging in this practice, followed by

Figure 1: Incidence and social desirability bias of clientelism

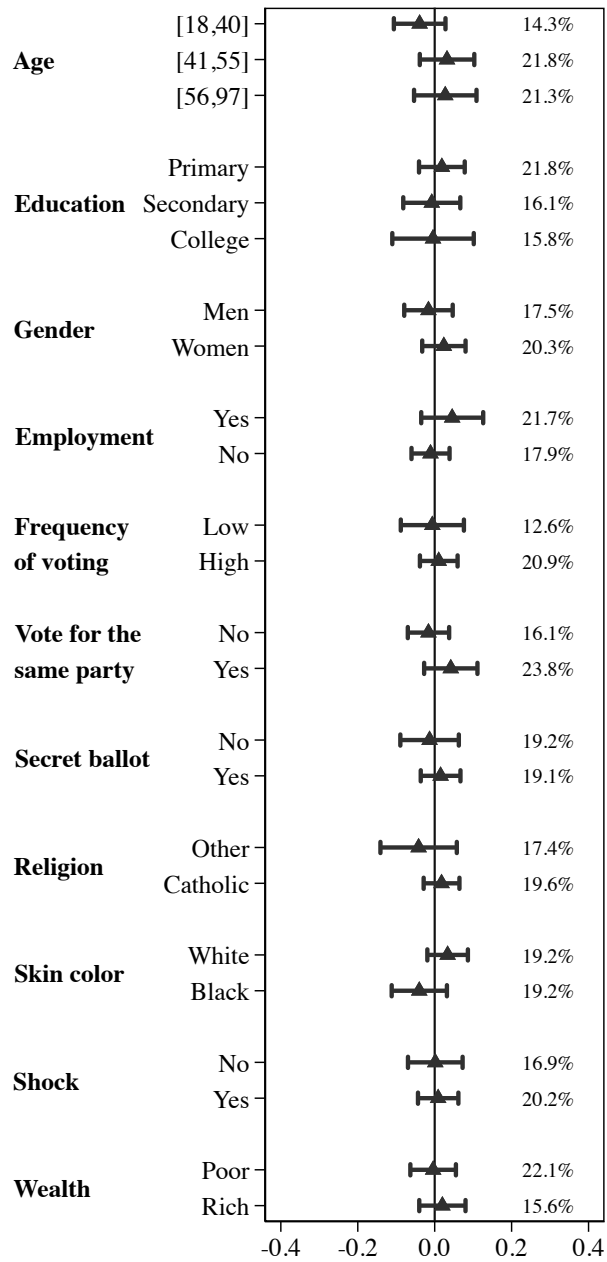


Notes: Incidence of clientelism as implied by the list experiment (diamond), direct question (square), and the difference between these two measures, capturing the extent of social desirability bias (triangle). Lines mark 95% confidence bounds. Estimates in this figure control for the set of individual characteristics listed in Figure 2. The list experiment sample in urban and rural areas are 2,683 and 2,737 households, respectively. The direct question sample on the other hand has 2,537 urban households and 2,554 rural households.

people are just as willing to admit to clientelism when asked directly as when presented with a list, and that the lack of social desirability holds not just on average but for different types of respondents. In Figure 2 we illustrate this by following the approach described in Equation (3) to evaluate the possible correlates of social desirability bias. Across all respondent types we find no social desirability bias.

We conclude that most types of people are willing to openly admit the extent to which they exchange votes for favors. We find similar results when Christians, and then Shi'a.

Figure 2: Incidence and social desirability bias of clientelism across different covariates



Notes: This Figure shows the social desirability bias estimate (and its 95% confidence interval) across different individual characteristics following Equation (3). Percentage values in the right represent the estimated incidence of clientelism using the list experiments (all point estimates are significant at the 95% confidence). Standard errors are computed using Monte Carlo simulations, and estimations control for the set of variables listed in the table. See Table A-1 for a description of all variables. Table A-7 report the point estimates for this figure and the results for urban and rural areas.

exploring other possible individual characteristics. One possible explanation is that respondents in our survey are comfortable enough with survey officers, having been visited by the organization three years earlier for the baseline survey and answering a long questionnaire, to provide honest answers. While little evidence is available, it appears more likely that this is not simply a feature of our data, but rather a reflection of broad acceptance of the practice of vote buying in Colombia. For instance, [Garcia and Pantoja \(2015\)](#), in the 2015 Colombian chapter of the Latin American Public Opinion Project (Lapop), use list experiments to test for vote buying, positive (and negative) patronage – promising to give (or threatening to remove, respectively) benefits contingent on electoral support – and direct intimidation. Their estimates differ to ours: they find no significant evidence of vote buying or negative patronage in these areas, though the incidence of intimidation and positive patronage is estimated at approximately 7% of voters in both cases. Yet like us, they find no social desirability bias. Using their data, [Figure 3](#) shows estimates that, though much noisier given their smaller sample, reveal no evidence of social desirability bias.

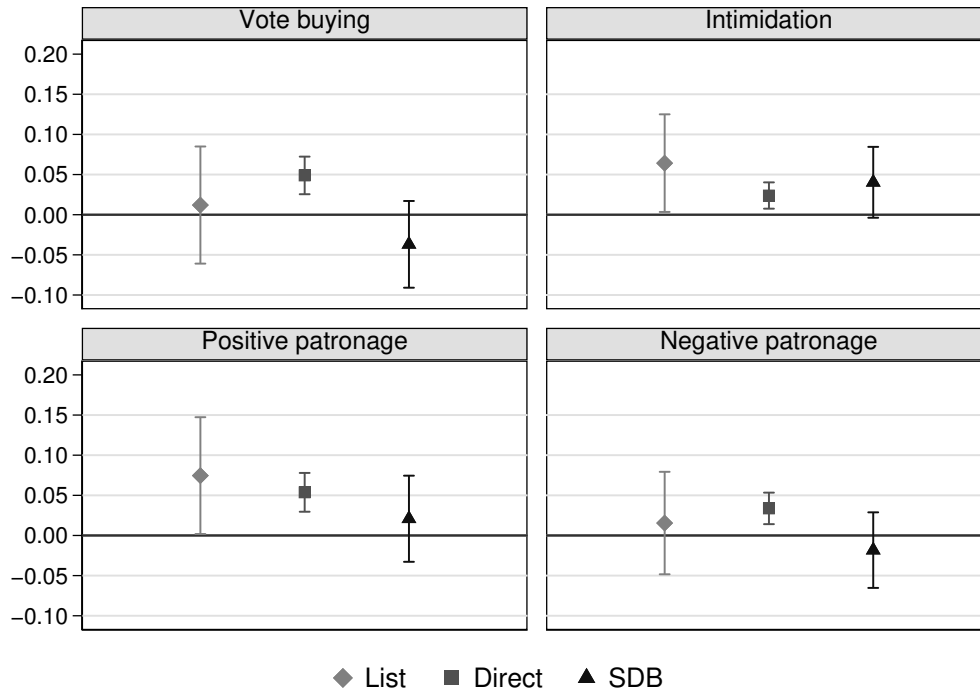
The difference in incidence, however, may reflect the very specific sample used in the Lapop study. Indeed, the Lapop survey also asked about electoral clientelism using a direct question in the full sample, and found numbers similar to ours: around 15% reported having at least one experience of electoral clientelism ([Carkoglu & Aytac, 2015](#)).²³ The similarity of these numbers again suggests that citizens are openly willing to admit the extent of electoral clientelism, a feature that may reflect the nature of a political equilibrium in which clientelism and a weak state reinforce each other as normal features of the political landscape ([Fergusson, Molina, & Robinson, 2017](#)).

This hypothesis is also consistent with citizens being willing to report whether they avoid paying VAT, which we analyze in [Fergusson, Molina, and Riaño \(2017\)](#).

Notice finally that though vote buying is very prevalent in our sample and not far from that in Lapop, the figure is smaller than in the papers using

²³The Lapop question asked about *offers* rather than *exchanges*: “And during the last election in YEAR, how often, if ever, did a candidate or someone from a political party offer you something, like food or a gift or money, in return for your vote?”

Figure 3: Incidence and social desirability bias in Lapop sample



Notes: For each electoral practice, the figure shows the incidence as implied by the list experiment (diamond), direct question (square), and the difference between these two measures, capturing the extent of social desirability bias (SDB, triangle). Lines mark 95% confidence bounds. Data from Lapop-Colombia, 2015 (García, Montalvo, & Seligson, 2015). Sample size in the list experiment and the direct question are respectively: 650 and 327 individuals for vote buying; 653 and 335 individuals for intimidation; 649 and 335 individuals for positive patronage; and 648 and 326 individuals for Negative patronage.

list experiments in other countries. Apart from higher real incidence, the distribution of abstention in the samples may play a role. Indeed, if one out of five people have sold their vote but only three out of five people actually turn out to vote, the impact of vote buying almost doubles. Using a self-reported turnout question for the local elections of 2010 and a question on “how frequently” the respondent participates in elections, we find that voters do engage in more vote buying than non-voters. Non-voters still report vote buying, but of course this may be due to their behavior in other elections in

which they did turn out. In both cases, vote buying is close to 20% for (frequent) voters whereas those of (infrequent) voters is in the 10-14% range.

4 Correlates of vote buying

In this section we present the main correlates of clientelism implementing the methodology described in Section 2.2.²⁴ Our inclusion of relevant variables was guided by a review of the most relevant literature, but in the text below we discuss the literature just briefly and only in connection with the most salient results. Our focus is on the most important debates on possible determinants, and on those for which we have particularly useful empirical measures in our dataset.

Table 2 reports descriptive statistics. Definitions and sources for all variables are in Table A-1. To facilitate identifying the magnitude of the correlations, we standardize all variables in the regression analysis. In Table 3 we report the results of the extreme bounds methodology analysis when applied to our direct question on clientelism (exchange of gifts, jobs, or benefits). Table 4 shows the results of a similar analysis for interaction terms (where the lower-order direct effects are always included in the regression). To easily identify the most important correlates, in these tables variables are sorted from most to least robustly (or “significantly”) correlated with clientelism.²⁵

4.1 Individual characteristics

A key focus of the literature is what types of individuals are more likely to be targeted with vote buying. A higher monetary value is more likely to produce an exchange for poorer voters or those for whom the marginal benefit of a given reward is higher, for instance because they are in a situation of distress.

We use rich data on household assets to construct a wealth index, and use questions on expenditure and on the incidence of negative shocks for the house-

²⁴Appendix A.4 also reports simple alternative bivariate and multivariate regressions, which produce similar conclusions.

²⁵Notice that we need to run $\binom{41}{3} + \binom{41}{2} + \binom{41}{1} = 11,521$ regressions to assess the robustness of each estimate.

hold to test if poorer people and those in distress sell their votes more. Table 3 suggests that wealthier households (row 6) are indeed “significantly” less likely to sell their votes; 100% of the coefficients are concentrated on the negatives (the value of $CDF(0)$ is 1 regardless of the precise assumptions used). Those experiencing negative shocks (row 22) are not more susceptible to clientelism at conventional levels (though a non-negligible 81% of the coefficients are greater than zero); nor are those with higher levels of expenditure per capita (row 19).

Individuals who believe more strongly in the rule of law and democracy should also presumably be less willing to compromise their ideals and sell their vote. We rely on five questions on the degree to which respondents agree with the following statements: it is important that rulers are elected by popular vote (not significant in row 36 of Table 3); it is sometimes justified to pay a bribe (positive and significant in row 2 of Table 3); when the government does not punish criminals, people can take justice into their own hands (positive and significant in row 9); sometimes, the use of violence is justified (positive and significant in row 12); and to capture criminals, authorities should sometimes violate the law (positive and marginally significant in row 15). Surprisingly, the most direct measure of belief in electoral democracy (“it is important that rulers are elected by popular vote”) does not appear to robustly correlate with clientelism. However, the next four variables, which capture different dimensions of tolerance of anti-social or illegal behavior, always point in the same direction, and are positively correlated.²⁶

We also analyze questions on political knowledge, political engagement, and partisan affinity to explore the idea (see, for example, [Corstange \(2010\)](#)) that those who are not interested in politics value their vote less, and can be easily swayed (i.e., they have low reservation prices just like the poor might). But party recall is positive and significant in row 5 of Table 3, party identity is positive and significant in row 14, and persuasion is positive and significant in row 1.²⁷

²⁶We also include variables that capture individuals’ beliefs about the role that the government (as opposed to the people) plays in determining individual welfare, labeled “Government role” (row 35) and “Own welfare” (row 33), and about how actively the state should combat inequality (row 34). None of these variables is robustly correlated with clientelism.

²⁷Other individual covariates included in the analysis are: age, which correlates negatively

4.2 Features facilitating enforcement

Vote buying is subject to severe commitment problems, as voters may renege and vote for an opposing candidate. Parties may also renege by failing to deliver promised benefits in the future.²⁸ Many of these problems are solved if there is no secret ballot and politicians can effectively monitor vote choice. Conversely, the more opaque and anonymous voting is, the less likely vote buying is to occur. Some papers (e.g., [Lehoucq and Molina \(2002\)](#) and [Fujiwara \(2015\)](#)) provide evidence that innovations like secret ballots or electronic voting prevent clientelism.

However, clientelism also persists with the use of secret ballots.²⁹ One possibility is that, where the ballot is effectively secret and monitoring is difficult or impossible, a feeling of reciprocal obligation may sustain clientelism ([Lawson & Greene, 2014](#); [Finan & Schechter, 2012](#)). Reciprocity not only overcomes commitment problems; politicians may also prefer reciprocal voters because they can better predict how they will vote and because reciprocal voters may be better at maintaining the long-term relationship of mutual cooperation often associated with clientelism.

Empirically, this motivates examining the connection between vote buying and available measures of positive reciprocity (agreement with the statement “you should always help those who help you”) and negative reciprocity (agreement with the statement “whoever harms me, pays for it”). In Table 3 the

with clientelism (row 10), and gender – women show no differential behavior (row 41). Religion is broken down into the two largest religious groups, Catholicism and Protestantism (rows 37 and 18) (see Table 2) and other religious minorities as a whole (row 3); only the latter group robustly reveals less clientelism (compared to atheists or agnostics, the excluded category). We are unable to explore a few potentially relevant variables due to lack of data. In particular, more risk-averse individuals may prefer a possibly more certain, targeted clientelistic exchange over an uncertain programmatic exchange. To the extent that clientelistic benefits accrue sooner, more impatient individuals may also be more willing to sell their votes.

²⁸[Robinson and Verdier \(2013\)](#) propose that clientelism often takes the form of (inefficient) public employment because it helps overcome the two-sided credibility problem.

²⁹Political machines’ insertion into dense voters’ social networks, another potential solution to the enforcement problem ([Stokes, 2005](#); [Stokes et al., 2013](#)), does not appear to be significant in our regressions, at least as captured by variables like the number of neighbors for whom respondents have a cell phone number or the number of neighbors from whom they can request a loan, or by comparing rural and urban areas, since social connections tend to be thinner in the latter.

expected positive connection is quite clear and robust (rows 4 and 7).

We also directly asked about belief in the secret ballot. As expected, believing that the ballot is secret is significantly and negatively correlated with clientelism in Table 3 (row 11). This belief is important in itself, and could also interact with the reciprocity and network measures, since where voting is thought to be secret, other enforcement mechanisms must be present (Stokes et al., 2013). Table 4 examines this issue by reporting the relevant interaction terms; none is robustly correlated with our key outcome variable.

4.3 Contextual factors

The political and socioeconomic context may play an important role in the prevalence of clientelism beyond simply influencing individual characteristics. For instance, the closer parties are to one another ideologically, the more likely vote buying is to occur, as citizens are more indifferent to the identity of the winner (Stokes, 2005, 2007). We explore this possibility by looking at whether responses correlate with the municipality having no right-wing or left-wing party (among the top two) in the most recent mayoral election³⁰ and find that there is a positive association (row 13). Other contextual factors that might facilitate clientelism, however, are not robustly correlated with vote buying in our sample.³¹

4.4 The core versus swing voters debate

A key element of the debate on who is targeted with vote buying is whether politicians focus on swing or core voters, or even opponents (Lindbeck & Weibull, 1987; Stokes, 2005). We have only implicitly referred thus far to

³⁰We use the party classification of Fergusson, Querubín, Ruiz, and Vargas (2017).

³¹These include economic inequality (land Gini coefficient, row 42), an index of municipality state presence (row 30), other proxies for law and order like the homicide rate (row 24), presence of guerrillas (row 32), and presence of paramilitaries (row 39), or an interaction between the degree of electoral competition with variables that presumably influence who gets targeted. This last exercise is motivated by the evidence in Corstange (2010), for Lebanon, suggesting that districts characterized by monopsonistic buyers are better able to discriminate and target voters more willing to sell than in areas with dueling machines.

this debate, because theoretical predictions on swing versus core voters are ambiguous unless one carefully measures the different types of clientelistic vote buying (e.g., turnout buying, abstention buying, persuasion), as well as the type of inducements delivered (G. Cox & Kousser, 1981; G. Cox & McCubbins, 1986; Calvo & Murillo, 2004; Nichter, 2008; Albertus, 2013; Gans-Morse, Mazzuca, & Nichter, 2014; Stokes, 2005). The comparative static predictions for each strategy also differ.

Thus the core versus swing voter debate is most likely to move forward with surveys explicitly designed to test some of the mechanisms mentioned above. This is not the case for our data, so we run a few correlations that must be examined these caveats in mind. Drawing on the approach of Gans-Morse et al. (2014), the salience of political preferences and the level of political polarization should reduce vote buying. We can use the questions on party identity and political persuasion to measure the salience of political preferences. As discussed above, these features correlate positively with clientelism. When instead examining the correlation using a polarization index (Reynal-Querol, 2002), Table 3 (row 17) suggests a negative sign, though the $CDF(0)$ values fall a bit short of the 95% benchmark (an index of electoral fractionalization is also not “significant” in row 21).

One very crude way to look at whether core, swing, or opponent voters are targeted is to combine people’s self-declared ideology with that of the “dominant” party in each location (defined as dummy variables for whether the incumbent mayor is right-wing or left-wing). Assuming that the dominant party has better clientelistic networks, then observing that in right-dominated (left-dominated) areas those with right-wing (left-wing) views are more likely to participate in clientelism could indicate that “core” voters are being targeted. Instead, if in right-dominated (left-dominated) areas those with left-wing (right-wing) views are more likely to participate in clientelism, this could signal persuasion buying.³² When we include such interactions in Table 4 we find

³²We are aware that ideally the notion of “core” or “loyal” voters should be based on individuals’ ideological predispositions towards the clientelistic party Stokes (2007), and measuring such tendencies is plagued with endogeneity problems: do people report an affinity with (or indifference to) a given party/ideology because they are ideologically predisposed, or because they have been bought in the past? Nor is identifying the “clientelistic” party simple. Much

a positive and significant coefficient for the interaction of right ideology with right domination, but not for the left interactions. This finding could reflect that in Colombia, core supporters are targeted and left-wing parties are excluded from clientelistic networks (Fergusson, Querubín, et al., 2017). Notice also that the direct correlation of clientelism with right ideology, in row 8 in Table 3, is positive.

5 Discussion and implications

In this section, we discuss the relevance of our findings, centering on the implications for economic policy and for some of the key debates in the economics of vote buying.

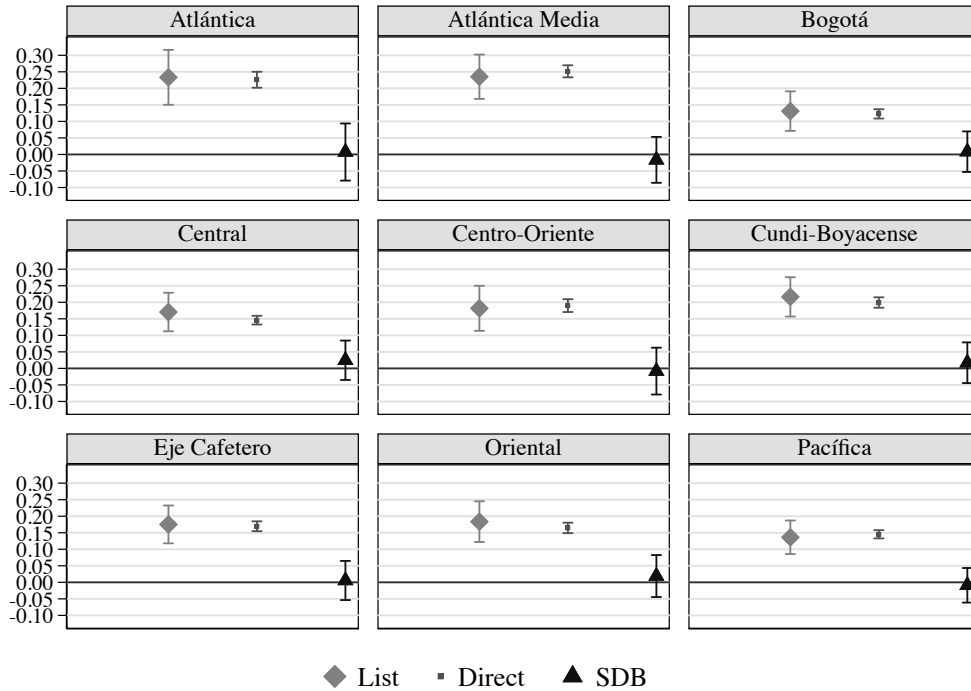
5.1 Social Desirability Bias and fighting clientelism

One of the most remarkable results presented above is the absence of social desirability bias in reports of clientelism. This result holds not only across different subgroups of the population (Figure 2 above), but also across different geographical regions as we show in Figure 4. Clientelism varies by region: the coast (regions *Atlántica* and *Atlántica Media*) features the highest incidence, nearing almost 25%, while in *Bogotá* it is approximately 12%. But one feature remains constant: respondents admit at similar rates when asked directly versus indirectly with the list experiment, implying a widespread lack of social stigma.

This result highlights possible obstacles to overcome clientelism. Like any other unethical behavior, vote buying is hard but not impossible to defeat. Several works have shown significant success in fighting corrupt or undesired behaviors using relatively cheap, simple interventions that leverage on people's concerns about their social image (see, for instance, Perez-Truglia and Troiano

of the literature focuses on monopsonistic vote buying (e.g., in the models of Stokes (2005) and Gans-Morse et al. (2014), only one party buys votes), since there is often one dominant clientelistic party with access to the necessary networks and resources. This could justify the assumption that the dominant party is the clientelistic party. However, dueling party machines are also possible, and in Colombia most parties are believed to engage in clientelism; competition for clients leads some parties to participate in a form of "market" clientelism (Gutiérrez & Andrés, 1998; Dávila Ladrón De Guevara, 1999; García, 2002; Gutiérrez, 2007).

Figure 4: Incidence and social desirability bias of clientelism across regions



Notes: Incidence of clientelism in different regions of Colombia as implied by the list experiment (diamond), direct question (square), and the difference between these two measures, capturing the extent of social desirability bias (triangle). Lines mark 95% confidence intervals. Estimates in this figure control for the set of individual characteristics listed in Figure 2.

For a distinction between urban and rural regions, see Table A-1.

(2015) for efforts to reduce tax evasion and DellaVigna, List, Malmendier, and Rao (2017) for social appeals to increase voting). However, that vote buying appears to be considered so “normal” in the Colombian context likely limits the effectiveness of these types of interventions. More precisely, rather than leveraging on an existing social norm, eliminating this practice might first require changing existing views to a new social understanding that clientelism is undesirable. But changing social norms with simple policy interventions is much more challenging (Bicchieri & Mercier, 2014).

5.2 On the significance of the main correlates of vote buying

In addition to investigating social desirability bias, our second main contribution is uncovering the robust correlates of clientelism. While we remain cautious by not providing causal interpretations, the results provide informative *prima facie* evidence that supports some theories and challenges others. We now discuss some of the most important findings and their implications.

Strong electoral institutions – In line with the literature emphasizing enforcement issues in clientelistic vote buying, we find that clientelism is more prevalent among more reciprocal voters and among those who believe that voting is not secret. That reciprocity facilitates clientelism is potentially troubling: it implies, at least in theory, that it could be sustained without a secret ballot via a self-enforcing relationship without monitoring. But that belief in a secret ballot is negatively correlated with vote buying is consistent with the idea that better electoral institutions and citizen confidence on them can help reduce vote buying.

Economic vulnerability – In our results, wealthier individuals appear to engage less in clientelism, fitting theories where poverty is both fertile ground for, and a consequence of, this practice.³³ By affecting more vulnerable voters, clientelism can consolidate a vicious cycle: It predates on (and reproduces) poverty and vulnerability, amplifying political and economic inequality (Fergusson, 2017). In this context, interventions that aim to improve the sources of income of the most vulnerable, and to reduce their dependency on politicians, could reduce the incidence of vote buying (see recent evidence from Brazil –Bobonis et al. (2017)– and Uganda –Blattman, Emeriau, and Fiala (2017)– along these lines).

Interest in politics – One intriguing finding, which runs contrary to several theories of clientelism (Lindbeck & Weibull, 1987; Dixit & Londregan, 1996, for example), is that individuals who are most interested in politics³⁴ are *more* clientelistic. Typically, these theories conclude that more engaged voters are

³³A caveat is that while wealth is a robust correlate of clientelism, the association is not very strong (clientelism is 0.09 standard deviations lower among households that are one-standard deviation wealthier than the average).

³⁴Captured empirically in our data by whether or not they persuade others to vote, identify with a party, or recall which party they voted for in the most recent election

either harder to persuade to vote for an opposing candidate in exchange for particularly targeted benefits, or a waste of resources in the case of sympathetic candidates – as they will turn out to vote enthusiastically in any case (Stokes, 2007; G. W. Cox, 2010).³⁵

Examining what underlies this correlation in greater detail is relevant to developing ways to combat vote buying. For instance, one hypothesis is that where clientelism is the “usual” form of political participation (a feature consistent with the lack of social desirability bias in our data), clientelism and interest in politics may be complements rather than substitutes. We hypothesize that this positive average correlation between these measures of political engagement and clientelism reveals how deeply entrenched clientelism is in Colombian society, to the extent that those who are more actively concerned about political matters are more likely to accept these forms of electoral exchange.³⁶ In other words, clientelism may be *the* way to get involved in politics in some contexts, rather than a substitute for doing so.

Notice also that since we are observing the outcome of a realized transaction, a related possibility is that more interested and engaged voters are disproportionately approached to sell their vote because they are easier to identify by candidates. Moreover, these citizens may be especially valuable for politicians, as they may have a multiplier effect by persuading other voters.

Beliefs – When examining the different sets of covariates, it is clear that individual “beliefs” tend to be robustly associated with clientelism. Specifically, individuals with less favorable views of the rule of law (in particular, those who believe that bribing, resorting to violence, or taking justice into their own hands may be justified) have a higher incidence of clientelistic vote buying. In contrast, some of the other household or individual features and, more clearly, contextual factors like the electoral features of jurisdiction (also discussed in the context of the swing versus core debate), are not very robustly associated with clientelism.

One possible reason for these results, in line with our interpretation of the

³⁵In line with these ideas, Corstange (2012) finds that disinterested voters are more likely to sell their vote in Lebanon. However, more consistent with our findings, Carkoglu and Aytac (2015) find in Turkey that those with weak partisan attachments sell their voters less.

³⁶For a related discussion, see Fergusson, Molina, and Robinson (2017).

lack of stigma for clientelism, is that only when individual's entire ethical principles change engagement with vote buying also changes. In other words, even though some contextual factors may matter, a key underlying root of clientelistic behavior appears to be a worldview that considers this practice natural.³⁷

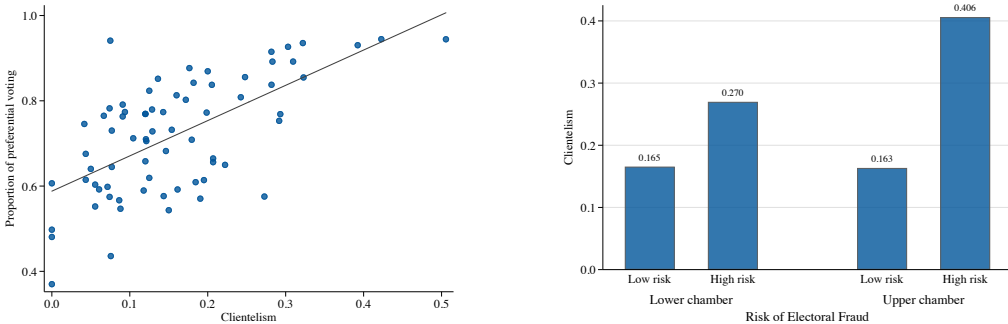
Non-robust correlates – Identifying covariates that are *not* robustly correlated is also informative. Consider one example. A common presumption is that education helps reduce clientelism. Yet our specifications suggest a non-robust correlation of clientelism with educational attainment (row 31 in Table 3), in spite of factors like reverse causality likely making it *more* likely to find a negative correlation. But our exercise also implies that researchers keen on finding this correlation might have focused on several combinations that produce significant results (65% of our coefficients are indeed less than zero!). While this does not fully prove that education is irrelevant, a researcher emphasizing a correlation that we have proven not to be robust will have to make a compelling case for his/her specification given the nature of our procedure.

Validating our measures of vote buying and further implications– Finally, while we focus on this paper mostly on determinants of clientelism, we end with a brief discussion of the implications of prevalent clientelistic vote buying, which also serves to gauge the validity of our vote buying estimates. In order to do so, in Figure 5 we present the municipality-level average of our vote buying measure and its relation with four outcomes of interest.

In panel A of Figure 5, we show that the average incidence of vote buying is strongly associated with the proportion of “preferential voting” (that is, for specific candidates in open party lists for Congress) at the municipality level in 2014. Preferential voting has long been associated with clientelistic practices (Putnam, Leonardi, & Nanetti, 1994) where personal relations supersede programmatic links. This exercise not only validates the relevance of our measure by showing that vote buying is higher where we would expect it to be, but also reveals that vote buying is part of a specific form of political exchange that is

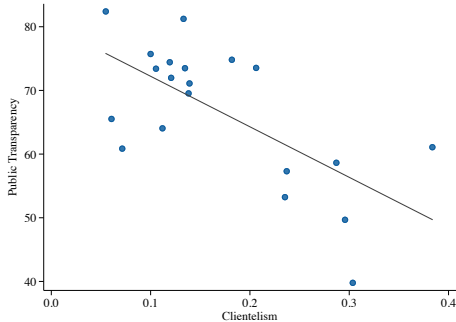
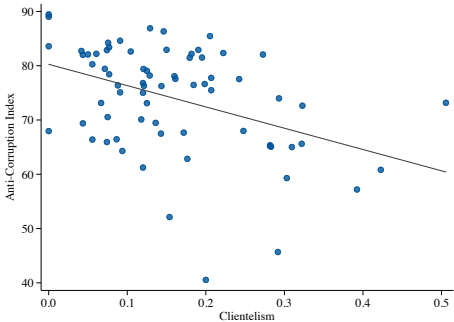
³⁷We must recognize, however, that particularly when comparing the importance of beliefs to more aggregate contextual variables, less variation in the latter may also partly explain the lack of significant correlations.

Figure 5: Municipality-level correlation with alternative measures of clientelism



Panel A. Clientelism and preferential voting

Panel B. Clientelism and Risk of Electoral Fraud



Panel C. Clientelism and Anti-Corruption Index

Panel D. Clientelism and Public Transparency

Notes: Panel A, B and C show correlations at the municipality level and Panel D at the departament level. Clientelism is the average incidence (at the municipality or departament level) of vote buying measured by the direct question. In each panel, we show the association between Clientelism and the proportion of preferential voting (Panel A), the Risk of Electoral Fraud (Panel B), The Anti-Corruption index of the municipal entities (Panel C) and The Transparency Index for departament entities (Panel D). See Table A-1 for a description of all variables.

often considered detrimental (Sani & Radaelli, 1992; Marsh, 1985). For further validation, Panel B shows that there is also a higher prevalence of vote buying in places with a higher risk of electoral fraud, as reported by The Colombian Electoral Observation Mission (MOE, for its Spanish acronym).

Likewise, and also important for the implications, Panel C and D show that vote buying is negatively correlated with measures on transparency and anti-corruption efforts. These correlations are consistent with the idea that clientelism, which relies on public funds for the reproduction of the clientelistic network, can create an incentive structure motivating corruption and arbitrary

redistribution and policy targeting.³⁸ Such distortions can be a key source of inefficiency (Olken & Pande, 2012).

6 Conclusions

As Stokes et al. (2013) put it, the transition from a situation where voters, rather than trading their votes for cash or other private benefits, vote to parties that offer public policies of which they approve, is a transition to a more democratic polity. The transition away from clientelism is thus part of a process of democratic consolidation. Nevertheless, clientelistic exchanges are prevalent in many societies, especially in new and poor democracies. Economic development and the processes that accompany it may help weaken clientelism, yet this is hardly a sufficient condition. Moreover, as several scholars emphasize, clientelism may not simply flourish in poverty, but can also help reproduce it. Understanding the key drivers of clientelism is therefore extremely important.

Nevertheless, empirical studies of clientelism must confront measures that, if at all available, are often imprecise or plagued with reporting biases. In this paper, we were able to estimate the incidence of clientelism using list experiments applied to a large sample of households, while demonstrating that social desirability in respondents' claims does not bias our measures. Estimates of illegal or sensitive activities are often unavailable, too coarse to relate with individual behavioral responses, or potentially plagued with reporting and other measurement errors. In our setting we have two key ingredients presenting a rare opportunity: (1) a very comprehensive survey with a wealth of information from survey respondents to examine the drivers and consequences of clientelistic vote buying, and (2) direct evidence that responses are not hampered by reporting biases.

We provided an overview of the resulting incidence and main correlates of clientelism. Since, as noted, clientelism is prevalent in many countries, our analysis is relevant beyond Colombia. Moreover, the richness of the microeconomic information should serve to carefully examine the mechanisms involved,

³⁸For related ideas and evidence see Stokes et al. (2013); Maiz and Requejo (2001); Singer (2009); Camacho and Conover (2011).

thus enabling researchers to draw useful lessons about possible general forces at play rather than restrict the analyses to simple observations of prevalence in the Colombian case.

In our review, we focused on the likely determinants of clientelism, highlighting the correlations that fall in line with some of the most prevalent existing theories and others that seem to challenge them and open potential avenues for new ideas. Advances in this direction are especially relevant to explore what factors may help explain the persistence or decline of clientelistic vote buying. We discuss the substantive significance and implications of the robust correlations we uncover, as well as of the apparent lack of social stigma associated with vote buying. Exploring the implications of clientelism for citizens (both the economic and political repercussions) is equally important, as we illustrated by exploring aggregate correlations with other electoral practices and overall corruption. This is another relevant area for future research that we expect to undertake using the wealth of information described in this paper.

References

- Acemoglu, D., Garcia-Jimeno, C., & Robinson, J. A. (2015). State capacity and economic development: a network approach. *American Economic Review*, 105(8), 2364-2409.
- Achen, C. H. (2005). Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong. *Conflict Management and Peace Science*, 22(4), 327-339.
- Albertus, M. (2013). Vote buying with multiple distributive goods. *Comparative Political Studies*, 46(9), 1082-1111.
- Angrist, J., & Pischke, J. (2010). The credibility revolution in empirical economics: how better research design is taking the con out of econometrics. *The Journal of Economic Perspectives*, 24(2), 3-30.
- Bates, R. (1981). *Markets and states in tropical Africa: the political basis of agricultural policies*. Berkeley: University of California Press.
- Bernal, R., Cadena, X., Camacho, A., Cárdenas, J., Fergusson, L., & Ibáñez, A. (2014). Encuesta longitudinal de la Universidad de los Andes (ELCA) – 2013. *Documento CEDE 42/2014, Universidad de Los Andes*.
- Bicchieri, C., & Mercier, H. (2014). Norms and beliefs: How change occurs. In M. Xenitidou & B. Edmonds (Eds.), *The complexity of social norms* (pp. 37–54). Cham: Springer International Publishing.
- Bjørnskov, C., Dreher, A., & Fischer, J. (2008). Cross-country determinants of life satisfaction: exploring different determinants across groups in society. *Social Choice and Welfare*, 30(1), 119-173.
- Blair, G., & Imai, K. (2012). Statistical analysis of list experiments. *Political Analysis*, 20(1), 47-77.
- Blattman, C., Emeriau, M., & Fiala, N. (2017, January). *Do Anti-Poverty Programs Sway Voters? Experimental Evidence from Uganda* (Working Paper No. 23035). National Bureau of Economic Research.
- Bobonis, G. J., Gertler, P., Gonzalez-Navarro, M., & Nichter, S. (2017). *Vulnerability and clientelism* (Tech. Rep.). NBER Working Paper No. 23589.
- Calvo, E., & Murillo, M. (2004). Who delivers? Partisan clients in the Argentine electoral market. *American Journal of Political Science*, 48, 742-757.

- Camacho, A., & Conover, E. (2011, May). Manipulation of social program eligibility. *American Economic Journal: Economic Policy*, 3(2), 41-65.
- Carkoglu, A., & Aytac, E. (2015). Who gets targeted for vote-buying? Evidence from an augmented list experiment in Turkey. *European Political Science Review*, 7(4), 547-566.
- Corstange, D. (2009). Sensitive questions, truthful answers? Modeling the list experiment with LISTIT. *Political Analysis*, 17(1), 45-63.
- Corstange, D. (2010). Vote buying under competition and monopsony: evidence from a list experiment in Lebanon. *Paper prepared for the 2010 Annual Conference of the American Political Science Association, Washington, D.C.*
- Corstange, D. (2012). Vote trafficking in Lebanon. *International Journal of Middle East Studies*, 44(3), 483-505.
- Cox, G., & Kousser, M. (1981). Turnout and rural corruption: New York as a test case. *American Journal of Political Science*, 25(4), 646-663.
- Cox, G., & McCubbins, M. (1986). Electoral politics as a redistributive game. *The Journal of Politics*, 48(2), 370-389.
- Cox, G. W. (2010). Swing voters, core voters, and distributive politics. In I. Shapiro, S. C. Stokes, E. J. Wood, & A. S. Kirshner (Eds.), *Political representation* (p. 342-357). Cambridge University Press.
- Dávila Ladrón De Guevara, A. (1999). Clientelismo, intermediación y representación política en Colombia: ¿Qué ha pasado en los noventa? *Revista de Estudios Políticos*, 15(1), 61-78.
- DellaVigna, S., List, J. A., Malmendier, U., & Rao, G. (2017). Voting to Tell Others. *Review of Economic Studies*, 84(1), 143-181.
- Dixit, A., & Londregan, J. (1996, 11). The determinants of success of special interests in redistributive politics. *The Journal of Politics*, 58, 1132-1155.
- Doppelhofer, G., Miller, R., & Sala-i Martin, X. (2004). Determinants of long-term growth: a bayesian averaging of classical estimates (bace) approach. *American Economic Review*, 94(4), 813-835.
- Fergusson, L. (2017, December). *Who wants violence? The political economy of conflict and state building in Colombia* (DOCUMENTOS CEDE No. 015890). UNIVERSIDAD DE LOS ANDES-CEDE.
- Fergusson, L., Molina, C., & Robinson, J. (2017). *The vicious cycle of clientelism*

- and state weakness*. (Unpublished manuscript)
- Fergusson, L., Molina, C. A., & Riaño, J. F. (2017). I evade taxes, and so what? a new database and evidence from Colombia. *Documento CEDE 21/2017, Universidad de Los Andes*.
- Fergusson, L., Querubín, P., Ruiz, A., & Vargas, J. (2017). The real winner's curse. *Documento CEDE 5/2017, Universidad de Los Andes*.
- Fergusson, L., & Riaño, J. (2014). La política Colombiana a la luz de la Elca: entre el desinterés y el clientelismo. In X. Cadena (Ed.), *Colombia en movimiento 2010-2013*. Ediciones Uniandes.
- Finan, F., & Schechter, L. (2012). Vote-buying and reciprocity. *Econometrica*, 80(2), 863-881.
- Fujiwara, T. (2015). Voting technology, political responsiveness, and infant health: evidence from Brazil. *Econometrica*, 83(2), 423-464.
- Gans-Morse, J., Mazzuca, S., & Nichter, S. (2014). Varieties of clientelism: machine politics during elections. *American Journal of Political Science*, 58(2), 415-432.
- García, M. (2002). La política bogotana un espacio de recomposición (1992-2001). In F. Gutiérrez (Ed.), *Degradación o cambio: evolución del sistema político Colombiano*. Bogotá: Grupo Editorial Norma.
- García, M., Montalvo, J., & Seligson, M. (2015). *Cultura política de la democracia en Colombia, 2015: actitudes democráticas en zonas de consolidación territorial*. Bogotá: Universidad de los Andes, Observatorio de la democracia, Lapop, Americas Barometer.
- García, M., & Pantoja, S. (2015). Incidencia del clientelismo según riesgo electoral y de violencia. In *Mapas y factores de riesgo electoral. Elecciones de autoridades locales. Colombia*. Bogotá: Misión de Observación Electoral.
- Gassebner, M., Lamla, M., & Vreeland, R. (2013). Extreme bounds of democracy. *Journal of Conflict Resolution*, 57(2), 171-197.
- Glynn, A. (2013). What can we learn with statistical truth serum? design and analysis of the list experiment. *Public Opinion Quarterly*, 77, 159-172.
- Gonzalez-Ocantos, E., de Jonge, C., Meléndez, C., Osorio, J., & Nickerson, D. (2012). Vote buying and social desirability bias: experimental evidence from Nicaragua. *American Journal of Political Science*, 56(1), 202-217.

- Gutiérrez, F. (2007). *¿lo que el viento se llevó? Los partidos políticos y la democracia en Colombia, 1958-2002*. Bogotá: Grupo Editorial Norma.
- Gutiérrez, F., & Andrés, D. (1998). *La ciudad representada: política y conflicto en Bogotá*. Bogotá: IEPRI-Tercer Mundo.
- Hafner-Burton, E. (2005). Right or robust? The sensitive nature of repression to globalization. *Journal of Peace Research*, 42(6), 679-698.
- Hegre, H., & Sambanis, N. (2006). Sensitivity analysis of empirical results on civil war onset. *Journal of conflict resolution*, 50(4), 508-535.
- Hicken, A. (2011). Clientelism. *Annual Review of Political Science*, 14(1), 289-310.
- Hlavac, M. (2015). Extreme bounds analysis in R. *Harvard University*, 47(4), 685-700.
- Holbrook, A., & Krosnick, J. (2010). Social desirability bias in voter turnout reports: tests using the item count technique. *Public Opinion Quarterly*, 74(1), 37-67.
- Imai, K. (2011). Multivariate regression analysis for the item count technique. *American Statistical Association*, 106(494), 407-416.
- Kitschelt, H. (2000). Linkages between citizens and politicians in democratic politics. *Comparative Political Studies*, 33(6-7), 845-79.
- Lawson, C., & Greene, K. (2014). Making clientelism work: how norms of reciprocity increase voter compliance. *Comparative Politics*, 47(1), 61-77.
- Leal Buitrago, F., & Dávila Ladrón de Guevara, A. (2010). *Clientelismo: el sistema político y su expresión regional*. Ediciones Uniandes-Universidad de los Andes.
- Leamer, E. (1985). Sensitivity analyses would help. *American Economic Review*, 75(3), 308-313.
- Lehoucq, F., & Molina, I. (2002). *Stuffing the ballot box: fraud, electoral reform, and democratization in Costa Rica*. Cambridge: Cambridge University Press.
- Levine, R., & Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review*, 82(4), 942-963.
- Lindbeck, A., & Weibull, J. (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice*, 52(3), 273-297.
- Maiz, R., & Requejo, R. (2001). Clientelism as a political incentive structure for corruption. In *Trabajo presentado en la conferencia de primavera del european*

- consortium for political research, grenoble.*
- Marsh, M. (1985). The Voters Decide?: Preferential Voting in European List Systems. *European Journal of Political Research*, 13(4), 365-378.
- Matanock, A., & García-Sánchez, M. (2011a). Fighting for hearts and minds: examining popular support for the military in Colombia. *Presented at the Governance, Development, and Political Violence Conference at the University of California at San Diego.*
- Matanock, A., & García-Sánchez, M. (2011b). *A tangled web: measuring and explaining popular support for the military and the paramilitaries in Colombia.* (Unpublished manuscript)
- Montalvo, J., & Reynal-Querol, M. (2005). Ethnic polarization, potential conflict and civil war. *American Economic Review*, 95(3), 796-816.
- Moser, C., & Sturm, J. (2011). Explaining IMF lending decisions after the cold war. *The Review of International Organizations*, 6(3-4), 307-340.
- Nichter, S. (2008). Vote buying or turnout buying? Machine politics and the secret ballot. *American Political Science Review*, 102(1), 19-31.
- Nichter, S. (2014). Conceptualizing vote buying. *Electoral Studies*, 35(2), 315-327.
- Olken, B. A., & Pande, R. (2012, July). Corruption in Developing Countries. *Annual Review of Economics*, 4(1), 479-509.
- Pachón, M., & Sánchez, F. (2014). Base de datos sobre resultados electorales. *Documento CEDE 29/2014, Universidad de Los Andes.*
- Perez-Truglia, R., & Troiano, U. (2015, June). *Shaming tax delinquents* (Working Paper No. 21264). National Bureau of Economic Research.
- Putnam, R. D., Leonardi, R., & Nanetti, R. Y. (1994). *Making democracy work: civic traditions in modern Italy.* Princeton university press.
- Reed, R. (2009). The determinants of US state economic growth: a less extreme bounds analysis. *Economic Inquiry*, 47(4), 685-700.
- Reynal-Querol, M. (2002). Ethnicity, political systems, and civil wars. *Journal of Conflict Resolution*, 46(1), 29-54.
- Robinson, J., & Verdier, T. (2013). The political economy of clientelism. *Scandinavian Journal of Economics*, 115(2), 260-291.
- Sala-i Martin, X. (1997). I just ran two million regressions. *American Economic Review*, 87(2), 178-183.

- Sani, G., & Radaelli, C. M. (1992). Preference voting: Before and after the 1991 referendum. *Il Politico*, 57(2 (162)), 193–203.
- Singer, M. (2009). Buying voters with dirty money: The relationship between clientelism and corruption.
- Steele, A., & Shapiro, J. (2012). *State-building, counterinsurgency, and development in Colombia*. (Unpublished manuscript)
- Stokes, S. (2005). Perverse accountability: a formal model of machine politics with evidence from Argentina. *American Political Science Review*, 99(3), 315-25.
- Stokes, S. (2007). Political clientelism. In C. Boix & S. Stokes (Eds.), *Oxford handbook of comparative politics* (p. 604-27). Oxford: Oxford University Press.
- Stokes, S., Dunning, T., Nazareno, M., & Brusco, V. (2013). *Brokers, voters, and clientelism: the puzzle of distributive politics*. Cambridge: Cambridge University Press.
- Sturm, J., & De Haan, J. (2005). Determinants of long-term growth: new results applying robust estimation and extreme bounds analysis. *Empirical Economics*, 30(3), 597-617.
- Wang, E. (2010). Determinants of R&D investment: the extreme-bounds-analysis approach applied to 26 OECD countries. *Research Policy*, 39(1), 103-116.

A Online appendix (Not for publication)

A.1 Variable definition and sources

Table A-1: Variable definition and sources

Variable	Description
<i>Age</i>	Respondent's age in years.
<i>Agree with bribery</i>	Equals 1 if respondent "totally agrees" or "agrees" with the statement: "As things are, sometimes paying a bribe is justified."
<i>Anti-Corruption Index</i>	Summary measure for the performance of strategic anti-corruption standards in municipalities using measures of: internal control, recruitment, administrative management systems and accountability. Computed by the <i>General Attorney</i> , available as <i>Índice de Gobierno Abierto</i> .
<i>Authorities violate the law</i>	Equals 1 if respondent "totally agrees" or "agrees" with the statement: "To capture criminals, authorities should sometimes violate the law."
<i>Catholic</i>	Equals 1 for Catholics (self-declared).
<i>Contributes to social security</i>	Equals 1 if respondent pays contributions to the social security system (thus excluding members in the subsidized social security system).
<i>Dark</i>	Respondent's skin color based on color palette (assessed by interviewer). The palette is numbered from 1 to 11 (1 = lightest color). <i>Dark</i> Equals 1 for colors greater than or equal to 5.
<i>Education</i>	Educational attainment. Equals 1 if respondent has some secondary education or more.
<i>Employed household head</i>	Equals 1 if household head was employed in the week preceding the survey.
<i>Employment</i>	Equals 1 if respondent was employed the week previous to the survey.
<i>Evangelical/Pentecostal</i>	Equals 1 for Evangelical/Pentecostal (self-declared).
<i>FEA</i>	Equals 1 if household is a <i>Familias en acción</i> (main conditional cash transfer for the poor with school-age children) beneficiary.
<i>Formal credit</i>	Equals 1 if household has any formal credit.
<i>Fractionalization</i>	$F_j = 1 - \sum_{i=1}^N \pi_{ij}^2$, where π_{ij} is the vote share for the mayoral candidate (in 2011) i in municipality j . See Montalvo and Reynal-Querol (2005) . Pachón and Sánchez (2014) .
<i>Frequency of voting</i>	Equals 1 if respondent "Always votes in elections" or "Votes in most elections", 0 otherwise ("Rarely votes in elections" or "Never voted").

Continued on next page

Table A-1 – Variable definition and sources, continued from previous page

Variable	Description
<i>Get help</i>	Equals 1 if household answers “yes” to at least one of the following questions: “During the past 12 months, do any members of the household receive money or in-kind aid...” a. “...from relatives or friends living in Colombia?”, b. “from relatives or friends living abroad?”, c. “for alimony?”, d. “from international organizations (WFP, UNICEF, ICRC)?”, e. “from NGOs?”, f. “from the church or other religious organizations?”, g. “from other persons, entities or organizations?”.
<i>Gini of land properties</i>	Municipal land Gini. <i>Cede Panel</i> .
<i>Gov. against inequality</i>	Equals 1 if respondent “totally agrees” or “agrees” with the statement: “The government should implement strong policies to reduce inequality between rich and poor.”
<i>Government role</i>	Equals 1 if respondent “totally agrees” or “agrees” with the statement: “Government is primarily responsible for ensuring the welfare of the people.”
<i>Guerrillas</i>	Number of violent events per year perpetrated by guerrillas per 100,000 inhabitants (average 2010–2012). <i>Conflict analysis resource center (CERAC) and Universidad del Rosario</i> .
<i>HH expenses</i>	Per capita household expenses. See Bernal et al. (2014) .
<i>HH food expenses</i>	Household expenses in food (Colombian pesos). See Bernal et al. (2014) .
<i>Homeowner</i>	Equals 1 if the household residence is “own, fully paid” or “own, being paid”. Equals 0 otherwise (“rented” or “in usufruct or other type of tenure”).
<i>Homicide rate</i>	Homicide rate by 100,000 inhabitants (average 2010–2012). <i>Medicina Legal</i> .
<i>Household with spouse</i>	Equals 1 if household is inhabited by household head and spouse.
<i>Independent</i>	Equals 1 if working independently is the most important job during the previous month.
<i>Justice into own hands</i>	Equals 1 if respondent “totally agrees” or “agrees” with the statement: “When the government does not punish criminals, it is okay that people take justice into their own hands.”
<i>Lands</i>	Equals 1 if respondent reports owning land.
<i>Left and Right Ideology</i>	Equals 1 if respondent reports “left” (“right”) or “center-left” (“center-right”) when asked: “Often, people speak of political leanings from left and right. According to the meaning that the terms ‘left’ and ‘right’ have for you, which political tendency do you sympathize with?”
<i>Left and Right dominated</i>	Equals 1 if municipal mayor belongs to a left- (right-) leaning party. Parties are coded following Fergusson, Querubín, et al. (2017) . Pachón and Sánchez (2014) .

Continued on next page

Table A-1 – Variable definition and sources, continued from previous page

Variable	Description
<i>Male household head</i>	Equals 1 if household head is male.
<i>Neighbor cell phones</i>	Equals 1 if person has the cell phone number of at least half of her neighbors.
<i>Neighbor loans</i>	Equals 1 if a person thinks that at least half of her neighbors would lend her money.
<i>No debts</i>	Equals 1 if household has not any type of debt. Equals 0 otherwise (“debt without credit”, “formal credit”, “informal credit” or “formal-informal credit”).
<i>No sewage</i>	Equals 1 if household dwelling has no sewage system.
<i>Non left-right contender</i>	Equals 1 if municipality had no right-wing or left-wing party (among the top two) in the most recent mayoral election (classification of Fergusson, Querubín, et al. (2017)).
<i>Not in organization</i>	Equals 1 if respondent does not belong to any organization (options included are <i>Juntas de acción comunal</i> , charity organization, community organization, religious organization, organizations supported or promoted by the state, ethnic organization, educational organization, labor union, cooperative of work or union of producers, organization of environment conservation, cultural or sports organization, other).
<i>Negative reciprocity</i>	Equals 1 if respondent “totally agrees” or “agrees” with the statement: “Whoever hurts me, pays for it.”
<i>Nuclear family</i>	Equals 1 household is comprised of: household head and spouse, with or without children; or, household head without spouse but with children).
<i>Other religion</i>	Equals 1 for believers of religions other than Catholic, Evangelical, or Pentecostal (self-declared).
<i>Overcrowded</i>	Equals 1 if ratio of number of residents to number of bedrooms is greater than three in rural households, or greater than or equal to three in urban households.
<i>Own welfare</i>	Equals 1 if respondent “totally agrees” or “agrees” with the statement: “Each individual is responsible for their own welfare.”
<i>Paramilitaries</i>	Number of violent events per year perpetrated by paramilitaries per 100,000 inhabitants (average 2010–2012). <i>Conflict analysis resource center (CERAC) and Universidad del Rosario</i> .
<i>Party identity</i>	Equals 1 if respondent answers “yes” to: “At the moment, do you sympathize with any political party?”
<i>Party recall</i>	Equals 1 if respondent answers “yes” to: “Do you remember which party you vote for mayor of your city or municipality?”
<i>People in household</i>	Number of household residents.

Continued on next page

Table A-1 – Variable definition and sources, continued from previous page

Variable	Description
<i>Persuasion</i>	Equals 1 if respondent answers “frequently,” “sometimes”, or “rarely” (i.e., leaves out “never”) to “During elections, some people try to convince others to vote for a particular party or candidate. How often have you tried to convince others to vote for a party or candidate?”
<i>Polarization</i>	Reynal-Querol (2002) polarization index. $P_j = 1 - \sum_{i=1}^N \pi_{ij} \left(\frac{1/2 - \pi_{ij}}{1/2} \right)^2$, where π_{ij} is the vote share for the mayoral candidate (in 2011) i in municipality j . Pachón and Sánchez (2014).
<i>Pop. density</i>	Population divided by total area (km^2) in the municipality.
<i>Popular vote</i>	Equals 1 if respondent “totally agrees” or “agrees” with the statement: “It is important that rulers are elected by popular vote.”
<i>Positive reciprocity</i>	Equals 1 if respondent “totally agrees” or “agrees” with the statement: “You always have to help those who help you.”
<i>Public Transparency</i>	Summary measure of public transparency (lower levels in this indicator imply a higher risk of corruption) based on three criteria: visibility (management of public information, open data and access to information on the website), transparency in public procurement, and local governments capacity to apply sanctions and compliance delivering information to control and regulation agencies. Computed by <i>Transparencia por Colombia</i> , available as <i>Índice de Transparencia Departamental</i> .
<i>Regions</i>	Regions included in fixed effects. Urban regions in the survey are: <i>Atlántica, Oriental, Central, Pacífica, Bogotá</i> . Rural regions include: <i>Atlántica-Media, Cundi-Boyacense, Eje Cafetero, Centro-Oriente</i> .
<i>Risk of Electoral Fraud</i>	Risk of fraud based on atypical electoral participation, limitations on electoral competition, and atypical levels of null votes and unmarked ballots. Classifies municipalities in four categories: “No Risk”, “Medium risk”, “High Risk” and “Extreme Risk”. In the figure, we aggregate the former two categories as “Low Risk” and the latter two categories as “High Risk”. Index computed by The Colombian Electoral Observation Mission (Misión de Observación Electoral, MOE).
<i>Rural population</i>	Proportion of rural population in the municipality (average 2006–2008). DANE.
<i>Savings</i>	Equals 1 if respondent answers “yes” to: “Do you usually save some of the income you receive?”.
<i>Secret ballot</i>	Equals 1 if respondent answers “yes” to: “Do you think that the ballot is secret?”

Continued on next page

Table A-1 – Variable definition and sources, continued from previous page

Variable	Description
<i>Send help</i>	Equals 1 if household answers “yes” to at least one of the following questions: “During the past 12 months, did any members of the household send money or in-kind aid...” a. “to relatives or friends who live in Colombia?”, b. “to relatives or friends who live abroad?”, c. “for alimony?”, d. “to other persons, entities or organizations?”.
<i>Shock</i>	Equals 1 if household reports any major destabilizing negative event during the previous three years.
<i>Social program beneficiary</i>	Equals 1 if household benefits from any of the following programs: <i>Familias en acción</i> (main conditional cash transfer for the poor with school-age children), programs for the elderly, <i>SENA</i> training programs, <i>Red Juntos - Unidos</i> (program that provides social services to displaced families with the lowest levels of poverty), <i>ICBF</i> programs for children, aid for displaced people, support to households affected by natural disasters, or “other programs”.
<i>Social security</i>	Equals 1 if respondent is affiliated to social security.
<i>State presence</i>	Raw total of local state agencies, local municipality employees, and national-level municipality employees (per capita in 1995). <i>Acemoglu, Garcia-Jimeno, and Robinson (2015)</i> .
<i>Stratum 1, 2</i>	Socio-economic stratum, based on classification of household residence (used to target utility subsidies).
<i>Use of violence</i>	Equals 1 if respondent “totally agrees” or “agrees” with the statement: “Sometimes the use of violence is justified.”
<i>Vote for the same party</i>	Equals 1 if respondent “Always votes for the same party” or “Almost always votes for the same party”. Equals 0 otherwise (“Votes for different parties” or “Always votes blank”).
<i>Voted in the last election</i>	Equals 1 if respondent voted in the last election for mayors (in 2011), 0 otherwise.
<i>Wealth</i>	First principal component following a <i>principal component analysis</i> on a set of reported household assets and dwelling characteristics. See <i>Bernal et al. (2014)</i> .
<i>Win margin</i>	Difference between the vote shares of the winner and runner-up in the 2011 mayoral election. <i>Pachón and Sánchez (2014)</i> .
<i>Woman</i>	Equals 1 if respondent is female.

Notes: Source is *Elca 2013* unless otherwise stated at the end of each description.

Table 2: Summary statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	Observations	Mean	Median	Std. Dev.	Min	Max
Panel A. Dependent variable						
Clientelism	4,962	0.183	0.000	0.387	0.000	1.000
Panel B. Covariates						
Age	4,962	46.641	46.000	12.464	19.000	97.000
Agree with bribery	4,962	0.137	0.000	0.344	0.000	1.000
Authorities violate law	4,962	0.337	0.000	0.473	0.000	1.000
Catholic	4,962	0.820	1.000	0.384	0.000	1.000
Education	4,962	0.457	0.000	0.498	0.000	1.000
Evangelical/Pentecostal	4,962	0.126	0.000	0.331	0.000	1.000
Fractionalization	4,962	0.621	0.641	0.102	0.448	0.834
Gini of land properties	4,962	0.661	0.693	0.132	0.404	0.921
Gov. against inequality	4,962	0.936	1.000	0.244	0.000	1.000
Government role	4,962	0.885	1.000	0.320	0.000	1.000
Guerrillas	4,962	0.208	0.000	1.551	0.000	21.085
HH expenses	4,962	0.092	0.000	0.153	0.000	1.000
Homicide rate	4,962	26.284	18.349	25.282	0.000	163.159
Independent	4,962	0.151	0.000	0.358	0.000	1.000
Justice into own hands	4,962	0.294	0.000	0.456	0.000	1.000
Lands	4,962	0.034	0.000	0.182	0.000	1.000
Left Ideology	4,962	0.055	0.000	0.228	0.000	1.000
Left dominated	4,962	0.087	0.000	0.209	0.000	0.750
Negative reciprocity	4,962	0.189	0.000	0.392	0.000	1.000
Neighbor cell phones	4,962	0.131	0.000	0.337	0.000	1.000
Neighbor loans	4,962	0.199	0.000	0.400	0.000	1.000
Non left-right contender	4,962	0.391	0.000	0.488	0.000	1.000
Other religion	4,962	0.021	0.000	0.144	0.000	1.000
Own welfare	4,962	0.949	1.000	0.220	0.000	1.000
Paramilitaries	4,962	1.440	0.000	7.261	0.000	68.367
Party Identity	4,962	0.163	0.000	0.370	0.000	1.000
Party recall	4,962	0.468	0.000	0.499	0.000	1.000
Persuasion	4,962	0.162	0.000	0.369	0.000	1.000
Polarization	4,962	0.833	0.858	0.090	0.537	1.000
Pop. density	4,962	0.001	0.000	0.002	0.000	0.013
Popular vote	4,962	0.955	1.000	0.206	0.000	1.000
Positive reciprocity	4,962	0.971	1.000	0.167	0.000	1.000
Right Ideology	4,962	0.264	0.000	0.441	0.000	1.000
Right dominated	4,962	0.109	0.000	0.140	0.000	0.750
Rural population	4,962	0.356	0.416	0.275	0.002	0.940
Secret Ballot	4,962	0.701	1.000	0.458	0.000	1.000
Shock	4,962	0.680	1.000	0.467	0.000	1.000
State presence	4,962	-0.060	0.564	1.880	-7.305	3.126
Use of violence	4,962	0.193	0.000	0.395	0.000	1.000
Wealth	4,962	-0.288	-0.719	2.650	-5.296	5.887
Win margin	4,962	0.160	0.131	0.113	0.001	0.423
Woman	4,962	0.575	1.000	0.494	0.000	1.000

Notes: Clientelism is a dummy variable that equals 1 if the respondent has taken into account the benefits, gifts, or jobs that a candidate offered in exchange for his vote. Covariates in Panel B, and their sources, are described in Table A-1. The sample of respondents for is the set of individuals in groups Control 1 and Control 2 as described in Section 2.1, namely, those asked directly about clientelism.

Table 3: Robust correlates of clientelism

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Leamer's		Coefficient		Standard error		CDF(0)		
	Lowest	Upperest	Weighted	Unweighted	Weighted	Unweighted	Weighted*	Weighted	Unweighted
1. Persuasion	0.076	0.152	0.116	0.116	0.015	0.015	1.000	1.000	1.000
2. Agree with bribery	0.044	0.114	0.081	0.082	0.015	0.015	1.000	1.000	1.000
3. Other religion	-0.085	-0.017	-0.047	-0.047	0.009	0.009	1.000	1.000	1.000
4. Negative reciprocity	0.023	0.097	0.064	0.064	0.014	0.014	1.000	1.000	1.000
5. Party recall	0.018	0.100	0.063	0.063	0.013	0.013	1.000	1.000	1.000
6. Wealth	-0.157	-0.027	-0.085	-0.085	0.021	0.021	1.000	1.000	1.000
7. Positive reciprocity	0.005	0.050	0.029	0.029	0.010	0.010	0.999	0.999	0.999
8. Right ideology	-0.010	0.070	0.038	0.038	0.014	0.014	0.993	0.996	0.996
9. Justice into own hands	-0.018	0.066	0.034	0.034	0.013	0.013	0.991	0.996	0.996
10. Age	-0.064	0.009	-0.028	-0.028	0.013	0.013	0.983	0.984	0.984
11. Secret ballot	-0.058	0.006	-0.027	-0.027	0.013	0.013	0.978	0.979	0.979
12. Use of violence	-0.022	0.061	0.030	0.030	0.014	0.014	0.978	0.986	0.986
13. Non left-right contender	-0.008	0.076	0.032	0.032	0.016	0.016	0.974	0.976	0.976
14. Party identity	-0.037	0.065	0.025	0.025	0.015	0.015	0.930	0.951	0.953
15. Authorities violate law	-0.027	0.051	0.020	0.020	0.014	0.014	0.913	0.925	0.926
16. Win margin	-0.141	0.029	-0.020	-0.020	0.016	0.016	0.900	0.903	0.903
17. Polarization	-0.235	0.050	-0.025	-0.025	0.021	0.021	0.882	0.885	0.884
18. Evangelical/Pentecostal	-0.065	0.091	0.015	0.015	0.013	0.013	0.866	0.883	0.882
19. HH expenses	-0.058	0.037	-0.018	-0.018	0.017	0.017	0.847	0.854	0.854
20. Pop. density	-0.090	0.051	-0.020	-0.020	0.020	0.020	0.836	0.846	0.848
21. Fractionalization	-0.229	0.070	0.016	0.016	0.018	0.018	0.813	0.805	0.804
22. Shock	-0.025	0.043	0.013	0.013	0.014	0.014	0.812	0.814	0.815
23. Rural population	-0.118	0.039	-0.021	-0.021	0.024	0.024	0.793	0.812	0.809
24. Homicide rate	-0.040	0.066	0.015	0.015	0.019	0.019	0.792	0.795	0.796
25. Neighbor cell phones	-0.029	0.046	0.011	0.011	0.014	0.014	0.774	0.778	0.781
26. Independent	-0.026	0.041	0.010	0.010	0.014	0.014	0.747	0.749	0.750
27. Neighbor loans	-0.038	0.059	0.012	0.012	0.019	0.019	0.734	0.736	0.739
28. Right dominated	-0.029	0.051	0.008	0.008	0.014	0.014	0.709	0.716	0.715
29. Left dominated	-0.099	0.092	-0.018	-0.018	0.035	0.035	0.695	0.696	0.696
30. State presence	-0.049	0.035	-0.007	-0.007	0.016	0.016	0.677	0.679	0.680
31. Education	-0.044	0.054	0.005	0.005	0.013	0.013	0.665	0.647	0.648
32. Guerrillas	-0.043	0.033	0.003	0.003	0.010	0.010	0.663	0.607	0.613
33. Own welfare	-0.033	0.027	-0.005	-0.005	0.013	0.013	0.643	0.644	0.645
34. Gov. against inequality	-0.035	0.026	-0.005	-0.005	0.014	0.014	0.629	0.629	0.629
35. Government role	-0.030	0.038	0.005	0.005	0.015	0.015	0.624	0.625	0.625
36. Popular vote	-0.039	0.031	-0.005	-0.005	0.016	0.016	0.620	0.621	0.621
37. Catholic	-0.090	0.089	0.003	0.003	0.014	0.014	0.620	0.597	0.598
38. Lands	-0.041	0.033	-0.004	-0.004	0.015	0.015	0.615	0.611	0.610
39. Paramilitaries	-0.033	0.029	0.002	0.002	0.011	0.011	0.593	0.561	0.561
40. Left ideology	-0.044	0.034	-0.002	-0.002	0.013	0.013	0.583	0.564	0.559
41. Woman	-0.036	0.044	0.000	0.000	0.015	0.015	0.562	0.507	0.503
42. Gini of land properties	-0.057	0.044	0.000	0.000	0.017	0.017	0.555	0.507	0.509

Notes: The table reports statistics based on the extreme bounds methodology described in Section 2.2. The dependent variable is a dummy indicating whether, when deciding who to vote for, the respondent has taken into account the benefits, gifts, or jobs that a candidate offered in exchange for the vote. The independent variable of interest is indicated in each row. In each case, a sensitivity analysis is conducted by also including all possible permutations of up to three of the additional covariates listed in the table (That is $\binom{41}{3} + \binom{41}{2} + \binom{41}{1} = 11,521$ regressions per covariate). Region fixed effects are always included, and standard errors are clustered at the community level. Lowest and Upperest are Leamer's lowest and upper extreme bounds for the coefficient of interest (that is, for the resulting coefficient of the variable listed in each row) at the 95% confidence level. The adjusted R-squared is used to weight statistics where indicated. The CDF(0) measures the largest accumulated density of the resulting coefficients of interest to the left or right of zero (whichever is largest). Column 7 reports the weighted CDF(0) assuming no normality, while Column 8 (9) reports the weighted (unweighted) CDF(0) assuming normality. See Table A-1 for a description of all variables.

Table 4: Clientelism: Interaction terms

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Leamer's		Coefficient		Standard error		CDF(0)		
	Lowest	Upperest	Weighted	Unweighted	Weighted	Unweighted	Weighted*	Weighted	Unweighted
1. Right ideology × Right dominated	0.008	0.077	0.042	0.042	0.016	0.016	0.996	0.996	0.996
2. Win margin × Negative rec.	-0.053	0.010	-0.022	-0.022	0.015	0.015	0.930	0.930	0.930
3. Secret ballot × Positive rec.	-0.032	0.011	-0.010	-0.010	0.010	0.010	0.850	0.851	0.850
4. Left ideology × Left dominated	-0.012	0.030	0.009	0.009	0.010	0.010	0.815	0.816	0.816
5. Secret ballot × Negative rec.	-0.044	0.018	-0.012	-0.012	0.015	0.015	0.795	0.795	0.795
6. Win margin × Positive rec.	-0.026	0.019	-0.004	-0.004	0.010	0.010	0.643	0.643	0.642
7. Win margin × Wealth	-0.039	0.043	0.005	0.005	0.016	0.016	0.623	0.617	0.616
8. Secret ballot × Rural population	-0.023	0.031	0.003	0.003	0.012	0.012	0.607	0.607	0.607
9. Win margin × Secret ballot	-0.026	0.032	0.003	0.004	0.013	0.013	0.603	0.603	0.604
10. Win margin × Rural pop.	-0.042	0.052	0.003	0.003	0.017	0.017	0.594	0.579	0.581
11. Secret ballot × Neighbor cell phones	-0.031	0.027	-0.002	-0.002	0.014	0.014	0.550	0.550	0.550

Notes: The table reports statistics based on the extreme bounds methodology described in Section 2.2. For each row, we run a set of regressions varying all possible groups of up to three covariates (see Table 3) in addition to the interaction listed in the row, the lower-order direct terms, and region fixed effects. The dependent variable of interest is a dummy indicating whether, when deciding who to vote for, the respondent has taken into account the benefits, gifts, or jobs that a candidate offered in exchange for the vote. Standard errors are clustered at the community level. Lowest and Upperest are Leamer's lowest and upper extreme bounds for the coefficient of interest (that is, for the resulting coefficient of the variable listed in each row) at the 95% confidence level. The adjusted R-squared is used to weight statistics where indicated. The CDF(0) measures the largest accumulated density of the resulting coefficients of interest to the left or right of zero (whichever is largest). Column 7 reports the weighted CDF(0) assuming no normality, while Column 8 (9) reports the weighted (unweighted) CDF(0) assuming normality. See Table A-1 for a description of all variables.

A.2 Balance between treatment and control groups in list experiments

We corroborate that respondents assigned to treatment and control lists, and the direct question, have similar observable characteristics. For a set of observables \mathbf{X} , we check both the bivariate relationship between group assignment and observables:

$$\Pr(T_i = m) = f(x_i\beta_m) \quad \text{with } x_i \in \mathbf{x},$$

and the multivariate regression,

$$\Pr(T_i = m) = f(\mathbf{x}'\beta_m),$$

where m represents each group (*Treatment*, *Control 1*, and *Control 2*). We estimate the marginal effects of multinomial probit models.

Since randomization was stratified at the regional level, in both types of regressions we include region fixed effects. We also estimated separate regressions for each region, with similar results, but present only these aggregate results to save space. Similarly, we also estimated simple probit and linear probability models for dichotomous indicators of each treatment condition as the dependent variable, and again found no systematic evidence of imbalance.

Table A-2 shows balance using observables in 2010 and Table A-3 in 2013 for the clientelism experiment.

Table A-2: Balance on covariates at the baseline (2010): vote-buying list experiment

Variables	Urban sample						Rural sample					
	Bivariate			Multivariate			Bivariate			Multivariate		
	Treatment	Control 1	Control 2	Treatment	Control 1	Control 2	Treatment	Control 1	Control 2	Treatment	Control 1	Control 2
Age	0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)	0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)	0.001 (0.00)	-0.001** (0.00)	0.001 (0.00)	0.001 (0.00)	-0.001** (0.00)	0.000 (0.00)
Male household head	-0.012 (0.02)	-0.008 (0.02)	0.021 (0.02)	0.012 (0.03)	-0.024 (0.02)	0.012 (0.03)	0.002 (0.02)	-0.012 (0.02)	0.010 (0.02)	0.008 (0.03)	-0.020 (0.03)	0.012 (0.03)
Education	-0.001 (0.00)	0.001 (0.00)	-0.000 (0.00)	0.000 (0.00)	0.002 (0.00)	-0.002 (0.00)	-0.002 (0.00)	0.002 (0.00)	-0.000 (0.00)	-0.000 (0.00)	-0.001 (0.00)	0.001 (0.00)
Employed household head	-0.021 (0.02)	0.010 (0.02)	0.011 (0.02)	-0.018 (0.02)	0.016 (0.02)	0.003 (0.02)	-0.001 (0.02)	0.019 (0.02)	-0.017 (0.02)	-0.000 (0.02)	0.017 (0.02)	-0.016 (0.02)
Savings	-0.009 (0.02)	-0.021 (0.02)	0.030 (0.02)	-0.008 (0.02)	-0.020 (0.02)	0.028 (0.02)	0.012 (0.03)	0.016 (0.02)	-0.028 (0.03)	0.021 (0.03)	0.008 (0.02)	-0.029 (0.03)
Not in organization	0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)	0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)	-0.001 (0.00)	-0.001 (0.00)	0.002* (0.00)	-0.001 (0.00)	-0.001 (0.00)	0.002 (0.00)
Social security	0.006 (0.03)	0.013 (0.03)	-0.019 (0.03)	0.014 (0.03)	0.008 (0.03)	-0.022 (0.03)	0.009 (0.04)	-0.002 (0.04)	-0.007 (0.03)	0.014 (0.04)	-0.008 (0.04)	-0.006 (0.03)
Contributes to social security	-0.028* (0.01)	0.007 (0.01)	0.021 (0.01)	-0.033* (0.02)	0.011 (0.02)	0.021 (0.02)	-0.011 (0.02)	-0.015 (0.02)	0.026 (0.03)	-0.002 (0.03)	-0.028 (0.03)	0.030 (0.03)
Household with spouse	0.022 (0.01)	0.000 (0.02)	-0.022 (0.02)	0.028 (0.03)	-0.023 (0.02)	-0.006 (0.02)	0.006 (0.02)	0.002 (0.02)	-0.008 (0.02)	0.013 (0.03)	0.004 (0.03)	-0.018 (0.02)
Wealth	-0.002 (0.01)	-0.002 (0.01)	0.004 (0.01)	0.006 (0.01)	-0.008 (0.01)	0.002 (0.01)	-0.017** (0.01)	0.008 (0.01)	0.009 (0.01)	-0.013 (0.01)	0.002 (0.01)	0.011 (0.01)
People in household	0.001 (0.00)	0.004 (0.00)	-0.005 (0.00)	0.001 (0.00)	0.001 (0.00)	-0.003 (0.01)	0.005 (0.00)	-0.000 (0.00)	-0.004 (0.00)	0.002 (0.01)	0.003 (0.01)	-0.005 (0.01)
Overcrowded	-0.013 (0.02)	0.020 (0.02)	-0.007 (0.02)	-0.016 (0.02)	0.032 (0.02)	-0.016 (0.02)	-0.020 (0.02)	-0.010 (0.02)	0.030 (0.02)	-0.018 (0.02)	-0.004 (0.02)	0.023 (0.02)
Homeowner	0.003 (0.01)	-0.012 (0.01)	0.009 (0.01)	0.003 (0.02)	-0.020 (0.01)	0.017 (0.01)	-0.008 (0.01)	-0.002 (0.02)	0.009 (0.01)	-0.012 (0.02)	0.004 (0.02)	0.008 (0.02)
No debts	-0.013 (0.01)	-0.005 (0.01)	0.018 (0.01)	-0.009 (0.02)	-0.001 (0.02)	0.010 (0.02)	-0.007 (0.02)	0.033** (0.01)	-0.026* (0.01)	-0.005 (0.02)	0.020 (0.02)	-0.014 (0.02)
Shock	0.001 (0.02)	0.020 (0.01)	-0.020 (0.02)	0.003 (0.02)	0.019 (0.02)	-0.023 (0.01)	-0.015 (0.01)	-0.004 (0.01)	0.019 (0.02)	-0.015 (0.02)	-0.005 (0.01)	0.020 (0.02)
HH expenses (log)	-0.004 (0.01)	0.005 (0.01)	-0.001 (0.01)	0.005 (0.02)	0.008 (0.02)	-0.014 (0.02)	-0.003 (0.01)	0.010 (0.01)	-0.007 (0.01)	-0.015 (0.02)	0.030 (0.02)	-0.014 (0.02)
HH food expenses (log)	-0.003 (0.01)	0.005 (0.01)	-0.002 (0.01)	-0.002 (0.02)	-0.002 (0.02)	0.003 (0.02)	0.004 (0.01)	-0.002 (0.01)	-0.002 (0.01)	0.018 (0.02)	-0.029 (0.02)	0.011 (0.02)
Nuclear family	-0.000 (0.01)	-0.031* (0.02)	0.031 (0.02)	0.016 (0.02)	-0.038* (0.02)	0.023 (0.02)	-0.019 (0.01)	0.021 (0.02)	-0.002 (0.02)	-0.018 (0.02)	0.033 (0.02)	-0.015 (0.02)
Formal credit	-0.016 (0.02)	-0.004 (0.02)	0.020 (0.01)	-0.001 (0.02)	-0.007 (0.02)	0.008 (0.02)	-0.010 (0.02)	0.036** (0.02)	-0.025 (0.02)	-0.001 (0.02)	0.020 (0.02)	-0.019 (0.02)
Get help	-0.020 (0.02)	0.015 (0.02)	0.004 (0.02)	-0.011 (0.02)	0.018 (0.02)	-0.008 (0.02)	0.013 (0.02)	-0.012 (0.02)	-0.001 (0.02)	0.021 (0.02)	-0.021 (0.02)	-0.000 (0.02)
Send help	-0.001 (0.02)	0.018 (0.02)	-0.017 (0.02)	-0.011 (0.02)	0.019 (0.02)	-0.008 (0.02)	0.020 (0.03)	0.011 (0.02)	-0.031 (0.02)	0.016 (0.03)	0.025 (0.02)	-0.041* (0.02)
Social program beneficiary	0.018 (0.01)	0.003 (0.01)	-0.020 (0.01)	0.029 (0.03)	-0.021 (0.03)	-0.008 (0.03)	0.003 (0.02)	0.000 (0.02)	-0.003 (0.01)	-0.008 (0.03)	0.029 (0.02)	-0.021 (0.03)
FEA	0.009 (0.02)	0.012 (0.02)	-0.021 (0.02)	-0.021 (0.04)	0.033 (0.04)	-0.012 (0.03)	0.006 (0.02)	-0.012 (0.02)	0.005 (0.02)	0.008 (0.03)	-0.048* (0.03)	0.041 (0.03)
Stratum 1	0.001 (0.02)	0.003 (0.02)	-0.004 (0.02)	-0.007 (0.02)	0.021 (0.03)	-0.014 (0.03)						
Stratum 2	-0.002 (0.01)	0.007 (0.02)	-0.005 (0.01)	-0.002 (0.02)	0.015 (0.02)	-0.014 (0.02)						
Wealth (rural): quintile 1							0.025 (0.02)	-0.016 (0.02)	-0.009 (0.02)	0.018 (0.04)	-0.035 (0.04)	0.017 (0.04)
Wealth (rural): quintile 2							0.005 (0.02)	0.006 (0.02)	-0.011 (0.02)	0.005 (0.04)	-0.016 (0.03)	0.011 (0.04)
Wealth (rural): quintile 3							0.012 (0.02)	-0.029 (0.02)	0.017 (0.02)	0.011 (0.03)	-0.038 (0.03)	0.027 (0.04)
Wealth (rural): quintile 4							-0.014 (0.02)	0.017 (0.02)	-0.003 (0.02)	-0.006 (0.03)	-0.006 (0.03)	0.011 (0.03)
No sewage							0.020 (0.02)	-0.014 (0.02)	-0.006 (0.02)	0.024 (0.02)	-0.014 (0.02)	-0.010 (0.02)

Notes: The table reports marginal effects from multinomial probit models, with standard errors in parentheses. All regressions include region fixed effects. Treatment refers to respondents who were presented with the list that included a sensitive item, Control 1 received the list without the sensitive item followed by the direct question, and Control 2 was asked the direct question. For variable definitions, see Appendix Table A-1. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table A-3: Balance on covariates at the follow-up (2013): vote-buying list experiment

Variables	Urban sample						Rural sample					
	Bivariate			Multivariate			Bivariate			Multivariate		
	Treatment	Control 1	Control 2	Treatment	Control 1	Control 2	Treatment	Control 1	Control 2	Treatment	Control 1	Control 2
Age	0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)	0.000 (0.00)	0.000 (0.00)	-0.001 (0.00)	-0.000 (0.00)	-0.001 (0.00)	0.001 (0.00)	0.000 (0.00)	-0.001** (0.00)	0.001 (0.00)
Male household head	0.002 (0.01)	-0.006 (0.01)	0.004 (0.01)	0.036* (0.02)	-0.014 (0.02)	-0.021 (0.02)	-0.010 (0.02)	0.003 (0.02)	0.007 (0.02)	-0.007 (0.02)	0.022 (0.02)	-0.015 (0.02)
Education	-0.000 (0.00)	0.001 (0.00)	-0.000 (0.00)	0.002 (0.00)	0.001 (0.00)	-0.003 (0.00)	-0.001 (0.00)	0.001 (0.00)	-0.000 (0.00)	-0.001 (0.00)	-0.001 (0.00)	0.002 (0.00)
Employed household head	-0.028* (0.02)	0.041** (0.02)	-0.013 (0.02)	-0.028 (0.02)	0.047** (0.02)	-0.019 (0.02)	0.029 (0.02)	-0.053*** (0.02)	0.024 (0.02)	0.037* (0.02)	-0.072*** (0.02)	0.035* (0.02)
Savings	-0.019 (0.03)	-0.004 (0.02)	0.023 (0.02)	-0.013 (0.03)	-0.009 (0.02)	0.023 (0.02)	0.006 (0.02)	0.002 (0.02)	-0.008 (0.02)	0.011 (0.02)	-0.004 (0.02)	-0.008 (0.02)
Not in organization	-0.026* (0.02)	0.033** (0.02)	-0.007 (0.02)	-0.017 (0.02)	0.032* (0.02)	-0.014 (0.02)	0.007 (0.02)	-0.012 (0.02)	0.005 (0.02)	0.008 (0.02)	-0.012 (0.02)	0.003 (0.02)
Social security	-0.036 (0.03)	0.074** (0.04)	-0.038 (0.03)	-0.037 (0.04)	0.069* (0.03)	-0.032 (0.03)	-0.032 (0.03)	0.035 (0.03)	-0.002 (0.03)	-0.033 (0.04)	0.024 (0.03)	0.009 (0.04)
Contributes to social security	-0.032** (0.02)	-0.003 (0.01)	0.035** (0.01)	-0.036** (0.02)	-0.002 (0.02)	0.038** (0.02)	0.007 (0.02)	-0.018 (0.02)	0.011 (0.02)	-0.000 (0.02)	-0.016 (0.02)	0.016 (0.02)
Home with spouse	0.015 (0.01)	0.003 (0.01)	-0.017 (0.01)	0.031 (0.02)	-0.004 (0.02)	-0.027 (0.02)	0.024 (0.02)	-0.011 (0.02)	-0.013 (0.02)	0.035 (0.02)	-0.002 (0.02)	-0.033 (0.02)
Wealth	-0.013* (0.01)	0.009 (0.01)	0.003 (0.01)	-0.008 (0.01)	0.015 (0.01)	-0.006 (0.01)	-0.005 (0.01)	0.010 (0.01)	-0.005 (0.01)	-0.002 (0.02)	-0.003 (0.02)	0.005 (0.02)
People in household	0.001 (0.00)	-0.003 (0.00)	0.003 (0.00)	0.000 (0.00)	-0.003 (0.01)	0.003 (0.01)	0.001 (0.00)	0.004 (0.00)	-0.005 (0.00)	-0.000 (0.00)	0.006 (0.00)	-0.006 (0.00)
Overcrowded	-0.035** (0.02)	0.043** (0.02)	-0.008 (0.02)	-0.028 (0.02)	0.038 (0.03)	-0.010 (0.03)	-0.017 (0.02)	0.012 (0.02)	0.004 (0.02)	-0.012 (0.02)	0.023 (0.02)	-0.011 (0.02)
Homeowner	0.001 (0.01)	-0.005 (0.01)	0.003 (0.01)	0.006 (0.02)	-0.011 (0.02)	0.005 (0.01)	-0.022 (0.01)	0.005 (0.01)	0.018 (0.01)	-0.022 (0.02)	0.004 (0.02)	0.017 (0.02)
No debts	0.031** (0.01)	-0.011 (0.01)	-0.020 (0.01)	0.051* (0.03)	-0.033 (0.03)	-0.017 (0.02)	-0.000 (0.01)	0.011 (0.01)	-0.011 (0.01)	-0.029 (0.03)	0.020 (0.02)	0.009 (0.03)
Shock	0.039*** (0.01)	-0.026* (0.01)	-0.013 (0.01)	0.030** (0.01)	-0.015 (0.01)	-0.015 (0.01)	0.021 (0.02)	0.007 (0.02)	-0.029* (0.02)	0.021 (0.02)	0.005 (0.02)	-0.026 (0.02)
HH expenses (log)	-0.008 (0.01)	-0.004 (0.01)	0.012 (0.01)	0.007 (0.02)	-0.028 (0.02)	0.021 (0.02)	0.010 (0.01)	0.010 (0.01)	-0.020* (0.01)	-0.018 (0.02)	0.023 (0.02)	-0.005 (0.02)
HH food expenses (log)	-0.010 (0.01)	0.004 (0.01)	0.006 (0.01)	-0.009 (0.02)	0.022 (0.02)	-0.014 (0.02)	0.022* (0.01)	0.001 (0.01)	-0.023** (0.01)	0.043** (0.02)	-0.025 (0.02)	-0.018 (0.02)
Nuclear family	0.005 (0.01)	0.002 (0.01)	-0.007 (0.02)	0.015 (0.02)	-0.004 (0.02)	-0.011 (0.02)	-0.006 (0.02)	-0.000 (0.02)	0.007 (0.02)	0.002 (0.02)	0.010 (0.02)	-0.012 (0.02)
Formal credit	0.014 (0.01)	0.001 (0.01)	-0.015 (0.01)	-0.025 (0.03)	0.031 (0.03)	-0.006 (0.03)	0.009 (0.02)	0.005 (0.01)	-0.013 (0.01)	0.036 (0.03)	-0.021 (0.02)	-0.014 (0.03)
Get help	-0.029** (0.01)	0.002 (0.01)	0.026* (0.01)	-0.018 (0.02)	-0.004 (0.02)	0.022 (0.02)	0.004 (0.02)	0.003 (0.02)	-0.007 (0.02)	0.008 (0.02)	0.002 (0.02)	-0.010 (0.02)
Send help	-0.022 (0.02)	0.028* (0.01)	-0.006 (0.02)	-0.021 (0.02)	0.028* (0.02)	-0.007 (0.02)	0.000 (0.02)	-0.004 (0.02)	0.004 (0.02)	0.001 (0.02)	-0.004 (0.02)	0.003 (0.02)
Social program beneficiary	0.021 (0.01)	-0.011 (0.01)	-0.010 (0.02)	0.017 (0.02)	-0.009 (0.02)	-0.008 (0.02)	0.008 (0.02)	0.006 (0.01)	-0.014 (0.01)	0.000 (0.02)	0.012 (0.02)	-0.012 (0.02)
FEA	0.014 (0.02)	-0.003 (0.02)	-0.011 (0.02)	-0.008 (0.03)	0.012 (0.03)	-0.005 (0.02)	0.013 (0.02)	-0.007 (0.02)	-0.006 (0.01)	0.012 (0.02)	-0.026 (0.02)	0.014 (0.02)
Stratum 1	-0.010 (0.02)	0.009 (0.02)	0.001 (0.01)	-0.025 (0.02)	0.026 (0.02)	-0.001 (0.02)						
Stratum 2	0.019 (0.02)	-0.004 (0.01)	-0.015 (0.01)	0.012 (0.02)	0.005 (0.02)	-0.017 (0.02)						
Wealth (rural): quintile 1							-0.008 (0.02)	-0.003 (0.02)	0.011 (0.02)	0.005 (0.05)	-0.027 (0.05)	0.022 (0.05)
Wealth (rural): quintile 2							0.034* (0.02)	-0.045** (0.02)	0.010 (0.02)	0.035 (0.05)	-0.056 (0.04)	0.021 (0.04)
Wealth (rural): quintile 3							0.010 (0.02)	0.009 (0.02)	-0.019 (0.02)	0.015 (0.04)	-0.013 (0.03)	-0.002 (0.04)
Wealth (rural): quintile 4							-0.031* (0.02)	0.020 (0.02)	0.011 (0.02)	-0.019 (0.03)	-0.001 (0.03)	0.021 (0.03)
No sewage							0.020 (0.02)	-0.005 (0.02)	-0.015 (0.02)	0.024 (0.02)	-0.008 (0.02)	-0.016 (0.02)

Notes: The table reports marginal effects from multinomial probit models, with standard errors in parentheses. All regressions include region fixed effects. Treatment refers to respondents who were presented with the list that included a sensitive item, Control 1 received the list without the sensitive item followed by the direct question, and Control 2 was asked the direct question. For variable definitions, see Appendix Table A-1. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

A.3 Testing the no design and no liar assumptions in list experiments

In this Appendix we test for the ‘no design’ and ‘no liar’ assumptions in our list experiments, following Blair and Imai (2012). The test for the former compares the predicted average difference in answers to control items under treatment vs. control. With $Y_i(0)$, $Z_{i,J+1}^*$, Y_i and T_i specified as above, let $\pi_{yz} = \Pr(Y_i(0), Z_{i,J+1}^* = z)$ represent the proportion of the population in each type $(Y_i(0), Z_{i,J+1}^*)$. If there are no design effects, these proportions can be computed for all $y = 0, \dots, J$ as follows:

$$\begin{aligned}\pi_{y1} &= \Pr(Y_i \leq y | T_i = 0) - \Pr(Y_i \leq y | T_i = 1), \\ \pi_{y0} &= \Pr(Y_i \leq y | T_i = 1) - \Pr(Y_i \leq y - 1 | T_i = 0).\end{aligned}$$

Proportions π_{y1} and π_{y0} always take positive values. But with design effects, estimated proportions can be negative (for example, see Table 5 in Blair and Imai (2012)). To test for design effects, one can therefore evaluate whether the proportion of the population in each type (π_{yz}) is jointly nonnegative.³⁹ Panel A in Table A-4 shows that no single estimated proportion is negative for either experiment, so the test suggests there is no evidence to reject the null hypothesis of no design effects.

To test the ‘no liar’ assumption, we can evaluate the two most common sources of untruthful answers: ceiling and floor effects. These occur when the respondent engages in either none or all of the behaviors, and thus feels exposed if he or she answers truthfully. In Table 1, the bulk of the answers in the treated lists (93.9%) are larger than zero and smaller than the maximum (five) number of items people can list. This reflects that, since the original instrument design, we included option items that are likely to be negatively correlated with each other, as well as at least one very frequent behavior.

We also test for floor and ceiling effects more formally by estimating the model under the no liar assumption, and comparing it to an alternative model

³⁹This test, however, has limitations: there can be design effects with positive π_{y1} and π_{y0} . Also, a higher probability of positive answers to the sensitive item reduces the likelihood of rejecting the null of no design effects.

allowing for floor and ceiling effects. Based on different information criteria, if the data supports the second model, there is evidence to reject the null of no floor or ceiling effects.⁴⁰ Panel B of Table A-4 reports the results. Regardless of the criterion used, Schwarz's BIC or Akaike's AIC, the preferred model includes no floor or ceiling effects, so this test fails to reject the null of no floor or ceiling effects. Furthermore, these results hold either with covariates (Columns 1 and 2) or when the basic set of covariates in Figure 2 are included.⁴¹

⁴⁰Since the model is identified under the no floor or ceiling effects assumption, we must make additional assumptions to estimate the alternative, allowing for these effects. To do so, we follow Blair and Imai (2012) and consider that respondents' truthful answers to the sensitive item are independent of their answers for control items, conditional upon the pretreatment covariates.

⁴¹We also find similar results using a different set of covariates.

Table A-4: Testing assumptions in the list experiments

	(1)	(2)	(3)	(4)
Panel A: No design effects				
Response value (y)	Estimated proportions with response y to control items and...			
	...not following sensitive behavior ($\hat{\pi}_{y0}$)		...following sensitive behavior ($\hat{\pi}_{y1}$)	
	<i>Estimate</i>	<i>Std. Error</i>	<i>Estimate</i>	<i>Std. Error</i>
0	0.045	0.004	0.020	0.006
1	0.407	0.011	0.055	0.014
2	0.288	0.012	0.053	0.010
3	0.067	0.008	0.015	0.006
4	0.033	0.005	0.015	0.002
Total	0.841		0.159	
P-value	1			
Panel B: No liar effects				
	Information criterion			
	<i>BIC</i>	<i>AIC</i>	<i>BIC</i>	<i>AIC</i>
<i>Clientelism</i>				
No boundary	9873.01	9885.44	9863.88	10050.28
Ceiling	9875.50	9894.14	9894.20	10173.80
Floor	9896.91	9915.55	9897.23	10176.84
Ceiling-Floor	9899.40	9924.26	9927.55	10300.36

Notes: Panel A reports the estimated proportion of respondent types as described in each column title. The design effects test evaluates whether the population proportions are jointly non-negative. For each experiment, the Bonferroni-corrected P-value for the null of no design effects is reported. Panel B reports Schwarz's (BIC) and Akaike's (AIC) information criteria when the model is estimated without including boundaries (No boundary), including ceiling effects (Ceiling), including floor effects (Floor) and including both ceiling and floor effects (Ceiling-Floor). In this panel, the first two columns estimate the models without covariates, while the final two columns include the set of characteristics listed in Figure 2.

A.4 Simple regression analysis

In the main text, we focus on the extreme bounds methodology to examine which variables are robustly correlated with clientelism. This section reports a simpler regression analysis, which produces similar conclusions.

Table A-5 runs linear regressions for clientelism on the same set of variables explored in the text. Odd columns, labeled “bivariate”, show the resulting coefficient for regressions including only one covariate at a time (in addition to region fixed effects, which are always included). Even columns show the coefficient for a multivariate regression, which simultaneously includes all variables listed in the table. As in our baseline analysis, variables are standardized to ease interpretation of the magnitude of the correlations. In the main text we also explored the role of a few interactions between correlates of interest. In Table A-6 we show the results of including such interaction terms in regressions for clientelism that include only region fixed effects and the relevant lower-order uninteracted terms (in the even, “bivariate” columns) as well as in regressions containing the full set of covariates in Table A-5. Again, there are few differences relative to the results using the extreme bounds methodology.

Table A-5: Correlates of clientelism
Simple regression analysis

	(1)	(2)		(3)	(4)
<i>Variables</i>	Bivariate	Multivariate	<i>Variables</i>	Bivariate	Multivariate
Persuasion	0.117*** (0.0149)	0.108*** (0.0153)	Shock	0.0136 (0.0143)	0.00467 (0.0145)
Agree with bribery	0.0831*** (0.0150)	0.0730*** (0.0154)	Rural population	-0.0139 (0.0239)	-0.104*** (0.0224)
Other religion	-0.0471*** (0.00937)	-0.0441*** (0.0164)	Homicide rate	0.0159 (0.0183)	-0.00410 (0.0246)
Negative reciprocity	0.0659*** (0.0141)	0.0509*** (0.0140)	Neighbor cell phones	0.0122 (0.0143)	0.00222 (0.0146)
Party recall	0.0645*** (0.0125)	0.0576*** (0.0131)	Independent	0.0105 (0.0145)	0.00320 (0.0143)
Wealth	-0.0835*** (0.0203)	-0.0976*** (0.0272)	Neighbor loans	0.0127 (0.0186)	0.00882 (0.0204)
Positive reciprocity	0.0292*** (0.00958)	0.0244** (0.00975)	Right dominated	0.00733 (0.0136)	0.0151 (0.0149)
Right Ideology	0.0392*** (0.0143)	0.0215 (0.0146)	Left dominated	-0.0208 (0.0350)	0.0335 (0.0404)
Justice into own hands	0.0378*** (0.0129)	0.00617 (0.0138)	State presence	-0.00712 (0.0159)	-0.0124 (0.0177)
Age	-0.0272** (0.0129)	-0.0261* (0.0142)	Education	0.00405 (0.0131)	0.00976 (0.0151)
Secret ballot	-0.0281** (0.0135)	-0.0216 (0.0133)	Guerrillas	0.00403 (0.00974)	0.00666 (0.00951)
Use of violence	0.0334** (0.0136)	0.00281 (0.0141)	Own welfare	-0.00507 (0.0127)	-2.03e-05 (0.0138)
Non left-right contender	0.0319* (0.0163)	0.0492*** (0.0174)	Gov. against inequality	-0.00456 (0.0140)	-0.00622 (0.0135)
Party identity	0.0271* (0.0148)	-0.000311 (0.0158)	Government role	0.00531 (0.0147)	-0.000734 (0.0142)
Authorities violate law	0.0232* (0.0135)	-0.00652 (0.0145)	Popular vote	-0.00501 (0.0157)	-0.00337 (0.0153)
Win margin	-0.0211 (0.0154)	-0.0684** (0.0308)	Catholic	0.00249 (0.0128)	0.000538 (0.0344)
Polarization	-0.0235 (0.0199)	-0.105** (0.0510)	Lands	-0.00405 (0.0150)	-0.00145 (0.0156)
Evangelical/Pentecostal	0.0128 (0.0120)	0.0122 (0.0290)	Paramilitaries	0.00235 (0.0109)	-0.00784 (0.0123)
HH expenses	-0.0182 (0.0168)	-0.0113 (0.0179)	Left ideology	-0.00102 (0.0131)	-0.0109 (0.0143)
Pop. density	-0.0219 (0.0193)	-0.0292 (0.0249)	Woman	-0.00163 (0.0153)	0.0136 (0.0164)
Fractionalization	0.0170 (0.0169)	-0.109** (0.0553)	Gini of land properties	0.000559 (0.0162)	-0.00799 (0.0227)

Notes: Ordinary least squares regressions. The dependent variable of interest is a dummy indicating whether, when deciding who to vote for, the respondent has taken into account the benefits, gifts, or jobs that a candidate offered in exchange for the vote. Region fixed effects are always included, and standard errors are clustered at the community level. Odd columns, labeled “bivariate”, show the resulting coefficient for regressions including only one covariate at a time. Even columns show the coefficient for a multivariate regression, simultaneously including all variables listed in the table. For variable definitions, see Appendix Table A-1. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table A-6: Clientelism: interaction terms
Simple regression analysis

<i>Variables</i>	(1)	(2)
	Bivariate	Multivariate
Right ideology × Right dominated	0.0416*** (0.0157)	0.0427** (0.0166)
Win margin × Negative rec.	-0.0216 (0.0146)	-0.0226 (0.0146)
Secret ballot × Positive rec.	-0.00980 (0.00955)	-0.0112 (0.00946)
Left ideology × Left dominated	0.00893 (0.0100)	0.00794 (0.00940)
Secret ballot × Negative rec.	-0.0117 (0.0145)	-0.0153 (0.0146)
Win margin × Positive rec.	-0.00332 (0.00984)	-0.00486 (0.00998)
Win margin × Wealth	0.00570 (0.0160)	0.00156 (0.0160)
Secret ballot × Rural population	0.00317 (0.0125)	0.00669 (0.0126)
Win margin × Secret ballot	0.00364 (0.0133)	0.00274 (0.0133)
Win margin × Rural pop.	0.00343 (0.0162)	-0.0137 (0.0186)
Secret ballot × Neighbor cell phones	-0.00172 (0.0141)	-0.00241 (0.0136)

*Notes: Ordinary least squares regressions. The dependent variable of interest is a dummy indicating whether, when deciding who to vote for, the respondent has taken into account the benefits, gifts, or jobs that a candidate offered in exchange for the vote. Standard errors are clustered at the community level. Region fixed effects are always included, and standard errors are clustered at the community level. Column 1 reports the coefficient of a “bivariate regression” containing only the region fixed effects, lower-order uninteracted terms, and the interaction of interest as regressors. Column 2 presents the results of a multivariate regression in which all variables in Table A-5. For variable definitions, see Appendix Table A-1. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.*

Table A-7: Point estimates of the incidence and social desirability bias of clientelism across different covariates

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Urban sample			Rural sample			Full sample		
Variables		List	Direct	SDB	List	Direct	SDB	List	Direct	SDB
Age	(18,40]	0.172*** (0.049)	0.161*** (0.012)	0.012 (0.051)	0.155*** (0.055)	0.209*** (0.014)	-0.054 (0.057)	0.143*** (0.033)	0.182*** (0.009)	-0.039 (0.034)
	(40,55]	0.138*** (0.049)	0.138*** (0.010)	0.000 (0.050)	0.288*** (0.053)	0.238*** (0.013)	0.051 (0.054)	0.218*** (0.035)	0.186*** (0.008)	0.032 (0.036)
	>55	0.219*** (0.065)	0.155*** (0.015)	0.064 (0.066)	0.244*** (0.055)	0.210*** (0.015)	0.035 (0.057)	0.213*** (0.040)	0.185*** (0.011)	0.027 (0.041)
		0.227*** (0.054)	0.154*** (0.012)	0.073 (0.055)	0.247*** (0.038)	0.220*** (0.009)	0.027 (0.039)	0.218*** (0.030)	0.200*** (0.007)	0.019 (0.030)
Education	Primary or less	0.112*** (0.041)	0.145*** (0.011)	-0.033 (0.042)	0.199*** (0.066)	0.218*** (0.017)	-0.019 (0.069)	0.161*** (0.037)	0.168*** (0.009)	-0.008 (0.038)
	Secondary	0.187*** (0.066)	0.152*** (0.014)	0.035 (0.067)	0.172 (0.139)	0.260*** (0.049)	-0.088 (0.146)	0.158*** (0.052)	0.163*** (0.014)	-0.004 (0.054)
	College	0.161*** (0.047)	0.141*** (0.011)	0.020 (0.048)	0.209*** (0.045)	0.232*** (0.012)	-0.023 (0.047)	0.175*** (0.031)	0.191*** (0.008)	-0.016 (0.032)
Gender	Men	0.172*** (0.037)	0.155*** (0.009)	0.018 (0.039)	0.256*** (0.042)	0.210*** (0.011)	0.046 (0.044)	0.203*** (0.028)	0.179*** (0.007)	0.024 (0.029)
	Women	0.147*** (0.051)	0.133*** (0.013)	0.014 (0.053)	0.303*** (0.059)	0.205*** (0.014)	0.098 (0.061)	0.217*** (0.040)	0.171*** (0.009)	0.046 (0.041)
Employment	No	0.176*** (0.036)	0.156*** (0.008)	0.020 (0.037)	0.200*** (0.037)	0.229*** (0.010)	-0.029 (0.039)	0.179*** (0.024)	0.190*** (0.007)	-0.011 (0.025)
	Yes	0.138** (0.053)	0.116*** (0.012)	0.022 (0.054)	0.117 (0.075)	0.161*** (0.018)	-0.044 (0.077)	0.126*** (0.040)	0.132*** (0.010)	-0.006 (0.042)
Frequency of voting	Low	0.181*** (0.035)	0.162*** (0.008)	0.019 (0.036)	0.256*** (0.035)	0.232*** (0.009)	0.025 (0.036)	0.209*** (0.024)	0.199*** (0.006)	0.011 (0.025)
	High	0.149*** (0.036)	0.147*** (0.008)	0.002 (0.037)	0.199*** (0.045)	0.219*** (0.011)	-0.020 (0.047)	0.161*** (0.027)	0.177*** (0.007)	-0.016 (0.027)
Vote for the same party	No	0.213*** (0.051)	0.155*** (0.013)	0.058 (0.052)	0.272*** (0.044)	0.222*** (0.012)	0.050 (0.045)	0.238*** (0.034)	0.196*** (0.009)	0.042 (0.035)
	Yes	0.198*** (0.057)	0.165*** (0.013)	0.033 (0.059)	0.243*** (0.059)	0.257*** (0.017)	-0.014 (0.061)	0.192*** (0.038)	0.205*** (0.010)	-0.013 (0.039)
Secret ballot	No	0.153*** (0.037)	0.142*** (0.008)	0.011 (0.038)	0.231*** (0.036)	0.208*** (0.009)	0.023 (0.037)	0.191*** (0.026)	0.176*** (0.006)	0.015 (0.026)
	Yes	0.104* (0.060)	0.162*** (0.016)	-0.058 (0.062)	0.264*** (0.084)	0.295*** (0.023)	-0.031 (0.087)	0.174*** (0.049)	0.216*** (0.013)	-0.042 (0.051)
Religion	None or other	0.187*** (0.034)	0.146*** (0.008)	0.040 (0.035)	0.231*** (0.034)	0.208*** (0.009)	0.023 (0.035)	0.196*** (0.023)	0.178*** (0.006)	0.018 (0.024)
	Catholic	0.165*** (0.037)	0.129*** (0.008)	0.036 (0.038)	0.229*** (0.037)	0.191*** (0.010)	0.038 (0.038)	0.192*** (0.026)	0.158*** (0.006)	0.034 (0.027)
Skin color	White	0.173*** (0.047)	0.192*** (0.013)	-0.019 (0.049)	0.243*** (0.055)	0.269*** (0.014)	-0.026 (0.056)	0.192*** (0.035)	0.233*** (0.010)	-0.040 (0.037)
	Black	0.140*** (0.041)	0.123*** (0.011)	0.017 (0.043)	0.213*** (0.057)	0.225*** (0.016)	-0.012 (0.059)	0.169*** (0.035)	0.168*** (0.009)	0.002 (0.036)
Shock	No	0.183*** (0.040)	0.164*** (0.009)	0.018 (0.041)	0.242*** (0.036)	0.219*** (0.009)	0.023 (0.038)	0.202*** (0.026)	0.193*** (0.007)	0.009 (0.027)
	Yes	0.186*** (0.043)	0.186*** (0.011)	-0.001 (0.044)	0.229*** (0.041)	0.252*** (0.011)	-0.023 (0.042)	0.221*** (0.029)	0.225*** (0.008)	-0.004 (0.030)
Wealth	Below median	0.152*** (0.042)	0.118*** (0.009)	0.034 (0.043)	0.242*** (0.047)	0.183*** (0.011)	0.059 (0.049)	0.156*** (0.030)	0.136*** (0.007)	0.020 (0.031)
	Above median									

Notes: For each area, rural, urban, and the full sample, the table shows the average incidence of clientelism using the list experiment (with methods as described in Section 2.1), the direct question (using a logit model), and the difference between the two or social desirability bias -SDB- (following Equation (3)). Standard errors are computed using Monte Carlo simulations, and estimations control for the set of variables listed in the table. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level. See Table A-1 for a description of all variables.