Small Aggregates, Big Manipulation: Vote Buying Enforcement and Collective Monitoring

Supplemental Material

(not intended for publication)

November 8, 2015

Table 1: Variable Definitions and Sources

Variable	Description
Armed group	Dummy that takes the value of 1 if there was combat in which either guerrillas or paramilitary forces were involved, or if there was a unilateral military action taken by any of these groups. Source: CERAC.
Closeness central government	Percentage of senators who voted in favor of legislation that the central government supported and who belong to the party of the mayor of the municipality. Two roll call votes are used. The first, from 2004, approved the constitutional change that allowed the first reelection of president Alvaro Uribe. The second, in 2009, decided in favor of holding a referendum for a constitutional change that allowed president Alvaro Uribe to run for office for the third time. Source: *Gaceta del Senado* and author's calculations.
Electorate size	Average of the total valid votes of all races in that particular year. For a regional election year it is the average of the valid votes in mayoral and local council members' races at the municipality level, and of assembly and gubernatorial races at the department level. For a national election year, it is the average of the valid votes of lower house members' races at the department level and of presidential and senatorial races at the national level. Source: National Registrar's Office and author's calculations.
Local revenues	Revenues from the local government as a share of the municipalities' total revenues. Source: National Planning Department.
Margin	Average of all margins of victory in races in a given year weighted by valid votes in each race in a municipality. Margins for plurality elections are calculated as the gap between the winner's and the runner-up's votes. For presidential elections, results of the first round are used. For proportional representation races after 2003, margins are the gap between the electoral quotient of the party winning the final seat and the electoral quotient of the closest loser. Before 2003, it is calculated as the gap between the votes of the party winning the final seat and the votes of the closest loser. Source: National Registrar's Office and author's calculations.
Polling place size	Population 20 years or older per polling place in the municipality. Source: DANE, National Registrar's Office, and author's calculations.
Poverty	The Unsatisfied Basic Needs Index of 1993 and 2005 is linearly interpolated for each municipality. Source: DANE and author's calculations.
Total population	Total population. Source: DANE.

Variable	Description
Age	Respondent's age. Source: LAPOP.
Corruption public sector	The variable is built with answers to the following question: "Taking into account your own experience or what you have heard, corruption among public officials is: very common, common, uncommon, or very uncommon." Variable is increasing in corruption perceptions. Source LAPOP.
Crime victim	The variable takes the value of 1 if the person has been the victim of a crime in the past year and 0 otherwise. Source: LAPOP.
Uninterested politics	The variable is built with answers to the following question: "How much interest do you have in politics: a lot, some, little, or none?" The variable is increasing in lack of interest in politics. Source: LAPOP.
Years education	Number of years of education. Source: LAPOP.
Female	It takes the value of 1 if respondent is a female and 0 otherwise. Source: LAPOP.
Income level	The 2010 and 2011 LAPOP surveys ask respondents to choose an income range out of ten into which their monthly income falls. For the year 2012, LAPOP increased the number of income ranges. For the year 2012, answers were modified to make them compatible with the ones from previous years. The variable is increasing in reported income. Source: LAPOP and author's calculations.
Involved community	Frequency with which the respondent helped to solve a problem in her community or neighborhood in the last year. The answers were originally grouped into four categories (once a week, once or twice a month, once or twice a year, never) and the values assigned to each category are modified to make the variable increasing in involvement in community affairs. Source: LAPOP and author's calculations.
News	The variable is built with the answers to the question "About how often do you pay attention to the news, whether on TV, the radio, newspapers, or the internet?" The answers are grouped into four categories (daily, a few times a week, a few times a month, rarely, never) and the values of each category are modified to make the variable increasing in attention to the news. Source: LAPOP and author's calculations.
Registered voter	It takes the value of 1 if the person is registered to vote and 0 otherwise. Source: LAPOP.
Religiosity	The variable is increasing in self-reported importance of religion. Source: LAPOP and author's calculations.
Trust community	The variable is built with answers to the question "Would you say that people in this community are very trustworthy, somewhat trustworthy, not very trustworthy, or untrustworthy?" Variable is increasing in trust. Source: LAPOP and author's calculations.

Summary Statistics

Table 2: Summary Statistics of Reports of Vote Buying Over Time

	Mean	Std. Deviation	Min.	Max.
		Monitors' reports		
2006	1.33	1.87	0	5
2007	1.08	2.98	0	27
2010	0.05	0.25	0	2
2011	0.09	0.35	0	4
		Citizens' reports		
2006	0.07	0.57	0	15
2007	0.33	1.23	0	22
2010	0.13	0.63	0	10
2011	0.35	1.24	0	29

Table 2 presents summary statistics of the number of citizens' and monitors' reports of vote buying per municipality. We see that while the monitors' reports tend to decrease over time, the citizens' reports have gone up. In particular, monitors have reported less vote buying on average in the general election of 2010 than in the previous general election of 2006. The same pattern holds for local elections in 2007 and 2011. Citizens, on the other hand, have reported an increase in vote buying incidents in general elections and a slight increase in local elections. This pattern is consistent with the fact that in the elections of 2006 and 2007 right wing paramilitaries were more active, which could induce more underreporting by citizens. When comparing the mean number of citizens' reports in 2006 with those in 2007 and the mean number of reports in 2010 with those in 2011, we see that vote buying tends to be more prevalent in local elections. Monitors' reports confirm that pattern for the years 2010 and 2011. This is expected as local elections are more competitive and have smaller electorate sizes. These observations highlight the need to account for competitiveness of the election, electorate size and presence of armed groups in the empirical analysis.

Table 3: Summary Statistics (Municipality-level Analysis)

	Observations	Mean	Std. Deviation	Min.	Max.
Panel A			Dependent Variables		
Vote buying (Citizens)	4,352	0.223	0.98	0	29
Vote buying (Monitors)	1,069	0.281	1.409	0	27
Turnout suppression (Monitors)	1,069	0.03	0.477	0	15
Panel B			Explanatory Variables		
Polling place size	4,352	318.393	84.051	108.046	1,110.75
Armed group	4,352	0.279	0.449	0	1
Closeness Central Government	3,969	0.412	0.433	0	1
Electorate size	4,352	2,734,943	2,522,170	1,585.75	5,565,864
Local revenues	4,352	23.722	24.572	0	100
Margin	4,352	0.104	0.058	0.001	0.546
Poverty	4,352	42.103	21.038	0	100
Total population	4,352	40,698.29	248,446.5	908	7,467,804

Table 4: Summary Statistics (Individual-level Analysis)

Variable	Observations	Mean	Std. Deviation	Min.	Max.
Panel A	Dependent variables				
Vote buying	3,636	0.181	0.385	0	1
Turnout suppression	3,629	0.008	0.088	0	1
Panel B	Explanatory variables				
Polling place size	3,636	644.734	307.841	340.7073	1,397.086
Armed group	3,636	0.445	0.497	0	1
Electorate size	3,636	2,843,710	196,618.2	2,540,350	3,260,987
Local revenues	3,636	43.905	32.866	1.039	100
Margin	3,636	0.116	0.042	0.0145	0.222
Poverty	3,636	22.839	19.769	4.793	100
Total population	3,636	1,603,756	2,673,352	2,726	7,467,806
Age	3,636	36.518	14.507	17	89
Corruption	3,499	3.404	0.79	1	4
Crime victim	3,633	0.203	0.402	0	1
Uninterested politics	3,636	2.861	0.945	1	4
Years education	3,636	9.93	4.559	0	18
Female	3,636	0.493	0.5	0	1
Income level	3,636	4.496	1.891	0	10
Involved community	3,636	1.427	0.781	1	4
News	3,636	3.614	0.777	0	4
Party supporter	3,636	0.292	0.455	0	1
Registered voter	3,636	0.816	0.387	0	1
Religiosity	3,636	3.481	0.8	1	4
Trust community	3,636	2.897	0.887	1	4

Municipality-Level Controls

Controlling for the size of the electorate is important as previous work has identified this variable as negatively correlated with vote buying (Stokes 2005; Stokes et al. 2013; Rueda 2015) and, at the same time, it is possible that areas with smaller electorates would tend to have smaller polling stations.

Whether parties target core or competitive districts with clientelistic benefits remains a contested issue in the literature (e.g. Calvo and Murillo 2004; Stokes 2005; Magaloni 2006; Nichter 2008; Calvo and Murillo 2013; Stokes et al. 2013). Here the objective is not to further explore this question, but rather to account for potential biases caused by the omission of competitiveness. If there is more homogeneity of political preferences or high margins of victory in isolated rural areas, omitting the margin of victory could bias the coefficient on polling place size upwards.

Poverty has also been hypothesized to increase electoral manipulation. Poor voters are cheaper to bribe, tend not to have strong political preferences, and have few years of schooling, all characteristics that facilitate vote buying. More importantly, the omission of income levels could bias the results in favor of my hypothesis, as richer municipalities are likely to have larger polling places. A measure of poverty calculated by DANE for each municipality for 1993 and 2005, the Unsatisfied Needs Index, is linearly interpolated and used as a control.[1] I also include as a control the share of local revenues in the total revenues of the municipality, which is taken from the National Planning Department. This is a second proxy for economic development, as richer municipalities rely less on transfers from the central government. This variable can also serve as a proxy for the rewards of holding office. Elected positions that assign greater control over public resources, should have more

[1] All of the results reported in the paper still hold if this variable is not included as a control.

politicians willing to engage in manipulation to attain them. In Colombia, local governments receive a large share of their revenues from the central government and those resources are tied to particular expenditures (mainly expenditures in health and education), which could reduce the discretion in how they are spent.

Given the importance of non-state armed actors in Colombian politics, the control set includes a variable that takes the value of one if there was combat involving either guerrillas or paramilitary forces in the municipality. The concern, if we do not control for the presence of armed groups, is that they tend to operate in areas where polling places are small. This variable is also included in all models of misreporting, as presence of armed actors that are involved in politics can deter citizens from filing reports about electoral crimes or answering truthfully when they are asked about their vote buying experiences in surveys.

Patterns of Misreporting with Municipality-Level Data

Determinants of Misreporting

Since election monitors often come from outside the communities to file their reports, it is less likely that the factors that affect the citizens' willingness to report crimes also affect the monitors' reporting. Nevertheless, in this section I take a different approach to deal with misreporting that does not rely on monitors being less sensitive than citizens to factors that affect their propensity to report. For this, I use a generalized Negative Binomial mixture Poisson model proposed by Li, Trivedi and Guo (2003). Under this approach, I model the misreporting process and then incorporate the model in the estimation of the parameters of interest when using the citizens' data. According to this model, the true count of vote buying instances, denoted by $y^*_{i,t}$, follows a Negative Binomial distribution with mean $exp(\tilde{\mathbf{x}}_{i,t}\beta)$. Conditional on $y^*_{i,t}$ being zero, $y_{i,t}$, the reported count of crimes, is assumed to be Poisson distributed with mean $exp(\mathbf{z}_{i,t}\phi)$. In a similar way, it is assumed that conditional on $y^*_{i,t}$ being positive, the observed count is Poisson distributed with mean $y^*_{i,t}exp(\mathbf{z}_{i,t}\theta)$. The model allows us to estimate the parameters β, ϕ and θ and hence, to identify what factors determine the true count of vote buying instances, over-reporting and underreporting. A detailed explanation of how the parameters of interest are recovered follows the results.

Panel A of Table 5 shows that the coefficient of polling place size has changed little relative to the original Negative Binomial model presented in the paper. Panels B and C show some interesting results regarding the variables that affect misreporting. A large electorate reduces the incentives to report false crimes. This can be explained by the fact that false accusations carry less weight whenever there are more voters, especially if it is expected that not many others will corroborate or report similar actions. It can also be seen that the presence of armed groups appears to be associated with underreporting. This result is intuitive, as people, out of fear, would avoid contacting the authorities to report irregularities in places where non-state armed actors operate. Larger margins of victory in previous elections also seem to induce underreporting. Voters in less competitive elections may feel that their reports could have even less of an impact if one candidate or party clearly dominated the election.

The previous results should be taken with caution as they rely on the assumptions about the underlying misreporting process being correct. Reassuringly, the next section shows that the predicted misreporting patterns of this approach do not markedly differ from those generated by the multiple imputation (MI) monitors' reports model.

Extent of Misreporting

Besides identifying the determinants of misreporting, the empirical strategy allows us to assess the extent of overreporting, underreporting, and accurate reporting of vote buying. I use the estimated coefficients of the Negative Binomial model in Table 3, column (5) in the paper and those of the true crime equation in Table 5 column (2) to simulate the true

Table 5: Determinants of Manipulation and Misreporting (Citizens' Reports)

Dependent variable:	Citizens' Vote Buying Reports	
	(1)	(2)
Panel A	Actual crimes (β)	
Polling place size	-1.776	-1.559
	(0.459)	(0.391)
Panel B	Misreporting with no crimes (ϕ)	
Poverty	-0.013	-0.009
	(0.004)	(0.005)
Armed group	-0.223	-0.161
	(0.187)	(0.392)
Margin	1.314	0.826
	(1.172)	(1.563)
Electorate size	-0.472	-0.367
	(0.048)	(0.041)
Panel C	Misreporting with crimes (θ)	
Poverty	-0.008	-0.009
	(0.004)	(0.011)
Armed group	-0.299	-0.265
	(0.167)	(0.582)
Margin	-2.220	-7.153
	(1.316)	(1.855)
Electorate size	-0.216	0.141
	(0.065)	(0.065)
Municipality controls	no	yes
Observations	3,968	3,968
Municipalities	1,098	1,098

Coefficients of the Negative Binomial model of true crimes are in Panel A. Coefficients of the Poisson count of reports conditional on not having crimes are in Panel B. Coefficients of the Poisson count of reports conditional on having a positive count of crimes are in Panel C. Misreporting equations include as additional controls the log of the total population and an index of proximity between local and central governments. Polling place size is measured as the log of the population older than 20 per polling station in the previous election. Electorate size is logged. Standard errors clustered at the municipality level are in parentheses.

count of electoral crimes for each observation.[2] I then compare the simulated counts to the citizens' reports to calculate the proportions of underreporting, overreporting, and accurate reporting in the sample.

Table 6: Simulated Proportions of Misreporting and Accurate Reporting

	Overreports	Underreports	Accurate reports
Structural model	0.108	0.084	0.809
	(0.002)	(0.004)	(0.004)
MI Monitors' model	0.071	0.319	0.610
	(0.001)	(0.004)	(0.004)

This table presents the average proportion of overreporting, underreporting, and correct reporting over 500 simulations of the true vote buying incidents count. The simulated true count is compared with the citizens' reports to calculate the proportions. Standard errors of the mean proportions are in parentheses.

Table 6 presents the average proportions and their standard errors for 500 simulated accurate counts. There are several observations. The first is that multiple imputation and the Negative Binomial model that accounts for misreporting give us similar proportions of overreporting. Moreover, both approaches agree that correct reporting constitutes a larger proportion of the observations. This indicates that two completely different methodologies that operate under different assumptions mostly agree in the overall patterns of misreporting. There are, however, some differences. In particular, while the multiple imputed monitor reports' model predicts a significantly larger proportion of underreporting cases than overreporting ones, the structural approach suggests that those proportions are very similar. The result given by the multiple imputation approach is consistent with conversations with election monitors who did not think overreporting was a major concern.

[2]For the multiple imputed coefficients, I use the average coefficients obtained with the completed datasets.

Estimation Details

Following the model's description as discussed above, it can be shown that the probability mass function of the observed count of crimes is

$$f(y_{i,t} \mid \mathbf{x}_{i,t}, \mathbf{z}_{i,t}, \beta, \phi, \theta) = \frac{e^{-\mu^0_{i,t}} {\mu^0_{i,t}}^{y_{i,t}}}{y_{i,t}!} \left(\frac{\nu}{\nu+\lambda_{i,t}}\right)^{\nu} + \sum_{y^*_{i,t}=1}^{\infty} \frac{e^{-y^*_{i,t}\mu_{i,t}}(y^*_{i,t}\mu_{i,t})^{y_{i,t}}}{y_{i,t}!} \frac{\Gamma(y^*_{i,t}+v)}{\Gamma(v)\Gamma(y^*_{i,t}+1)} \left(\frac{\nu}{\nu+\lambda_{i,t}}\right)^{\nu} \left(\frac{\lambda_{i,t}}{\nu+\lambda_{i,t}}\right)^{y^*_{i,t}}$$

where $\mu^0_{i,t} = exp(\mathbf{z}_{i,t}\phi)$, $\lambda_{i,t} = exp(\mathbf{x}_{i,t}\beta)$ and $\mu_{i,t} = exp(y^*_{i,t}\mathbf{z}_{i,t}\theta)$, and ν is one over the overdispersion parameter.

Maximum likelihood estimation cannot be directly implemented for this model, given the presence of the infinite series in the above expression. Li, Trivedi and Guo (2003) propose implementing a Simulated Maximum Likelihood estimator. For this, we require an unbiased simulator for the probability mass function of the observed count, $\widetilde{f}(y_{i,t}, \mathbf{x}_{i,t}, \mathbf{z}_{i,t}, u; , \beta, \phi, \theta)$

$$E[\widetilde{f}(y_{i,t}, \mathbf{x}_{i,t}, \mathbf{z}_{i,t}, u; \beta, \phi, \theta) \mid y_{i,t}, \mathbf{x}_{i,t}, \mathbf{z}_{i,t}] = f(y_{i,t} \mid \mathbf{x}_{i,t}, \mathbf{z}_{i,t}, \beta, \phi, \theta)$$

with the expectation taken over an appropriate distribution of u.

In this case, the simulator is

$$\widetilde{f}(y_{i,t}, \mathbf{x}_{i,t}, \mathbf{z}_{i,t}, u; \beta, \phi, \theta) = \frac{e^{-\mu^0_{i,t}} {\mu^0_{i,t}}^{y_{i,t}}}{y_{i,t}!} \left(\frac{\nu}{\nu+\lambda_{i,t}}\right)^{\nu} + \frac{\frac{e^{-u\mu_{i,t}}(u\mu_{i,t})^{y_{i,t}}}{y_{i,t}!} \frac{\Gamma(u+v)}{\Gamma(v)\Gamma(u+1)} \left(\frac{\nu}{\nu+\lambda_{i,t}}\right)^{\nu} \left(\frac{\lambda_{i,t}}{\nu+\lambda_{i,t}}\right)^{u}}{p(u|\mathbf{x}_{i,t})},$$

where $p(u \mid \mathbf{x}_{i,t})$ is the truncated at zero Negative Binomial distribution with parameters estimated from a "naive" Negative Binomial model of the observed count. The naive Negative Binomial model is the one that does not account for misreporting.

The Simulated Maximum Likelihood estimates are obtained by maximizing the following expression

$$\sum_{t=1}^{T}\sum_{i=1}^{N}\log\frac{1}{S}\sum_{s=1}^{S}\widetilde{f}(y_{i,t},\mathbf{x}_{i,t},\mathbf{z}_{i,t},u_{i,t}^{s};\beta,\phi,\theta),$$

over (β,ϕ,θ). In this expression $u_{i,t}^{s}$, with $s = 1, ...S$, are draws taken from the truncated at zero Negative Binomial distribution $p(u \mid \mathbf{x}_{i,t})$.

Modeling Underreporting with Individual-level Data

Results

Given that underreporting of bribe attempts is still a problem that must be addressed when using individual data, I follow an approach that is similar to the one in the previous section. I estimate a Logit model that accounts for potential misclassification in the dependent variable proposed by Hausman, Abrevaya and Scott-Morton (1998).[3] Table 7 presents the results of a simple Logit and those of the Hausman-Abrevaya-Scott-Morton estimator for models that share the same specification. Besides reporting the coefficient of polling place size, the table presents the coefficients on all individual-level covariates included in the model.[4] We see that the main finding of the paper still holds after accounting for social desirability bias.

The simple Logit results show that those who support a party are more likely to receive a bribe. This evidence supports theories that predict core voters to be the targets of manipulation. However, although the coefficient on supporter is positive after accounting for social desirability bias, it is not precisely estimated. Other characteristics of voters that both models indicate are conducive to receiving a vote buying offer are: being younger, more involved in community affairs, not considering religion very important, and being registered to vote. While the coefficient on being interested in politics is significant in the simple Logit,

[3]For more information on the properties of this estimator see Hug (2010).

[4]Results of municipality-level controls are available upon request.

it is not in the underreporting-corrected model. This is also the case with other variables that capture attitudes towards the democratic processes. Other specifications find that people who are not satisfied with democracy appear to be more likely to be targeted but the result is not robust to accounting for social desirability bias.[5] Moreover, this variable does not seem to affect the probability of underreporting bribes. Since some of these explanatory variables might be determined by vote buying, it is important to be cautious about giving a causal interpretation to the conditional correlations that we observe.

As we did with the analysis of misreporting at the municipality-level, it is possible to assess the extent of underreporting in the LAPOP survey. A model that only includes a constant in the underreporting equation gives a probability of underreporting of 15.30% (with a standard error of 0.0105).

[5]The coefficient on dissatisfaction with democracy when added to model 1 is 0.176 with a standard error of 0.077. For that model the coefficient on polling place size is still negative and significant.

Table 7: Determinants of Manipulation and Underreporting

	(1)	(2)
Panel A	Vote buying attempts (β)	
Polling place size	-3.141	-3.291
	(1.184)	(1.384)
Age	-0.013	-0.037
	(0.006)	(0.008)
Years education	0.013	-0.017
	(0.015)	(0.022)
Female	-0.249	-0.134
	(0.081)	(0.150)
Income level	-0.035	-0.068
	(0.027)	(0.046)
Interest in politics	0.096	0.090
	(0.058)	(0.062)
Involved community	0.176	0.183
	(0.061)	(0.067)
News	0.039	0.012
	(0.076)	(0.079)
Registered voter	0.485	0.395
	(0.127)	(0.153)
Religiosity	-0.121	-0.144
	(0.05)	(0.059)
Rural	-0.189	-0.343
	(0.141)	(0.270)
Party supporter	0.324	0.056
	(0.133)	(0.195)
Trust community	-0.14	-0.162
	(0.058)	(0.073)
Panel B	Underreporting (ξ)	
Age		-0.147
		(0.045)
Corruption		-0.149
		(0.195)
Crime victim		-0.819
		(0.433)
Years education		-0.165
		(0.065)
Female		0.562
		(0.539)
Income level		-0.122
		(0.129)
Party supporter		-1.458
		(0.800)
Municipality controls	yes	yes
Observations	3,636	3,497
Municipalities	77	77

Panel A reports coefficients on individual-level controls in the vote buying equation. Panel B reports coefficients on individual-level covariates included in the underreporting equation for model (2). Additional municipality-level variables in the underreporting equation for model (2) are: presence of armed group, margin of victory, poverty indicator, and size of the electorate. Both models include municipality fixed effects in vote buying equation. Polling place size is the logged number of people older than 20 per polling station. All municipality-level controls are averages of the last two previous elections. Standard errors clustered at the municipality-level are in parentheses.

Estimation Details

The model proposed by Hausman, Abrevaya and Scott-Morton (1998) extends the standard Logit model as follows. Let v_j take the value of 1 whenever respondent j reports being a target of a given type of manipulation and 0 otherwise, and let v_j^* be an unobserved variable that takes the value of 1 only when she has in fact been the target of manipulation. I am interested in learning how certain regressors in $\mathbf{x}_j$ affect the probability of observing v_j^* being 1, and so I assume that $Pr(v_j^* = 1|x_j) = F(\mathbf{x}_j\beta)$, with $F(.)$ being the cumulative logistic distribution. The difference between this and the standard Logit model is that $Pr(v_j = 0|v_j^* = 1)$ is allowed to be positive.[6] This probability is modeled as a function of regressors z_j with $Pr(v_j = 0|v_j^* = 1, z_j) = G(z_j\xi)$, where $G(.)$ is also a cumulative logistic function.[7] The probability of a respondent j answering that she has been the target of manipulation is then

$$Pr(v_j = 1|x_j, z_j) = (1 - G(z_j\xi))F(x_j\beta).$$

The maximum likelihood estimated parameters are obtained by maximizing the log

[6]The original model setup also allows for the possibility of false reporting. For this application I assume that the probability of claiming to be the target of manipulation when the person has not is zero, as there are no clear benefits of lying in this way when answering the survey. The assumption also ensures that a monotonicity condition needed for consistent estimation of β is automatically satisfied. For more details see Hausman, Abrevaya and Scott-Morton 1998, p.242.

[7]For the results presented in the paper, the vector z_j includes: age, years of education, gender, income level, perception of corruption in the public sector, and a dummy variable for whether the person has been a victim of a crime in the previous year.

likelihood function

$$\sum_{j=1}^{M} v_j \ln((1 - G(z_j\xi))F(x_j\beta)) + (1 - v_j)\ln(1 - ((1 - G(z_j\xi))F(x_j\beta)))$$

over (ξ, β).

Vote Buying, Turnout Suppression and Polling Place Size

To confirm that there are differences in the effect of polling place size between vote buying and turnout suppression models, I estimate linear Seemingly Unrelated Equation (SUR) models. These models also allow for a nonzero correlation between the error terms of vote buying and turnout suppression equations. This feature is particularly appealing for this application, as it is likely that the decision to engage in vote buying is not completely independent of turnout suppression efforts by the same party or by its competitors. The results in Table 8 show that, for all our measures of manipulation, we can reject the null hypothesis of the coefficient on polling place size being the same in vote buying and in turnout suppression equations.

Table 8: Vote Buying, Turnout Suppression, and Polling Place Size (SUR models)

Data:	Citizens' reports		Monitors' reports		LAPOP	
	(1)		(2)		(3)	
Equation:	Vote Buying	T. Suppression	Vote Buying	T. Suppression	Vote Buying	T. Suppression
Polling place size	-0.332	-0.104	-0.262	0.094	-0.35	-0.033
	(0.065)	(0.077)	(0.188)	(0.068)	(0.181)	(0.041)
Cultural individual controls	-	-	-	-	yes	yes
Municipality controls	yes	yes	yes	yes	yes	yes
Municipality fixed effects	no	no	no	no	yes	yes
χ^2_1 (H_0: equal αs)		6.49		5.43		2.98
p-value		0.011		0.020		0.084
Observations		4,352		1,068		3,705
Municipalities		1,098		632		81

This table presents coefficients of Seemingly Unrelated Equation linear models. Model (3) includes additional individual controls as well as municipality fixed effects. A Wald test statistic for the null that the coefficient on polling place size is equal across equations and its corresponding p-value are at the bottom of the table. Standard errors are in parentheses.

Other Figures

Figure 1: Discontinuities in Polling Place Size Induced by Institutional Rules

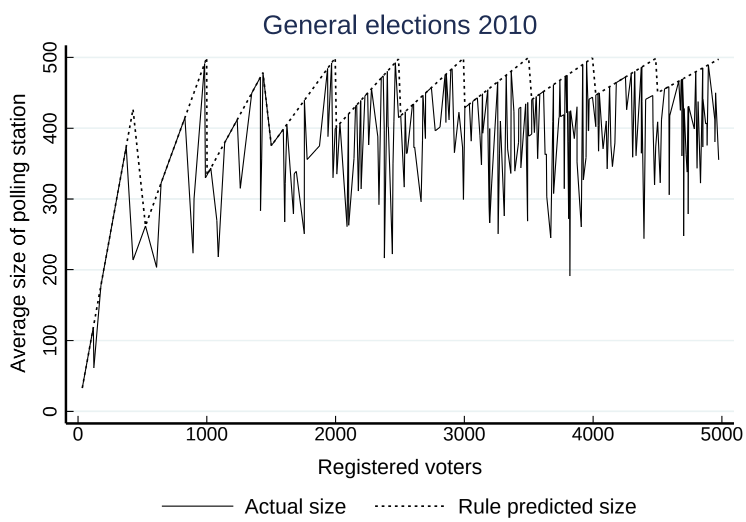
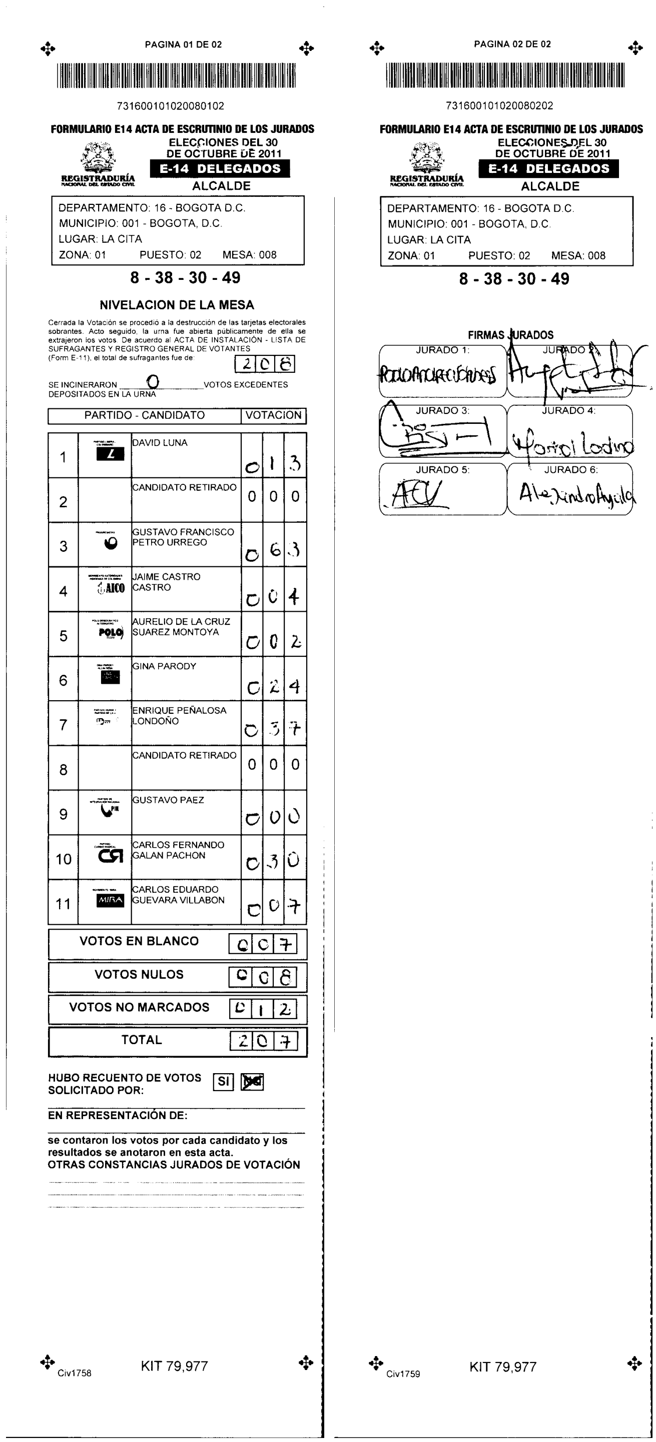


Figure 2: Published Electoral Results Form E14 (Mayoral Elections 2011)



PAGINA 01 DE 02

731600101020080102

FORMULARIO E14 ACTA DE ESCRUTINIO DE LOS JURADOS

ELECCIONES DEL 30 DE OCTUBRE DE 2011

REGISTRADURÍA NACIONAL DEL ESTADO CIVIL

E-14 DELEGADOS

ALCALDE

DEPARTAMENTO: 16 - BOGOTA D.C.
MUNICIPIO: 001 - BOGOTA, D.C.
LUGAR: LA CITA
ZONA: 01 PUESTO: 02 MESA: 008

8 - 38 - 30 - 49

NIVELACION DE LA MESA

Cerrada la Votación se procedió a la destrucción de las tarjetas electorales sobrantes. Acto seguido, la urna fue abierta públicamente de ella se extrajeron los votos. De acuerdo al ACTA DE INSTALACIÓN - LISTA DE SUFRAGANTES Y REGISTRO GENERAL DE VOTANTES (Form E-11), el total de sufragantes fue de: 2 0 8

SE INCINERARON 0 VOTOS EXCEDENTES DEPOSITADOS EN LA URNA

	PARTIDO - CANDIDATO	VOTACION
1	DAVID LUNA	0 1 3
2	CANDIDATO RETIRADO	0 0 0
3	GUSTAVO FRANCISCO PETRO URREGO	0 6 3
4	JAIME CASTRO CASTRO	0 0 4
5	AURELIO DE LA CRUZ SUAREZ MONTOYA	0 0 2
6	GINA PARODY	0 2 4
7	ENRIQUE PEÑALOSA LONDOÑO	0 3 7
8	CANDIDATO RETIRADO	0 0 0
9	GUSTAVO PAEZ	0 0 0
10	CARLOS FERNANDO GALAN PACHON	0 3 0
11	CARLOS EDUARDO GUEVARA VILLABON	0 0 7

VOTOS EN BLANCO	0 0 7
VOTOS NULOS	0 0 8
VOTOS NO MARCADOS	0 1 2
TOTAL	2 0 7

HUBO RECUENTO DE VOTOS SOLICITADO POR: SI ~~NO~~

EN REPRESENTACIÓN DE:

se contaron los votos por cada candidato y los resultados se anotaron en esta acta.
OTRAS CONSTANCIAS JURADOS DE VOTACIÓN

Civ1758 KIT 79,977

PAGINA 02 DE 02

731600101020080202

FORMULARIO E14 ACTA DE ESCRUTINIO DE LOS JURADOS

ELECCIONES DEL 30 DE OCTUBRE DE 2011

REGISTRADURÍA NACIONAL DEL ESTADO CIVIL

E-14 DELEGADOS

ALCALDE

DEPARTAMENTO: 16 - BOGOTA D.C.
MUNICIPIO: 001 - BOGOTA, D.C.
LUGAR: LA CITA
ZONA: 01 PUESTO: 02 MESA: 008

8 - 38 - 30 - 49

FIRMAS JURADOS

JURADO 1: JURADO 2:
JURADO 3: JURADO 4:
JURADO 5: JURADO 6:

Civ1759 KIT 79,977

References

Calvo, Ernesto and Maria V. Murillo. 2004. "Who Delivers? Partisan Clients in the Argentine Electoral Market." *American Journal of Political Science* 48(4):742–757.

Calvo, Ernesto and Maria Victoria Murillo. 2013. "Voters: Assessing Political Linkages Through Partisan Networks and Distributive Expectations in Argentina and Chile." *Comparative Political Studies* 46(7):851–882.

Hausman, Jerry A., Jason Abrevaya and Fiona M. Scott-Morton. 1998. "Misclassification of the Dependent Variable in a Discrete-Response Setting." *Journal of Econometrics* 87:239–269.

Hug, Simon. 2010. "The Effect of Misclassification in Probit Models: Monte Carlo Simulations and Applications." *Political Analysis* 18:78–102.

Li, Tong, Pravin K. Trivedi and Jiequn Guo. 2003. "Modeling Response Bias in Count: A Structural Approach with an Application to the National Crime Victimization Survey Data." *Sociological Methods and Research* 31(4):513–544.

Magaloni, Beatriz. 2006. *Voting for Autocracy: Hegemonic Party Survival and its Demise in Mexico.* New York: Cambridge University Press.

Nichter, Simeon. 2008. "Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot." *American Political Science Review* 102(1):19–31.

Rueda, Miguel R. 2015. "Buying Votes with Imperfect Local Knowledge and a Secret Ballot." *Journal of Theoretical Politics* 27(3):428–456.

Stokes, Susan C. 2005. "Perverse Accountability: A Formal Model of Machine Politics with Evidence from Argentina." *American Political Science Review* 99(3):315–325.

Stokes, Susan C., Thad Dunning, Marcelo Nazareno and Valeria Brusco. 2013. *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. New York: Cambridge University Press.