

Using the Predicted Responses from List Experiments as Explanatory Variables in Regression Models

Kosuke Imai

Department of Politics, Princeton University, Princeton, NJ 08544
e-mail: kimai@princeton.edu (corresponding author)

Bethany Park

Department of Politics, Princeton University, Princeton, NJ 08544
email: bapark@princeton.edu

Kenneth F. Greene

Department of Government, University of Texas, Austin
email: kgreene@austin.utexas.edu

Edited by R. Michael Alvarez

The list experiment, also known as the item count technique, is becoming increasingly popular as a survey methodology for eliciting truthful responses to sensitive questions. Recently, multivariate regression techniques have been developed to predict the unobserved response to sensitive questions using respondent characteristics. Nevertheless, no method exists for using this predicted response as an explanatory variable in another regression model. We address this gap by first improving the performance of a naive two-step estimator. Despite its simplicity, this improved two-step estimator can only be applied to linear models and is statistically inefficient. We therefore develop a maximum likelihood estimator that is fully efficient and applicable to a wide range of models. We use a simulation study to evaluate the empirical performance of the proposed methods. We also apply them to the Mexico 2012 Panel Study and examine whether vote-buying is associated with increased turnout and candidate approval. The proposed methods are implemented in open-source software.

1 Introduction

In the social sciences, the list experiment, also known as the item count technique or unmatched count technique, is becoming increasingly popular as an indirect questioning method to elicit truthful answers to sensitive survey questions. List experiments have been used in various disciplines to measure behaviors as varied as employee theft and drug use (e.g., Wimbush and Dalton 1997; Biemer and Brown 2005). In political science, list experiments have been used to measure socially undesirable attitudes such as prejudice concerning race, religion, and gender (e.g., Kuklinski, Cobb, and Gilens 1997; Kane, Craig, and Wald 2004; Streb et al. 2008), and illicit conduct such as voter fraud and vote-buying (e.g., Corstange 2012b; Gonzalez-Ocantos et al. 2012; Ahlquist, Mayer, and Jackman 2013). List experiments have also been used to study seemingly less sensitive matters that are still prone to measurement error, such as turnout and attitudes toward voting rights (e.g., Corstange 2009; Holbrook and Krosnick 2010). When successfully applied, list experiments can enhance measurement validity by reducing social desirability bias that may result when survey respondents are directly asked sensitive questions (e.g., Tourangeau and Yan 2007).

In recent years, scholars have also made several methodological advances that widen the applicability of list experiments. First, multivariate regression models have been developed to help

Authors' note: The proposed methods are implemented via the open-source software `list: Statistical Methods for the Item Count Technique and List Experiments`, which is available for download at the Comprehensive R Archive Network (<http://cran.r-project.org/package=list>). Supplementary materials for this article are available on the *Political Analysis* Web site. The replication archive is available as Imai, Park, and Greene (2014). We thank Adam Glynn and anonymous reviewers for helpful comments.

researchers identify the respondent characteristics that are associated with certain answers to sensitive questions (Corstange 2009; Imai 2011; Blair and Imai 2012). Second, statistical methods have been developed for testing whether the key assumptions of list experiments are violated and for reducing the potential biases that arise from violation of such assumptions (Blair and Imai 2012). Third, methodological recommendations have been made to improve the design and validity of list experiments (Blair and Imai 2012; Glynn 2013; Blair, Imai, and Lyall 2014).

Despite these advances, no method currently exists for using the predicted responses to sensitive items as an explanatory variable in another regression model. And yet, there are many instances in which researchers wish to use responses to sensitive questions to predict behaviors or opinions. In such cases, the central feature of list experiments becomes a key obstacle. Although the anonymity afforded to respondents enhances measurement validity, researchers do not directly observe their responses to sensitive questions and thus cannot use such responses as predictors in regression models. For example, Stokes (2005) and Stokes et al. (2013) argue that vote-selling affects candidate approval, whereas Nichter (2008) argues that it increases turnout. Other work has shown that list experiments appear to improve the measurement of vote-selling behavior (Gonzalez-Ocantos et al. 2012). Nevertheless, no existing method allows researchers to use this information about vote-selling in a regression model to predict candidate approval and turnout.

We fill this gap in the methodological literature by first improving the performance of a naive two-step estimator, which we show is potentially biased and is too statistically inefficient to be useful in practice (Section 3.3). The improved two-step estimator we propose is relatively easy to implement and is computationally stable. However, this estimator is not fully efficient and, more importantly, it can only be applied to linear models. To overcome these problems, we propose the maximum likelihood (ML) estimator that is statistically efficient and is applicable to a wide range of models (Section 3.4). Both of our estimators build upon the general likelihood framework proposed by Imai (2011) and Blair and Imai (2012). The proposed methods are implemented in the open-source software `list: Statistical Methods for the Item Count Technique and List Experiments`, which offers a comprehensive set of statistical methods for analyzing list experiments (Blair, Imai, and Park 2014).

We conduct a simulation study to assess the empirical performance of the proposed methods (Section 4). We find that especially in a small sample the one-step estimator, though computationally less stable than the improved two-step estimator, has less bias and is much more efficient than the two-step estimators. This improvement in statistical properties is particularly notable regarding the coefficient for the sensitive item. Both estimators significantly outperform the naive two-step estimator in terms of both bias and efficiency.

Finally, we apply the proposed methods to the Mexico 2012 Panel Study (Section 2), focusing on the association between vote-selling and both turnout and candidate approval. This application motivates the development of our new methodology. We first show that the direct question leads to severe under-reporting of vote-selling (around 6%) when compared to the list experiment (around 20%). These differences in measurement are reflected in the relationship between vote-selling and candidate approval. Whereas a standard regression model using answers to the direct question shows that vote-selling is associated with lower candidate approval, our proposed methodology based on the list experiment implies that vote-sellers hold higher approval ratings. We also demonstrate that vote-selling is associated with lower turnout, even among party supporters that the literature identifies as prime targets for turnout buying. This finding suggests that parties may be targeting voters who are unlikely to turn out in the first place. This analysis illustrates that our methodology can allow analysts to evaluate hypotheses about the effects of sensitive attitudes or behaviors that are well measured with list experiments.

2 Empirical Application: The Mexico 2012 Panel Study

In this section, we describe our motivating empirical application: the association between vote-selling and both turnout and candidate approval in Mexico's 2012 general elections. After discussing substantive arguments that drive our subsequent empirical analysis, we introduce the survey instrument and conduct an initial descriptive analysis.

2.1 *Vote-Selling, Turnout, and Candidate Approval*

As scholarly interest in clientelism has burgeoned in political science, a primary task has been to understand the efficacy of vote-buying as a campaign strategy in elections around the world (e.g., Brusco, Nazareno, and Stokes 2004; Calvo and Murillo 2004; Stokes 2005; Penfold Becerra 2007; Magaloni, Diaz-Cayeros, and Estévez 2007; Stokes et al. 2013). At the heart of this inquiry is a concern over how vote-buying affects both candidate approval and turnout, yet appropriate data have been scarce and the proposed associations have been evaluated mainly with bivariate relationships.

The most recent debate in the literature about vote-buying has focused on which voters receive selective benefits. Theoretical work establishes that machines should target swing voters (Lindbeck and Weibull 1987; Stokes 2005). Consistent with this claim, Stokes et al. (2013) show that voters who receive selective benefits in Argentina, Mexico, Venezuela, and India hold higher evaluations of the patron's party. However, they also find that brokers over-target loyalists who already hold positive opinions of the party, indicating that vote-buying may only work for narrow slices of the electorate. In contrast, Nichter (2008) argues that parties use selective benefits to increase turnout among unmobilized supporters. Such supporters may also be more likely to respond to selective benefits (Cox and McCubbins 1986).¹ Thus, the current literature largely agrees that loyal voters receive the most largesse but disagrees about whether selective benefits mainly help increase approval among swing voters or increase turnout among loyalists.

Although these core claims are theoretically compelling, empirical work has been hampered by two problems, which our methodology addresses. First, most works rely on surveys that ask directly whether respondents sold their vote (e.g., Cornelius 2004; Greene 2007; Diaz-Cayeros, Estévez, and Magaloni 2009; Kramon 2009; Carreras and İrepoğlu 2013). The illegality of such exchanges in many countries, including Mexico, likely makes respondents reticent to admit to vote-selling. List experiments offer one solution to this under-reporting problem by allowing respondents to communicate truthful information without openly admitting to illicit behaviors. In their study of vote-buying in Nicaragua, Gonzalez-Ocantos et al. (2012) show that only 2% of respondents admitted they sold their vote when asked directly, whereas almost 25% did so when queried through the list experiment. We follow the same measurement strategy by using a list experiment on vote-selling.

Furthermore, even if list experiments are successfully implemented, researchers cannot simply use this measure to predict turnout or voter opinions. The reason is that we do not observe for each respondent whether they report selling their votes. The very feature of list experiments to provide a degree of anonymity to respondents makes it impossible to estimate standard regression models in a straightforward fashion. Although multivariate regression models have been developed to identify respondent characteristics that are associated with certain answers to sensitive questions (Imai 2011; Blair and Imai 2012), simply taking the predicted probabilities from these models and using them as an explanatory variable in the outcome regression, as shown later, fail to recover the true coefficient and result in the underestimation of standard error. To the best of our knowledge, no method exists to appropriately use the responses from list experiments as explanatory variables in regression models to predict other outcomes, such as (in this case) turnout or opinions about the candidates. The goal of our article is to fill this methodological gap.

2.2 *The List Experiment*

We measure vote-selling and turnout using a list experiment embedded in the Mexico 2012 Panel Study, designed and fielded by Greene et al. (2012). Allegations of vote-buying featured prominently in the presidential election of that year. Enrique Peña Nieto of the formerly dominant Institutional Revolutionary Party (PRI)—a party with a long history of electoral clientelism (Cornelius and Craig 1991)—won 39% of the vote, besting Andrés Manuel López Obrador from

¹But see Dixit and Londregan (1996), who provide a general framework that generates two sets of conditions: those under which parties will favor their core voters and those under which parties will target swing voters.

a leftist coalition led by his Party of the Democratic Revolution (PRD) with 32%, and Josefina Vázquez Mota of the incumbent conservative National Action Party (PAN) with 26%. López Obrador claimed that Peña Nieto won through a massive vote-buying scheme (see Greene [2014] for details).

The 2012 panel study interviewed a nationally representative sample of ordinary citizens with a valid voter registration card, before and after the election. All interviews were face-to-face. In this article, we focus on the data from the postelection survey with a total of 1150 respondents who answered the list experiment questions. Each respondent was randomized either to the control or to the treatment group. The script for the control group was read aloud, as follows:

I am going to read you a list of three activities that appear on this card and I want you to tell me how many of these activities you have done in recent weeks. Please don't tell me which ones, just HOW MANY. The three activities are... [SHOW CARD AND READ]

- (a) See television news that mentions a candidate
- (b) Attend a campaign event
- (c) Talk about politics with other people

Respondents assigned to the treatment group were read aloud the following script that includes an additional sensitive item about vote-selling (item c below):

I am going to read you a list of four activities that appear on this card and I want you to tell me how many of these activities you have done in recent weeks. Please don't tell me which ones, just HOW MANY. The four activities are... [SHOW CARD AND READ]

- (a) See television news that mentions a candidate
- (b) Attend a campaign event
- (c) Exchange your vote for a gift, favor, or access to a service
- (d) Talk about politics with other people

We emphasize that the sensitive item specifically asks respondents whether they “exchanged” their vote, not whether they “received” a benefit. We believe that this wording better measures vote-selling by eliminating exchanges that respondents did not deem worth their vote. Campaigns routinely deliver token benefits to voters such as pins, stickers, key chains, and the like. If the data scored the receipt of benefits that voters consider to be tokens as vote-selling, we would likely underestimate the effect of vote-selling on outcomes of interest.

Table 1 summarizes the data from this list experiment and the direct question approach. Estimating the proportion of vote-sellers using the list experiment is straightforward. Because the two subsamples were selected at random, the mean number of nonsensitive activities that respondents affirmed should be equal across the two lists, implying that any difference in means is attributable to vote-selling. An obtrusive measure of vote-selling was also asked later in the survey; however, it is important to note that this question asked whether the respondent was offered a gift, service, or other benefit in exchange for their vote, not whether they accepted such an offer. The lower bar for a positive response to the obtrusive measure should at least narrow the gap between it and the list experiment. Nevertheless, the difference in results from the two measurement approaches is striking. The obtrusive measure indicates that just 5.5% of respondents received a benefit, whereas the list experiment reveals that a whopping 19.4% sold their vote. The confidence intervals are [4.2, 6.8] for the obtrusive measure and [18.1, 20.7] for the list experiment estimate, respectively.

The validity of the estimate from the list experiment relies upon the assumption that respondents do not lie about the sensitive item and that adding the sensitive item does not change responses to the control item list (Imai 2011; Blair and Imai 2012). Although these assumptions are not directly testable, Blair and Imai (2012) propose a statistical test that can be used to detect certain violations.

Table 1 Summary of the responses to list experiment and direct questioning about vote-selling

Response value	List experiment				Direct question		
	Control group		Treatment group		Response value	Frequency	Proportion
	Frequency	Proportion	Frequency	Proportion			
0	33	5.7%	39	6.8%	No	1087	94.5%
1	256	44.5	208	36.1	Yes	63	5.5
2	197	34.3	197	34.3			
3	73	12.7	89	15.5			
4			28	4.9			
Non-response	16	2.8%	14	2.4%	Non-response	0	0%
Total	575	100	575	100	Total	1150	100

According to the simple difference-in-means estimation, the list experiment indicates that 19.4% (with the 95% confidence interval of [18.1, 20.7]) of respondents sold their votes, which is much higher than the corresponding figure (5.5% with [4.2, 6.8]) from the direct question.

This generates a p -value of 0.27, indicating no clear violation of the required assumptions. Like any statistical test, this result may be due to the lack of statistical power. Given this promising initial result, however, we proceed to develop a new statistical methodology, which allows researchers to use the responses from list experiments as predictors in regression models.

3 The Proposed Methodology

In this section, we propose methods to use the responses from list experiments as predictors in an outcome regression model. We begin by briefly reviewing the multivariate regression analysis of list experiments proposed by Imai (2011) and Blair and Imai (2012). We then propose two new estimators. First, we describe a two-step estimator, which uses the estimates from the aforementioned multivariate regression model and improves the performance of a naive two-step estimator. Despite its simplicity, the major limitation of this approach is that it cannot be generalized beyond simple linear regression models. The estimator is also not statistically efficient, often resulting in large standard errors. To overcome these problems, we propose a general ML estimator that can be computed in a single step. Although this estimator is more computationally challenging, it is fully efficient and is applicable to a wide range of statistical models.

3.1 The Setup

Suppose that we have a simple random sample of N respondents from a population. In the simplest list experiment, each respondent is randomly assigned to either the treatment group, $T_i = 1$, or the control group, $T_i = 0$. The respondents in the control group are asked to report the number of affirmative items from a list of J control items. Often, these control items are not sensitive, but that is not a requirement. In contrast, the respondents in the treatment group are asked to count the number of affirmative items in the list of $J + 1$ items, containing the same J control items and one sensitive item of interest. Let Y_i represent the reported count for this list experiment. For the respondents in the control group, Y_i takes an integer between 0 and J . For those in the treatment group, Y_i is an integer ranging from 0 to $J + 1$. Finally, we observe a vector of K covariates for each respondent, which is denoted by a column vector $X_i \in \mathcal{X}$, where \mathcal{X} is the support of X_i .

The goal of this article is to develop new methods for estimating an outcome regression model, whose predictors include the latent (i.e., unobserved) response to the sensitive item from a list experiment. To formalize this, we follow the notation used in Imai (2011) and Blair and Imai (2012) and denote the latent response to the j th item in the list as Z_{ij}^* . Without a loss of generality, we assume that the sensitive item of interest appears at the end of the list, with $Z_{i,J+1}^*$ representing the latent response to the sensitive item. In practice, the order of items in list experiments is often randomized to minimize the ordering effect, and the proposed methodology is applicable regardless

of a particular ordering of items. Finally, we use $Y_i^* = \sum_{j=1}^J Z_{ij}^*$ to denote the number of affirmative responses to the J control items, which is observed for the respondents in the control group but may be unknown for those in the treatment group. Thus, the relationship between the observed outcome and these (possibly) latent variables is given by

$$Y_i = Y_i^* + T_i Z_{i,J+1}^*. \quad (1)$$

Throughout this article, as mentioned in Section 2.2, we assume no design effects and no liars; these assumptions are formally described in Imai (2011) and Blair and Imai (2012).

Under this setup, researchers may be interested in using the latent responses from list experiments as predictors in regression models for an outcome variable $V_i \in \mathcal{V}$, where \mathcal{V} is the support of V_i . For example, one of the simplest such models is the following linear regression model:

$$V_i = \alpha + \beta^\top X_i + \gamma Z_{i,J+1}^* + \zeta Y_i^* + \epsilon_i, \quad (2)$$

where the usual exogeneity assumption is made with respect to both observed and unobserved variables, that is, $\mathbb{E}(\epsilon_i | X_i, Z_{i,J+1}^*, Y_i^*) = 0$. Under this model, the main parameter of interest is the coefficient for the latent response to the sensitive item, that is, γ . Beyond this linear regression model, researchers may be interested in estimating nonlinear models such as logistic regression, where the responses from list experiments are used as predictors. The problem, of course, is that we do not directly observe $Z_{i,J+1}^*$ and Y_i^* for many respondents. The goal of this article is to show how to estimate these outcome regression models in a principled manner.

3.2 The Multivariate Regression Models for List Experiments: A Review

Next, we briefly review the multivariate regression models developed by Imai (2011) and Blair and Imai (2012) to analyze list experiments. These models form the basis of our new methodology introduced in Sections 3.3 and 3.4. Imai (2011) and Blair and Imai (2012) propose a class of the multivariate regression models that are represented by the following two submodels:

$$g_\delta(x) = \Pr(Z_{i,J+1}^* = 1 | X_i = x; \delta) \quad (3)$$

$$h_\psi(y | x, z) = \Pr(Y_i^* = y | X_i = x, Z_{i,J+1}^* = z; \psi) \quad (4)$$

for $z = 0, 1$, $y = 0, 1, \dots, J$, and $x \in \mathcal{X}$, where δ and ψ are vectors of unknown parameters.

The first submodel given in equation (3) models the probability of an affirmative response to the sensitive item, whereas the second submodel given in equation (4) represents a model for the response to the J control items given the response to the sensitive item. Both models include a vector of covariates X_i . A common submodel for the sensitive item is logistic regression, that is, $g_\delta(x) = \exp(\delta_0 + x^\top \delta_1) / \{1 + \exp(\delta_0 + x^\top \delta_1)\}$ and $\delta = (\delta_0, \delta_1)$. Similarly, a binomial model with a logistic link is often used as the submodel for the control items, that is, $h_\psi(y | x, z) = \binom{J}{y} p(x, z)^y \{1 - p(x, z)\}^{J-y}$, where $p(x, z) = \exp(\psi_0 + \psi_1 z + x^\top \psi_2) / \{1 + \exp(\psi_0 + \psi_1 z + x^\top \psi_2)\}$ and $\psi = (\psi_0, \psi_1, \psi_2)$.

Imai (2011) and Blair and Imai (2012) show how to obtain consistent estimates of δ and ψ using nonlinear least squares and ML estimation methods. In particular, they develop an expectation-maximization (EM) algorithm for computing fully efficient ML estimates.

3.3 The Two-Step Estimator

We first consider two-step estimators using the multivariate regression model described in Section 3.2 as the first step. The main advantage of this estimator is its simplicity. However, the estimator is only applicable when the outcome model is linear. In addition, the simplicity of the estimator comes at the cost of statistical efficiency, as we demonstrate through a simulation study in Section 4. We show that our proposed two-step estimator improves the performance of a more naive two-step estimator.

We consider the linear outcome regression model given in equation (2). Note that the treatment variable T_i is not included as a regressor in this model because the randomization makes it independent of $(X_i, Z_{i,J+1}^*, Y_i^*)$. The inclusion of T_i , although possible, does not affect large sample statistical properties of our proposed methodology under the standard assumptions of list experiments. Moreover, it is also possible not to include Y_i^* as a regressor. Here, we simply include it for completeness. All of our theoretical results apply to the model without this variable.

3.3.1 The naive two-step estimator

We begin by the examination of a naive two-step estimator, which is an approach applied researchers may consider. In this approach, we first fit the multivariate regression models, such as the one described in Section 3.2. Using the fitted model, we then compute, for each respondent, the predicted probability of affirmatively answering the sensitive item as a function of respondents' characteristics. We denote this quantity by $\hat{Z}_{i,J+1}^* = g_{\hat{\delta}}(X_i)$ (see equation (3)). Similarly, if Y_i^* is included in the outcome regression model, we may also estimate the expected response to the control items, that is, $\hat{Y}_i^* = \sum_{y=1}^J y \{h_{\hat{\psi}}(y | X_i, 1) \hat{Z}_{i,J+1}^* + h_{\hat{\psi}}(y | X_i, 0)(1 - \hat{Z}_{i,J+1}^*)\}$. In the second step, we include these predicted values in the outcome regression. For example, we would fit the linear regression model given in equation (2) by replacing $Z_{i,J+1}^*$ and Y_i^* with their predicted values, $\hat{Z}_{i,J+1}^*$ and \hat{Y}_i^* , respectively.

Unfortunately, this naive two-step estimator suffers from several problems. First, the estimator will be biased if the outcome regression is nonlinear. Second, the usual standard error will be underestimated because it will not take into account the first-stage estimation uncertainty, which is typically large. Third, and perhaps most importantly, even if the outcome regression model is linear so that the estimator is unbiased, the naive two-step estimator is too statistically inefficient to be useful in practice. The reason is that the predicted values, $\hat{Z}_{i,J+1}^*$ and \hat{Y}_i^* , are highly correlated with other regressors in the model, X_i . In fact, if a simple linear regression is used in the first stage, the model cannot be fitted, due to perfect collinearity. One way to circumvent this problem is to exclude some covariates from the second-stage regression by treating them as instrumental variables. However, the assumption of such exclusion restriction is often difficult to justify.

3.3.2 The new two-step estimator

We improve the naive two-step estimator by utilizing all the information from the list experiment. As stated above, we focus on the case of linear outcome regressions. The naive two-step estimator is based on the conditional expectation of the outcome given X_i . Here, we condition on all the information from the list experiment, that is, (Y_i, T_i) , as well as the characteristics of respondents, X_i . This yields the following regression equation:

$$\mathbb{E}(V_i | X_i, T_i, Y_i) = \alpha + \beta^\top X_i + \gamma \Pr(Z_{i,J+1}^* = 1 | X_i, T_i, Y_i) + \zeta \{Y_i - T_i \Pr(Z_{i,J+1}^* = 1 | X_i, T_i, Y_i)\}, \quad (5)$$

where we used the relationship given in equation (1) as well as the fact that the exogeneity assumption of the original regression model given in equation (2), that is, $\mathbb{E}(\epsilon_i | Z_{i,J+1}^*, Y_i^*, X_i) = 0$, implies $\mathbb{E}(\epsilon_i | X_i, T_i, Y_i) = \mathbb{E}\{\mathbb{E}(\epsilon_i | Z_{i,J+1}^*, Y_i^*, T_i, X_i) | X_i, T_i, Y_i\} = \mathbb{E}\{\mathbb{E}(\epsilon_i | Z_{i,J+1}^*, Y_i^*, X_i) | X_i, T_i, Y_i\} = 0$. Furthermore, using Bayes's rule, the unknown probability in equation (5) can be written as

$$\Pr(Z_{i,J+1}^* = 1 | X_i, T_i, Y_i) = \frac{h_{\psi}(Y_i - T_i | 1, X_i)g_{\delta}(X_i)}{h_{\psi}(Y_i - T_i | 1, X_i)g_{\delta}(X_i) + h_{\psi}(Y_i | 0, X_i)\{1 - g_{\delta}(X_i)\}}. \quad (6)$$

Under this approach, the linear regression model given in equation (5) can be fitted via the following two-step procedure. First, we estimate the unknown parameters (δ, ψ) from the multivariate regression models $g_{\delta}(x)$ and $h_{\psi}(y | x, z)$ given in equations (3) and (4) using the estimation

methods described in Imai (2011) and Blair and Imai (2012). Replace $g_\delta(x)$ and $h_\psi(y|x, z)$ in equation (5) with their estimates, $g_\delta(x)$ and $h_\psi(y|x, z)$, and then fit this outcome linear regression model. We emphasize that the estimate of the probability given in equation (6) is readily available from the E-step of the EM algorithm developed by Imai (2011) and Blair and Imai (2012). The standard errors that incorporate the first-stage estimation uncertainty can be calculated using either the bootstrap resampling method or analytically based on the (straightforward but tedious) method of moments framework (Imai 2011).

3.4 The General Full ML Estimator

Although the two-step estimator developed above improves the naive estimator, it still suffers from the fact that the estimator is only applicable when the outcome model is linear. In addition, the estimator is not yet statistically fully efficient because the first step does not use the information contained in the outcome regression. Statistical efficiency is an important consideration when analyzing list experiments because indirect questioning techniques such as list experiments reduce bias at the cost of losing statistical efficiency.

To address this limitation, we develop the general full ML estimator. This one-step estimator incorporates all the information from the data in the likelihood framework and therefore is fully efficient. The estimator is also general in that it can be extended to a wide range of outcome models.

Consider the general outcome model defined as follows:

$$f_\theta(V_i | X_i, Y_i^*, Z_{i,J+1}^*). \quad (7)$$

For example, if the normal linear regression model is assumed, we will have

$$f_\theta(V_i | X_i, Y_i^*, Z_{i,J+1}^*) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{1}{2\sigma^2}(V_i - \alpha - \beta^\top X_i - \gamma Z_{i,J+1}^* - \zeta Y_i^*)^2\right\}, \quad (8)$$

where $\theta = \{\alpha, \beta, \gamma, \zeta, \sigma\}$. Given this outcome model, we can write the observed-data full likelihood function as

$$\begin{aligned} & L_{\text{obs}}(\theta, \delta, \psi | \{T_i, V_i, X_i, Y_i\}_{i=1}^n) \\ &= \prod_{i=1}^n [f_\theta(V_i | X_i, Y_i, 0)h_\psi(Y_i | X_i, 0)\{1 - g_\delta(X_i)\}]^{\mathbf{1}\{Y_i=0\}T_i} \\ & \quad \times [f_\theta(V_i | X_i, Y_i - 1, 1)h_\psi(Y_i - 1 | X_i, 1)g_\delta(X_i)]^{\mathbf{1}\{Y_i=J+1\}T_i} \\ & \quad \times [f_\theta(V_i | X_i, Y_i - 1, 1)h_\psi(Y_i - 1 | X_i, 1)g_\delta(X_i) + f_\theta(V_i | X_i, Y_i, 0)h_\psi(Y_i | X_i, 0)\{1 - g_\delta(X_i)\}]^{\mathbf{1}\{1 \leq Y_i \leq J\}T_i} \\ & \quad \times [f_\theta(V_i | X_i, Y_i, 1)h_\psi(Y_i | X_i, 1)g_\delta(X_i) + f_\theta(V_i | X_i, Y_i, 0)h_\psi(Y_i | X_i, 0)\{1 - g_\delta(X_i)\}]^{1-T_i}. \end{aligned} \quad (9)$$

Unfortunately, maximizing the logarithm of this observed-data likelihood function is difficult due to its complicated mixture structure. However, as done in the multivariate regression models for the list experiments described in Section 3.2, the EM algorithm can be developed to solve this optimization problem.

To do this, first we derive the complete-data likelihood function, which takes the following much simpler form (when compared to the observed-data likelihood function given in equation (9)),

$$\begin{aligned} & L_{\text{com}}(\theta, \delta, \psi | \{T_i, V_i, X_i, Y_i, Z_{i,J+1}^*\}_{i=1}^n) \\ &= \prod_{i=1}^n \{f_\theta(V_i | X_i, Y_i - T_i, 1)h_\psi(Y_i - T_i | X_i, 1)g_\delta(X_i)\}^{Z_{i,J+1}^*} \\ & \quad \times [f_\theta(V_i | X_i, Y_i, 0)h_\psi(Y_i | X_i, 0)\{1 - g_\delta(X_i)\}]^{1-Z_{i,J+1}^*}. \end{aligned} \quad (10)$$

In each iteration, the M-step of the EM algorithm maximizes the conditional expectation of the complete-data log-likelihood function, which is given here:

$$\sum_{i=1}^n \tilde{w}(X_i, Y_i, T_i, V_i) \cdot \{\log g_\delta(X_i) + \log f_\theta(V_i | X_i, Y_i - T_i, 1) + \log h_\psi(Y_i - T_i | X_i, 1)\} \\ + (1 - w_i) [\log \{1 - g_\delta(X_i)\} + \log f_\theta(V_i | X_i, Y_i, 0) + \log h_\psi(Y_i | X_i, 0)], \quad (11)$$

where the weight $\tilde{w}(X_i, Y_i, T_i, V_i)$ is given by the following E-step,

$$\Pr(Z_{i,J+1}^* = 1 | T_i = t, V_i = v, X_i = x, Y_i = y) \\ = \frac{f_\theta(v | x, y - t, 1) h_\psi(y - t | x, 1) g_\delta(x)}{f_\theta(v | x, y - t, 1) h_\psi(y - t | x, 1) g_\delta(x) + f_\theta(v | x, y, 0) h_\psi(y | x, 0) \{1 - g_\delta(x)\}}. \quad (12)$$

In the algorithm, this weight is evaluated using the values of the parameters (θ, δ, ψ) obtained at the previous iteration. Notice that this weight conditions on the outcome variable V_i , whereas the weight used in the two-step estimator given in equation (6) does not. This difference is the source of relative efficiency gain for this one-step estimator. Our EM algorithm starts after suitable starting values are selected for all parameters, and it is iterated until convergence. The standard errors are calculated analytically by computing the sample average of the cross-product of score vectors. The detailed expression of this analytical asymptotic variance is provided in Supplementary Appendix A.1.

Despite its attractive properties and flexibility, the proposed ML estimator has a disadvantage over the two-step estimator in that it is more difficult to compute. The likelihood function exhibits a complicated mixture structure, and as such the EM algorithm may end up in a local maximum. Moreover, adding too many predictors may lead to the computational instability. A possible solution is to develop a Markov chain Monte Carlo algorithm, but we leave this to future research.

4 A Simulation Study

We now conduct a simulation study to compare the two estimators proposed above, as well as the naive two-step estimator, using a linear regression outcome model (so that all three estimators can be compared on an equal footing). We sample a single covariate X_i from a univariate normal distribution with mean -1 and variance 1 . For the sensitive model $g_\delta(X_i)$, we use logistic regression with an intercept and X_i , where the true values of the coefficients are 0.5 and 1 , respectively. For the responses to the control items, we use binomial logistic regression $h_\psi(Y_i | X_i, Z_{i,J+1}^*)$, where the true values of the coefficients for an intercept X_i and the latent variable $Z_{i,J+1}^*$ are both equal to 1 . Finally, the outcome model $f_\theta(V_i | X_i, Y_i^*, Z_{i,J+1}^*)$ is linear regression, where the true values for the coefficients, corresponding to an intercept, X_i , Y_i^* , and $Z_{i,J+1}^*$, are equal to $(-1, 1, 1, 0.5)$.

We examine three sample sizes: 1000 , 1500 , and 2500 . For each sample size, we independently simulate a total of $100,000$ data sets using the data-generating process specified here. We fit the naive two-step, one-step, and two-step estimators to each simulated data set and then calculate absolute median deviation and root mean squared error (RMSE) across these simulated data sets. We use the absolute median deviation rather than bias as a measure of systematic error for each estimator because the naive estimator is extremely variable and its bias cannot be estimated reliably in a small sample size even over $100,000$ simulated data sets.

Figure 1 summarizes the results of this simulation study. Each row represents a parameter of the outcome model. Note that RMSE is plotted using $\log(\text{RMSE} + 0.01)$ for a large variance of the naive two-step estimator though the axis labels are on the original scale. The dashed lines display the results from the two-step estimator, whereas the solid lines represent those of the one-step estimator. The dash-dotted lines represent the naive two-step estimator, which is consistently the worst performer, in terms of both absolute median deviation and RMSE. Although, as expected, both of our proposed estimators recover the truth as the sample size increases, the one-step estimator outperforms the two-step estimator in small sample sizes on both metrics. What is

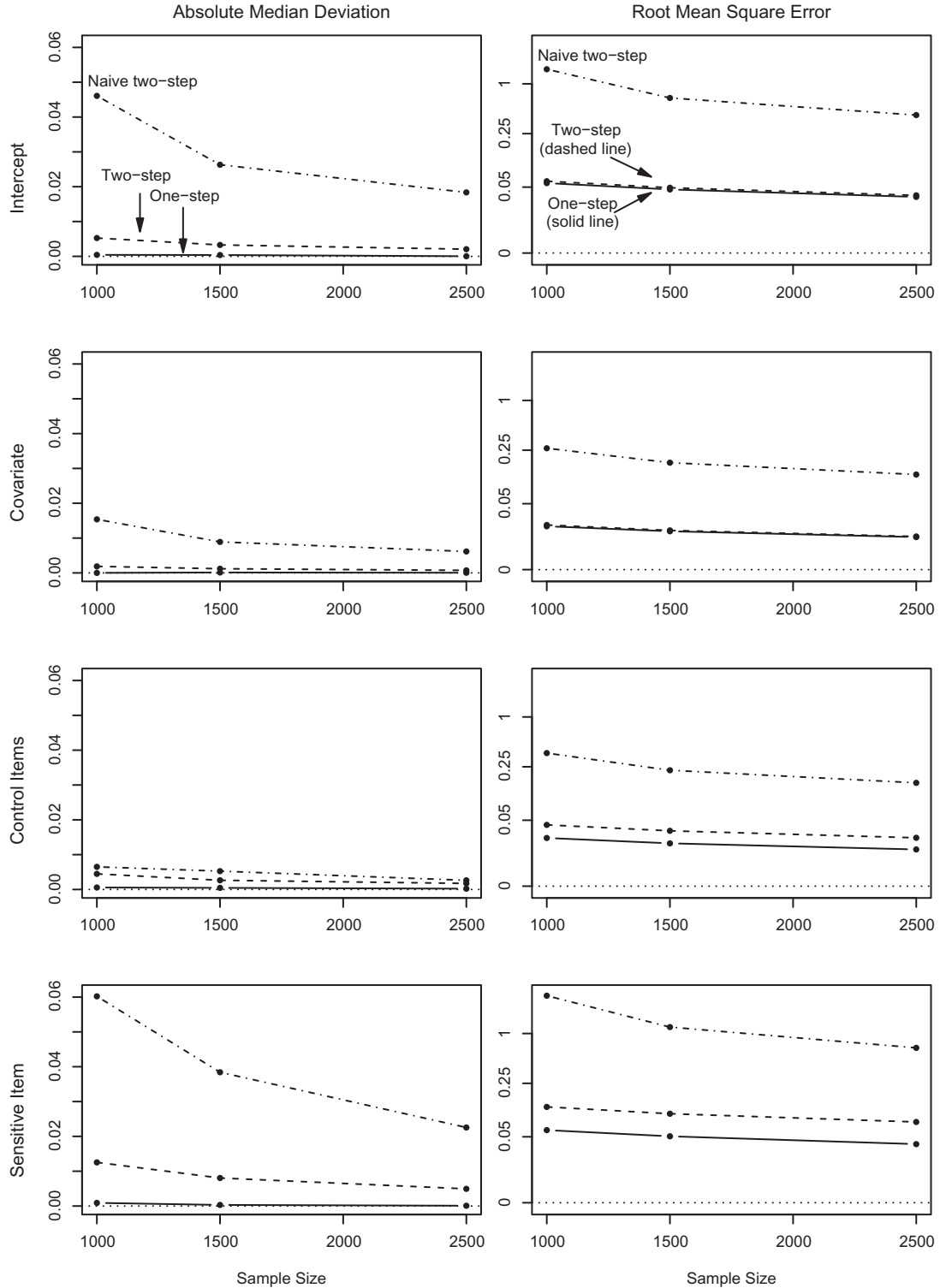


Fig. 1 A simulation study comparing the performance of the one-step estimator (solid lines), the two-step estimator (dashed lines), and the naive two-step estimator (dash-dotted lines). The sample size varies from 1000, 1500, to 2500. Although both proposed estimators significantly outperform the naive two-step estimator, the one-step estimator outperforms the proposed two-step estimator in terms of absolute median deviation and RMSE especially for the coefficient of the sensitive item. The RMSE is transformed before plotting; the plot shows $\log(\text{RMSE} + 0.01)$, but the y-axis labels are on the original scale.

interesting, however, is that the performance difference between the two estimators is the greatest for the coefficient of the sensitive item, which is the main quantity of interest here. When the sample size is 1000 (which is a typical sample size in social science research), we observe that the median deviation and RMSE of the one-step estimator are less than half of those of the two-step estimator. This clearly demonstrates the relative advantages of the one-step estimator, which is not only more flexible but also less biased and more efficient.

5 Empirical Analysis

In this section, we apply the proposed one-step estimator to the Mexico 2012 Panel Study described in Section 2. We assess the association between vote-selling and two core outcomes in the clientelism literature: turnout and approval for the winning presidential candidate, Enrique Peña Nieto.

5.1 *Effects of Vote-Selling on Turnout*

5.1.1 Data and model

We begin by introducing our measure of turnout. As an alternative to self-reported turnout that often generates an overestimate (e.g., Burden 2000), enumerators in the 2012 Mexico Panel Study asked respondents to show their voter registration card. Voters who showed cards that were marked appropriately by poll workers were scored as having voted, whereas those whose cards were unmarked were scored as nonvoters. Some respondents did not present cards when asked; we code these as nonvoters. Whereas self-reported turnout was 84.9% in the sample, the alternative approach measured turnout at 60.2%, coming much closer to the official turnout of 64.5%.

In probing the plausibility of our new methodology, we provide a basic model of turnout that includes the variables most often suggested by researchers in the field. We include a set of standard demographic variables: wealth, age, gender, and education (Wolfinger and Rosenstone 1980; Blais 2000). Citizens that are more engaged in politics are generally more likely to vote, so we also include a measure of political interest as well as an enumerator-scored variable capturing whether or not the respondent displayed political propaganda on the outside of their home.² By the same token, citizens that are disillusioned with politics are more likely to stay home, so we include a binary measure of whether respondents believed the elections were clean or not (Domínguez and McCann 1995). All of these covariates are also used for the model that predicts vote-selling.³

In addition to these standard variables, we include measures of strong partisanship for the three main parties to operationalize claims in the vote-buying literature, described in Section 2.1. Although such operationalization is a nontrivial task, this measure is one way to assess the argument of Nichter (2008) that bosses seeking higher turnout target selective benefits to supporters. In this case, the inclusion of strong partisanship allows us to explore the argument of Cox and McCubbins (1986) that core supporters respond to selective benefits more than other potential voters.⁴

5.1.2 Empirical findings

The estimated coefficients from the model described above are presented in Supplementary Appendix A.2. Here, we focus on the presentation of substantive results by calculating the average predicted probabilities from the fitted model.

²We note that respondents who display such propaganda may be displaying their partisanship (see Corstange 2012a) or as a way to seek patronage (Nichter and Palmer-Rubin 2015).

³Unlike the candidate approval model discussed in Section 5.2, we do not include regional indicator variables, as their inclusion appears to worsen the model fit when the outcome variable is turnout.

⁴The question used to measure partisanship is a typical one used in Mexican polls and surveys: “Generally, do you consider yourself to be a PANista, PRIista or PRDista? Do you think of yourself as a strong or weak [PANista/PRIista/PRDista]?” Respondents who identified strongly with any of the three major parties were coded as party supporters. All parties may engage in vote-buying, but the minor parties are less likely to be able to do so effectively.

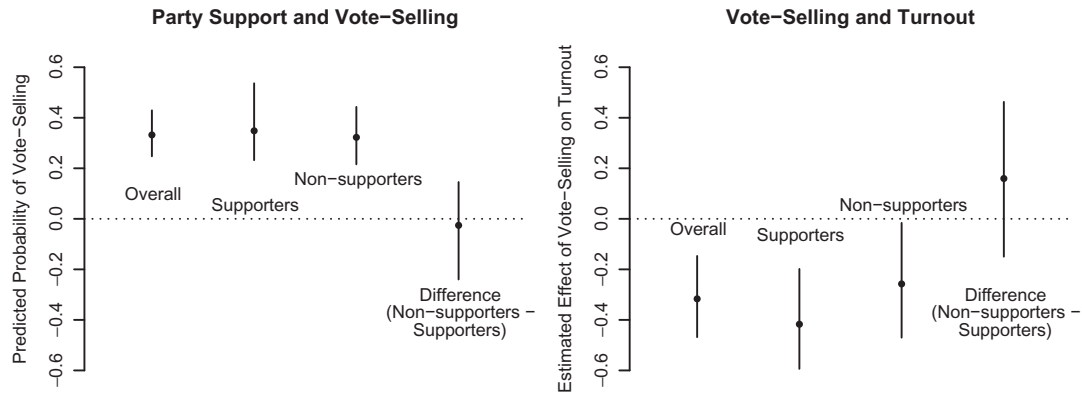


Fig. 2 The predicted probability of vote-selling and the estimated effect of vote-selling on turnout from the list experiment. The left panel displays the predicted probabilities of selling one's vote with 95% confidence intervals for the overall population, and for supporters and nonsupporters of the three main parties. There is no statistically significant difference between party supporters and nonparty supporters. The right panel displays the estimated effect of vote-selling on turnout with 95% confidence intervals for the same three groups of voters. Among supporters and nonsupporters, those who sold their vote are less likely to turn out to vote. Again, there is no statistically significant difference between party supporters and nonsupporters.

The left panel of Fig. 2 shows the predicted probability of vote-selling. We fit two models separately: the first to a subsample of strong partisans and the second to the remaining sample. Our models predict that about 33.2% of respondents (with a 95% confidence interval of [24.7, 42.9]) sold their votes and that there is no statistically significant difference between strong partisans and others. Although there are many ways to measure the concept of a “core” voter, our findings indicate that Mexico's parties probably do not target selective benefits to strong partisans more than others, a finding that is inconsistent with a core-voter model such as Cox and McCubbins (1986) or Nichter's assertion that parties should buy the turnout of their loyal supporters. However, it is consistent with other accounts that emphasize parties' targeting of various constituencies, either due to error (Stokes et al. 2013) or as part of a strategy (Magaloni, Diaz-Cayeros, and Estévez 2007; Gans-Morse, Mazzuca, and Nichter 2014).

The right panel of Fig. 2 presents the results from the outcome model regression on turnout. Vote-selling appears to decrease the probability of turnout by about 31.7 percentage points, on average, regardless of partisanship (with the 95% confidence interval of [19.8, 59.3] and [1.6, 47.0] for strong partisans and others, respectively). Among both subsamples of party supporters and nonsupporters, we find a relatively strong negative correlation between vote-selling and turnout.

We caution that due to their small sample sizes the analyses of subsamples are less robust. For example, the model fitted to the subsample of party supporters produces large standard errors for some parameters in the sensitive-item model. This could be an indication for poor model fit given a smaller sample size. Dropping these variables (i.e., urban and propaganda indicators) does not substantively change the results, reducing the estimated probability of vote-selling somewhat from 34.8% to 24.9% (95% confidence interval of [11.4%, 46.4%]). However, the estimated effect of vote-selling on turnout among party supporters under this alternative model specification is reduced to -5.5 percentage points with a 95% confidence interval of [-38.6%, 20.5%].

At first glance, these findings would seem to undermine the contention that vote-buying enhances turnout, and instead support the much more disturbing idea that bosses buy abstention (Cox and Kousser 1981). Though possible, field reports do not support the claim of a massive abstention-buying scheme in Mexico. Rather, if bosses direct benefits to voters that are the least likely to turn out, and their efforts are not always successful, then vote-sellers may still be less likely to turn out, on average, than nonsellers. This discrepancy can exist even if payoffs do increase the probability of turnout.

We note that straightforward behavioral findings like these have not been discussed in the literature. Instead, existing studies have mainly focused on broad theoretical relationships between vote-buying and turnout. Tools such as ours have a potential to move literatures forward not only by assessing empirical relationships but by uncovering logical yet untheorized

ones. As Fig. 3 demonstrates, only by using the list experiment responses as a predictor—rather than the responses to the direct question—can we detect an effect of vote-selling on turnout.

5.2 *Effects of Vote-Selling on Candidate Approval*

5.2.1 Data and model

As discussed in Section 2.1, vote-selling may also be associated with candidate approval. In Mexico's 2012 election, Peña Nieto's campaign was accused of trading selective benefits for political support. Such payoffs could be associated with higher approval ratings of the winning candidate if voters responded to the benefits. It is also possible that he targeted his party's supporters, who naturally hold him in higher esteem. However, although we find an association between vote-selling and approval of Peña Nieto, we find no evidence that party-specific support is strongly associated with vote-selling.

A limitation of the list experiment in the Mexico 2012 Panel Study is that it does not identify which party or candidate bought the vote of the respondent. Mexico's 2012 elections featured simultaneous contests for president, Congress, and a host of state and local offices. As a result, party brokers could have offered to buy citizens' votes in exchange for their support of the entire political party, not just the presidential candidate. In addition, respondents could have had down-ballot candidates in mind when responding to the list experiment. Nevertheless, presidential coattails, the national campaigns' control over advertising, and a dearth of split-ticket voting lead us to conjecture that voters considered the presidential elections when responding to the list experiment. In the follow-up to the direct vote-buying question, asking respondents which party attempted to buy their support, many respondents said they received a payoff from a party or presidential campaign; none mentioned down-ballot candidates.

In order to explore the relationship between vote-selling and candidate approval, we fit a linear outcome regression model predicting the respondents' opinion of Peña Nieto, measured on a scale from 1 to 11, where 1 represents a least favorable opinion and 11 indicates a most favorable opinion. This model includes the same set of variables as the turnout model, with a few additions. First, we control for residence in the country's four main geographic regions (North, Central, South, and the Metropolitan area of Mexico City), because political support and vote-buying strategies may vary by geographic region in Mexico. We also adjust for identification with the three main political parties, as separate indicators this time. This allows us to see whether those identifying with Peña Nieto's party are more likely to report selling their vote, and to adjust for partisan identification when predicting opinions about the candidate.

5.2.2 Empirical findings

We fit the model described above using the responses to the direct question as well as the predicted responses from the list experiment. The estimated coefficients of these models are presented in Supplementary Appendix A.3. Interestingly, the results from these two models are substantively different. The left panel of Fig. 4 presents the estimates based on the direct question. In this plot, we observe that the estimated difference between vote-sellers and nonsellers is approximately -0.82 (with the 95% confidence interval of $[-1.50, -0.16]$), suggesting that vote-selling may decrease one's approval rating for Peña Nieto. However, the conclusion based on the list experiment is the exact opposite. In the right panel of the same figure, we see that although the estimated support remains at a similar level, the difference between vote-sellers and nonsellers is estimated to be positive and around 0.71 points (with the 95% confidence interval of $[-0.04, 1.47]$). That is, as one would expect, vote-sellers appear to assess Peña Nieto somewhat more highly.

The different (and opposing) estimates generated by the two models indicate the real costs of incorrectly measuring vote-selling. If we were to take the direct question model at face value, we would conclude that vote-selling does not work as bosses intend and in fact backfires, contradicting the theoretical and empirical claims made by the multiple authors we reviewed in Section 2. We do not necessarily believe this conclusion is warranted. Rather, measurement error in the direct

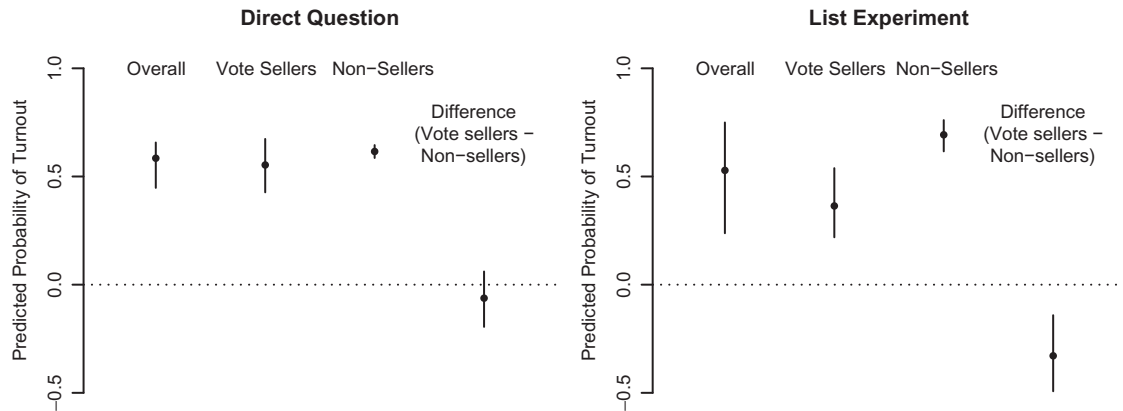


Fig. 3 Predicted turnout and the estimated effect of vote-selling on turnout from the direct question and the list experiment. The figure displays the predicted turnout, among those who reported they had sold their vote and those who did not, as measured by the direct question (left panel) and the list experiment (right panel). The vertical bars represent 95% confidence intervals. According to the direct question, vote-sellers are not measurably less likely to turn out than nonvote-sellers. However, the list experiment yields a different conclusion: vote-selling has a negative impact of around 31.7 percentage points on the predicted probability of turnout.

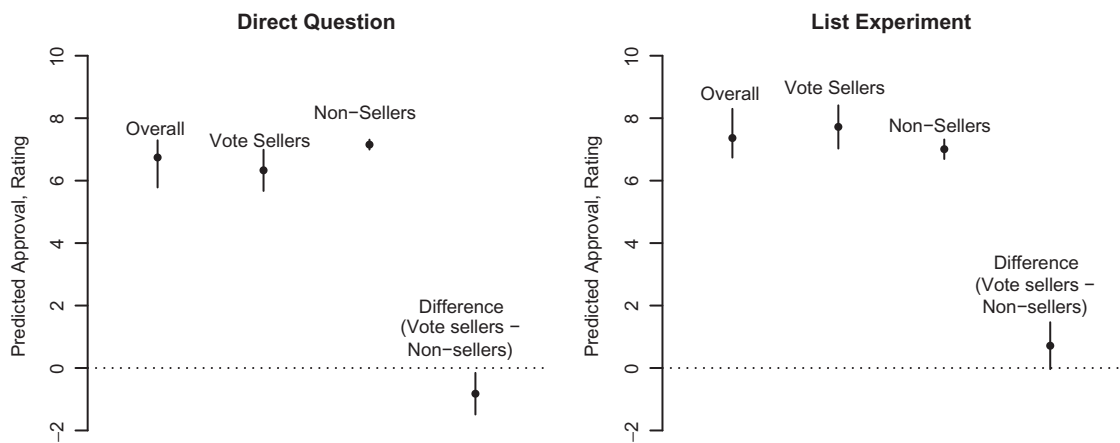


Fig. 4 The predicted approval rating of Enrique Peña Nieto (PRI) and the estimated effect of vote-selling on the rating from the direct question and the list experiment. The figure displays the predicted approval rating, which ranges from 1 (least supportive) to 11 (most supportive), of Enrique Peña Nieto, the winning presidential candidate (of the PRI party), among those who reported they had sold their vote and those who did not, as measured by the direct question (left panel) and the list experiment (right panel). The vertical bars represent 95% confidence intervals. According to the direct question, vote-sellers rate the candidate around 1 point lower than those who did not sell their vote. However, the list experiment yields the opposite conclusion: vote-selling has a positive impact of 0.71 point on the candidate's approval.

question is the likely culprit. A logistic model of response to the direct question demonstrates that Peña Nieto supporters may have been less likely to admit that they received payoffs (the coefficient on PRI support is negative and significant at the 0.1 level), whereas estimates from the sensitive-item model from the list experiment show that his supporters were just as likely to sell their votes as others (see Supplementary Appendix A.3 for full tables of coefficients). We conclude that Peña Nieto supporters declined to declare their activities openly; the legal battle over vote-buying that continued after the election may have discouraged the winner's supporters from admitting that they received payoffs.

The list experiment not only provides a better measurement of vote-selling activity but also yields the more intuitive result that vote-selling is associated with higher approval of Peña Nieto. His party's history of vote-buying (Cornelius and Craig 1991; Greene 2007) and other evidence from field research during the 2012 general election campaigns suggest that the Peña Nieto campaign attempted to buy more votes than his rivals. The association between vote-selling and approval of Peña Nieto could indicate either that his campaign targeted payoffs to his supporters or that vote-selling worked as intended. The targeting hypothesis is less plausible given that PRI partisanship is not associated with vote-selling in the sensitive-item model. It is more likely that vote-selling increased evaluations of Peña Nieto, a finding consistent with Stokes et al. (2013) and that we investigate further in other work (see Greene 2014).

6 Concluding Remarks

The list experiment has recently gained popularity among social scientists as a survey methodology to elicit truthful response for sensitive questions. Although statistical methods have been proposed to model responses to sensitive items as a dependent variable, no existing method has enabled the use of the predicted responses from list experiments as the explanatory variables in an outcome regression model. And yet, researchers may wish to ascertain the effects of sensitive behavior and attitudes on other actions and opinions. We address this methodological gap by proposing a new estimation technique. Our one-step estimator can be applied to a wide range of outcome regression models, thereby providing a flexible tool for survey researchers.

The methodological approach proposed in this article can be extended to other settings. For example, one can generalize this method to other designs of list experiments such as double list experiments (Droitcour et al. 1991; Glynn 2013) or other indirect questioning methods such as endorsement experiments (Bullock, Imai, and Shapiro 2011) and randomized response methods (Warner 1965; Blair, Imai, and Zhou 2014). These and other related research projects are currently being conducted by the authors and other researchers.

Finally, as the use of list experiments becomes popular, it is important for applied researchers to take seriously the required assumptions. Statistical tests for detecting possible violations of such assumptions and the modeling strategies for addressing them have been developed (Blair and Imai 2012). A more effective strategy, however, is to incorporate some type of empirical validation within one's research design. For example, researchers can use different indirect questioning techniques, such as endorsement experiments and randomized response methods, and examine whether they provide similar estimates (Blair, Imai, and Lyall 2014). Another strategy is to aggregate individual-level list experiment results and validate them against available true information (Rosenfeld, Imai, and Shapiro 2014). We believe that like any survey research, an extensive pilot study and the use of different data sources are essential for making empirical analyses credible.

Used carefully, list experiments and other indirect questioning techniques can be powerful tools for understanding sensitive behaviors and attitudes in a variety of disciplines. The proposed method increases the utility of the list experiment by allowing researchers to use sensitive information in ways that were previously impossible.

References

- Ahlquist, J., K. Mayer, and S. Jackman. 2013. Alien abduction and voter impersonation in the 2012 US general election: Evidence from a survey list experiment. Technical Report, University of Wisconsin, Madison.
- Biemer, P., and G. Brown. 2005. Model-based estimation of drug-use prevalence using item count data. *Journal of Official Statistics* 21(2):287–308.
- Blair, G., and K. Imai. 2012. Statistical analysis of list experiments. *Political Analysis* 20(1):47–77.
- Blair, G., K. Imai, and J. Lyall. 2014. Comparing and combining list and endorsement experiments: Evidence from Afghanistan. *American Journal of Political Science* 58(4):1043–63.
- Blair, G., K. Imai, and B. Park. 2014. List: Statistical methods for the item count technique and list experiment. Available at the Comprehensive R Archive Network (CRAN). <http://CRAN.R-project.org/package=list> (accessed June 1, 2014).

- Blair, G., K. Imai, and Y.-Y. Zhou. 2014. Design and analysis of randomized response technique. Technical report, Pittsburgh, PA: Princeton University.
- Blais, A. 2000. *To vote or not to vote: The merits and limits of rational choice theory*. Pittsburgh, PA: University of Pittsburgh Press.
- Brusco, V., M. Nazareno, and S. Stokes. 2004. Vote buying in Argentina. *Latin American Research Review* 39(2):66–88.
- Bullock, W., K. Imai, and J. N. Shapiro. 2011. Statistical analysis of endorsement experiments: Measuring support for militant groups in Pakistan. *Political Analysis* 19(4):363–84.
- Burden, B. C. 2000. Voter turnout and the National Election Studies. *Political Analysis* 8(4):389–98.
- Calvo, E., and M. Murillo. 2004. Who delivers? Partisan clients in the Argentine electoral market. *American Journal of Political Science* 48(4):742–57.
- Carreras, M., and Y. İrepoğlu. 2013. Electoral studies. *Electoral Studies* 32(4):609–19.
- Cornelius, W. 2004. Mobilized voting in the 2000 elections: The changing efficacy of vote buying and coercion in Mexican electoral politics. In *Mexico's Pivotal Democratic Election*, eds. J. I. Domínguez and C. Lawson, 47–65. Stanford, CA: Stanford University Press.
- Cornelius, W., and A. Craig. 1991. *The Mexican political system in transition*. Monograph Series. La Jolla, CA: Center for U.S.-Mexican Studies.
- Corstange, D. 2009. Sensitive questions, truthful answers? Modeling the list experiment with LISTIT. *Political Analysis* 17(1):45–63.
- . 2012a. Religion, pluralism, and iconography in the public sphere: Theory and evidence from Lebanon. *World Politics* 64(1):116–60.
- . 2012b. Vote-trafficking in Lebanon. *International Journal of Middle East Studies* 44:483–505.
- Cox, G. W., and J. M. Kousser. 1981. Turnout and rural corruption: New York as a test case. *American Journal of Political Science* 25(4):646–63.
- Cox, G. W., and M. D. McCubbins. 1986. Electoral politics as a redistributive game. *Journal of Politics* 48(2):370–89.
- Díaz-Cayeros, A., F. Estévez, and B. Magaloni. 2009. Welfare benefits, canvassing, and campaign handouts. In *Consolidating Mexico's democracy: The 2006 presidential campaign in comparative perspective*, eds. J. Domínguez, C. Lawson, and A. Moreno, 229–45. Baltimore, MD: Johns Hopkins University Press.
- Dixit, A., and J. Londregan. 1996. The determinants of success of special interests in redistributive politics. *Journal of Politics* 58:1132–55.
- Domínguez, J., and J. McCann. 1995. Shaping Mexico's electoral arena: The construction of partisan cleavages in the 1988 and 1991 national elections. *American Political Science Review* 89(1):34–48.
- Droitcour, J., R. A. Caspar, M. L. Hubbard, and T. M. Ezzati. 1991. The item count technique as a method of indirect questioning: A review of its development and a case study application. In *Measurement errors in surveys*, eds. P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman, 185–210. New York: John Wiley & Sons.
- Gans-Morse, J., S. Mazza, and S. Nichter. 2014. Varieties of clientelism: Machine politics during elections. *American Journal of Political Science* 58(2):415–32.
- Glynn, A. N. 2013. What can we learn with statistical truth serum? Design and analysis of the list experiment. *Public Opinion Quarterly* 77:159–72.
- Gonzalez-Ocantos, E., C. K. de Jonge, C. Meléndez, J. Osorio, and D. W. Nickerson. 2012. Vote buying and social desirability bias: Experimental evidence from Nicaragua. *American Journal of Political Science* 56(1):202–17.
- Greene, K. F. 2007. *Why dominant parties lose: Mexico's democratization in comparative perspective*. New York: Cambridge University Press.
- . 2014. Back from the dead: Vote-selling zombies and the return of Mexico's PRI. Paper presented at the Latin American Studies Association Congress, Chicago.
- Greene, K. F., J. Domínguez, C. Lawson, and A. Moreno. 2012. *The Mexico 2012 Panel Study*. Wave 2. Original public opinion survey, available at <http://kfgreene.webhost.utexas.edu>.
- Holbrook, A. L., and J. A. Krosnick. 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. *Journal of the American Statistical Association* 106(494):407–16.
- Imai, K., B. Park, and K. F. Greene. 2014. Replication data for: Using the predicted responses from list experiments as explanatory variables in regression models. The Dataverse Network. <http://dx.doi.org/10.7910/DVN/27083> (accessed June 1, 2014).
- Kane, J. G., S. C. Craig, and K. D. Wald. 2004. Religion and presidential politics in Florida: A list experiment. *Social Science Quarterly* 85(2):281–93.
- Kramon, E. 2009. Vote-buying and political behavior: Estimating and explaining vote-buying's effect on turnout in Kenya. *Afrobarometer Working Papers* 1–31.
- Kuklinski, J. H., M. D. Cobb, and M. Gilens. 1997. Racial attitudes and the “New South.” *Journal of Politics* 59(2):323–49.
- Lindbeck, A., and J. W. Weibull. 1987. Balanced-budget redistribution as the outcome of political competition. *Public Choice* 52(3):273–97.
- Magaloni, B., A. Díaz-Cayeros, and F. Estévez. 2007. Clientelism and portfolio diversification. *Patrons, Clients, and Policies: Patterns of Democratic Accountability and Political Competition*, 182–204. New York, NY: Cambridge University Press.

- Nichter, S. 2008. Vote buying or turnout buying? Machine politics and the secret ballot. *American Political Science Review* 102(1):19–31.
- Nichter, S., and B. Palmer-Rubin. 2015. Clientelism, declared support, and Mexico's 2012 campaign. In *Mexico's evolving democracy: A comparative study of the 2012 elections*, eds. J. Domínguez, K. Greene, C. Lawson, and A. Moreno, 220–26. Baltimore, MD: Johns Hopkins University Press.
- Penfold Becerra, M. 2007. Clientelism and social funds: Evidence from Chávez's Misiones. *Latin American Politics and Society* 49(4):63–84.
- Rosenfeld, B., K. Imai, and J. Shapiro. 2014. An empirical validation study of popular survey methodologies for sensitive questions. Working paper available at <http://imai.princeton.edu/research/Validate.html> (accessed June 1, 2014).
- 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., T. Dunning, M. Nazareno, and V. Brusco. 2013. *Brokers, voters, and clientelism: The puzzle of distributive politics*. Cambridge Studies in Comparative Politics. Cambridge, UK: Cambridge University Press.
- Streb, M. J., B. Burrell, B. Frederick, and M. A. Genovese. 2008. Social desirability effects and support for a female American president. *Public Opinion Quarterly* 72(1):76–89.
- Tourangeau, R., and T. Yan. 2007. Sensitive questions in surveys. *Psychological Bulletin* 133(5):859–83.
- Warner, S. L. 1965. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association* 60(309):63–69.
- Wimbush, J. C., and D. R. Dalton. 1997. Base rate for employee theft: Convergence of multiple methods. *Journal of Applied Psychology* 82(5):756–63.
- Wolfinger, R., and S. Rosenstone. 1980. *Who votes?* New Haven, CT: Yale University Press.