

## Usando as respostas previsíveis da abordagem *list-experiments* como variáveis explicativas em modelos de regressão

*Using the predicted responses from List experiments as explanatory variables in regression models*

---

Kosuke Imai  
Bethany Park  
Kenneth F. Greene

---

### **Resumo**

O modelo de lista de experimento (*List Experiment*), também conhecido como o modelo de contagem de itens (*the Item Count Technique*), vem se tornando bastante popular como uma metodologia para se obter respostas confiáveis para questões complexas e sensitivas. Recentemente, técnicas de múltiplas variáveis têm sido desenvolvidas para prever respostas, não observadas, para questões complexas usando a abordagem características dos entrevistados (*respondent characteristics*). Não obstante, não existe nenhum método para o uso dessas respostas, como uma variável explicativa em outro modelo de regressão. Tentamos resolver esta lacuna primeiro, melhorando a performance de um *naive two-step estimator*. Apesar de sua simplicidade, este aperfeiçoado método de duas etapas pode apenas ser aplicado em modelos lineares e é ineficiente estatisticamente. Nós, portanto, desenvolvemos um método para estimativas de máxima verossimilhança, de completa eficiência e aplicável a uma gama muito grande de modelos. Usamos um estudo simulado para avaliar o desempenho empírico dos métodos propostos. Nós também os aplicamos ao México 2012 Estudo de Panel para analisar se a compra de votos está associada com o aumento do comparecimento eleitoral e o nível de aprovação do candidato. Os métodos propostos são implementados in open-source software.

### **Palavras-chave**

Lista de Experimento; Compra de Votos; México 2012.

### **Abstract**

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.

### **Keywords**

List Experiment; Vote Buying; México 2012.

## Introdução<sup>1</sup>

Nas Ciências Sociais, o modelo de lista de experimento (*List Experiment*), também conhecido como modelo de contagem de itens sem emparelhamento (*the Item Count Technique* ou *Unmatched Count Technique*), vem se tornando bastante popular como um método de questionamento indireto para obter respostas confiáveis para questões complexas e sensíveis. O modelo *List Experiment* tem sido usado em diversas disciplinas para medir uma grande variedade de comportamentos como roubo por funcionários e uso de drogas (WIMBUSH e DALTON, 1997; BIEMER e BROWN, 2005). Em Ciência Política, *List Experiment* tem sido usado para medir atitudes socialmente não desejáveis, tais como preconceito envolvendo raça, religião e gênero (KUKLINSKI, COBB e GILENS, 1997; KANE, CRAIG e WALD, 2004; STREB *et al.*, 2008), e condutas ilícitas como fraude eleitoral e compra de votos (CORSTANGE, 2012; GONZALEZ-OCANTOS *et al.*, 2012; AHLQUIST, MAYER e JACKMAN, 2013). *List Experiment* tem também sido usado para estudar matérias aparentemente menos complexas como comparecimento eleitoral e atitudes relacionadas ao direito de voto (CORSTANGE, 2009; HOLBROOK e KROSNICK, 2010). Quando aplicado corretamente, o método pode melhorar a validade da medição, reduzindo o viés de desajustabilidade social que pode ocorrer quando alguém responde diretamente a um questionário com questões sensíveis (TOURANGEAU e YAN, 2007).

Nos anos recentes, pesquisadores têm conseguido, também, diversos avanços metodológicos que ampliaram o campo de aplicabilidade do modelo *List Experiment*. Em primeiro lugar, modelos de regressão multivariada foram desenvolvidos para ajudar pesquisadores a identificar as características dos entrevistados, as quais estão associadas a certas respostas a perguntas sensíveis (CORSTANGE, 2009; IMAI, 2011; BLAIR e IMAI, 2012). Segundo, métodos estatísticos têm sido desenvolvidos para testar e verificar se os principais pressupostos do modelo *List Experiment* foram violados, e para reduzir o potencial de distorção ocasionado pela violação de tais pressupostos (BLAIR e IMAI, 2012). Terceiro, recomendações metodológicas têm sido feitas para aperfeiçoar o design e capacidade de validação do *List Experiment* (BLAIR e IMAI, 2012; GLYNN, 2013; BLAIR, IMAI, e LYALL, 2014).

---

<sup>1</sup> Os métodos propostos são implementados via software livre: Os métodos estatísticos para o *Item Technique* e *List Experiment*, o qual está disponível para download no Comprehensive R Archive Network (<<http://cran.r-project.org/package=list>>). Material suplementar para este artigo está disponível no *Political Analysis* Web site. O arquivo para cópia está disponível como Imai, Park, and Greene (2014). Nos agradecemos a Adam Glynn e revisores anônimos pelos seus úteis comentários.

Apesar destes avanços, não existe nenhum método atualmente para usar as respostas para itens sensitivos como uma variável explicativa, em outro modelo de regressão. E ainda, existem muitos exemplos nos quais pesquisadores desejam ter respostas para questões sensitivas para prever comportamentos e opiniões. Nestes casos, a principal característica do *List Experiment* se torna um obstáculo importante. Muito embora o anonimato concedido ao entrevistado melhore a qualidade da validação, os pesquisadores não observam diretamente as suas respostas para questões sensitivas e, desta forma, não podem usar tais respostas para fazer previsões em modelos de regressão. Por exemplo, Stokes (2005) e Stokes *et al.* (2013) argumentam que a compra de votos afeta os índices de aprovação do candidato, ao passo que Nichter (2008) diz que isto, na verdade, aumenta o comparecimento eleitoral. Outro trabalho mostra que o modelo *List Experiment* parece aprimorar a medição da compra de votos (GONZALEZ-OCANTOS *et al.*, 2012). Não obstante, nenhum método existente permite aos pesquisadores usarem esta informação sobre a compra de votos em um modelo de regressão para prever índices de aprovação de candidato ou comparecimento eleitoral.

Nós tentamos preencher esta lacuna na literatura sobre metodologia primeiramente aperfeiçoando a *performance* do ingênuo método de duas etapas para fazer previsões, o qual, nós mostramos ser potencialmente tendencioso e estatisticamente ineficiente para ser útil na prática. O método renovado para fazer previsões que nós estamos propondo é relativamente fácil de ser implementado e é computacionalmente estável. No entanto, este método não é completamente eficiente e, mais importante, ele somente pode ser aplicado em modelos lineares. Para superar estas deficiências nós propomos um método com a máxima verossimilhança (ML: *Maximum Likelihood*, sigla em inglês) que é estatisticamente eficiente e é aplicável em uma variada gama de modelos. Ambas as nossas abordagens partem do modelo geral apresentado por Imai (2011) e Blair e Imai (2012). Os métodos propostos podem ser implementados pela lista de *softwares* livres: Métodos Estatísticos para *Item Count Technique* e *List Experiments*, os quais oferecem um conjunto abrangente de métodos estatísticos para análise de experimentos (BLAIR, IMAI e PARK, 2014).

Nós conduzimos uma simulação de estudo para avaliar a *performance* empírica dos métodos propostos. Nós achamos que, especialmente com amostras pequenas, o modelo de uma etapa, embora computacionalmente menos estável do que o aperfeiçoado modelo de duas etapas, apresenta menos distorções e é bem mais eficiente do que o modelo de duas etapas. Este aperfeiçoamento nas propriedades estatísticas é particularmente notável no que diz respeito a itens sensitivos. Ambos os

modelos superam o primeiro, e ultrapassado, método de duas etapas em termos de distorção de resultados e eficiência.

Finalmente, nós aplicamos os métodos propostos no *México 2012 Panel Study*, focando principalmente na associação entre compra de votos, comparecimento eleitoral e aprovação do candidato. Esta aplicação motivou o desenvolvimento da nossa nova metodologia. Nós primeiramente mostramos que a questão direta leva a uma importante porcentagem subestimada de compra de votos (aproximadamente 6%), quando comparado com o resultado do modelo *List Experiment* (aproximadamente 20%). Estas diferenças em medição refletem o relacionamento entre a compra de votos e a aprovação do candidato. Ao passo que um modelo padrão de regressão usando respostas para perguntas diretas mostra que a compra de votos está associada com baixos níveis de aprovação do candidato, nossa metodologia baseada no *List Experiment* sugere que os que vendem votos têm maior taxa de aprovação. Nós também demonstramos que a compra de votos está associada com baixo comparecimento eleitoral, mesmo entre os apoiadores dos partidos os quais a literatura identifica como os alvos preferenciais dos que compram votos. Este resultado sugere que os partidos estão mirando exatamente os votantes que mais provavelmente não irão votar. Esta análise ilustra que a nossa metodologia permite aos analistas avaliar hipóteses sobre os efeitos de atitudes sensitivas ou comportamentos que são bem capturados pelo método *List Experiment*.

### **Aplicação Empírica: The México 2012 Panel Study**

Nesta seção, nós descrevemos a nossa instigante aplicação empírica: a associação entre compra de votos e ambos, comparecimento eleitoral e aprovação do candidato, nas eleições gerais no México em 2012. Depois de exaustivamente discutir importantes argumentos que direcionaram a nossa subsequente análise empírica, nós introduzimos o instrumento *survey* e conduzimos uma análise descritiva inicial.

### **Compra de Votos, Comparecimento Eleitoral, e Aprovação do Candidato**

Como o interesse pelo clientelismo tem crescido no mundo acadêmico, em especial em Ciência Política, uma primeira tarefa tem sido entender qual a eficácia da compra de votos como estratégia de campanha em eleições ao redor do mundo (BRUSCO, NAZARENO e STOKES, 2004; CALVO e MURILLO, 2004; STOKES, 2005; PENFOLD-BECERRA, 2007; MAGALONI, DIAZ-CAYEROS e ESTÉVEZ, 2007; STOKES *et al.*, 2013). No coração deste questionamento está a preocupação de como a compra de votos afeta ambos, o comparecimento do eleitor

para votar, e a aprovação do candidato. No entanto, dados apropriados ainda são escassos e as associações propostas tem sido avaliadas, principalmente, com relações bivariadas.

O mais recente debate na literatura sobre a compra de votos tem focado sua atenção em quais eleitores recebem o seletivo benefício. Trabalhos teóricos estabelecem que a estratégia deveria focar nos eleitores indecisos (LINDBECK e WEIBULL, 1987; STOKES, 2005). Consistentes com esta afirmação, Stokes *et al.* (2013) mostram que os eleitores que receberam os seletivos benefícios na Argentina, México, Venezuela e Índia fazem uma avaliação mais positiva dos partidos que forneceram os benefícios. No entanto, eles também acham que os protagonistas destas ações concentram demasiadamente seu foco em grupos de eleitores os quais já possuem uma opinião positiva sobre o partido, indicando que a compra de votos pode funcionar apenas com pequenas fatias do eleitorado. Em contraste, Nichter (2008) argumenta que partidos usam os seletivos benefícios com o objetivo de aumentar o comparecimento eleitoral entre os grupos em que já é conhecida a baixa capacidade de mobilização. Tais grupos podem ser também os mais prováveis a responder a esses benefícios (COX e MCCUBBINS, 1986)<sup>2</sup>. Desta forma, a atual literatura concorda amplamente que os eleitores mais leais recebem a maior generosidade, mas ela discorda sobre se os benefícios seletivos ajudam mais a aumentar a aprovação entre os eleitores indecisos ou aumenta o comparecimento eleitoral entre os que já são leais.

Muito embora estas principais afirmações sejam teoricamente convincentes, trabalhos empíricos têm sido dificultados por dois problemas, os quais nossa metodologia tenta resolver. Primeiro, a maioria dos trabalhos confia em *surveys* que perguntam diretamente se o entrevistado vendeu o seu voto (CORNELIUS, 2004; GREENE, 2007; DIAZ-CAYEROS, ESTÉVEZ e MAGALONI, 2009; KRAMON, 2009; CARRERAS e IREPOGLU, 2013). A ilegalidade deste tipo de ato em muitos países, incluindo México, faz com que os entrevistados fiquem temerosos de admitir que tenham vendido os seus votos. *List Experiment* oferece uma solução para este problema, permitindo que os entrevistados passem informações verdadeiras sem admitir o comportamento ilícito. No seu estudo sobre a compra de votos na Nicarágua, Gonzalez-Ocantos *et al.* (2012) mostra que apenas 2% dos entrevistados

---

<sup>2</sup> Ver Dixit e Lontregan (1991), que fornece um quadro geral o qual produz dois grupos de condições: Aquelas sob a qual os partidos favorecem seus apoiadores e aqueles sob o qual os partidos focam no voto dos indecisos.

admitiram a venda dos seus votos, quando perguntados diretamente, ao passo que quase 25% o fez questionados através do método *List Experiment*.

Além disso, mesmo que o *List Experiment* tenha sido implementado com sucesso, pesquisadores não podem simplesmente usar esta medida para prever o comparecimento eleitoral, ou intenção de voto. A razão é que nós não observamos, em cada entrevistado, se eles afirmam vender os seus votos. A principal característica do *List Experiment*, ou seja, garantir certo grau de sigilo de identidade para os entrevistados, torna impossível estimar modelos de regressão de uma forma simples. Muito embora modelos de regressão multivariadas tenham sido desenvolvidos para identificar as características do entrevistado, as quais estão associadas com certas respostas para questões sensíveis (IMAI, 2011; BLAIR e IMAI, 2012), simplesmente pegar as probabilidades previstas por estes modelos e usá-las como variável explicativa no resultado da regressão, como mostrado mais tarde, falha na recuperação do verdadeiro coeficiente e resulta em erros subestimados. Para o melhor do nosso conhecimento, nenhum método existe para usar as respostas do *List Experiment* como variáveis explicativas em modelos de regressão para prever outros resultados, como (neste caso), o comparecimento eleitoral ou as opiniões sobre candidatos. O objetivo do nosso artigo é exatamente preencher esta lacuna.

### ***The List Experiment***

Medimos a compra de votos e o comparecimento eleitoral usando *List Experiment* incorporado ao *México 2012 Panel Study*, elaborado e apresentado por Greene *et al.* (2012). Houve denúncias de compra de votos, com destaque para as eleições presidenciais daquele ano. Henrique Peña Nieto, do anteriormente dominante Partido Revolucionário Institucionalista (PRI) – um partido com uma longa história de clientelismo eleitoral (CORNELIUS e CRAIG, 1991) –, recebeu 39% dos votos, superando Andrés Manuel López Obrador de uma coalisão de esquerda liderada pelo Partido da Revolução Democrática (PRD) com 32% dos votos, e Josefina Vasquez Mota do conservador Partido da Ação Nacional (PAN), o qual já ocupava o poder, com 26% dos votos. López Obrador alegou que Peña Nieto ganhou as eleições com um massivo esquema de compra de votos (GREENE, 2014).

O *2012 Panel Study* entrevistou uma amostra nacional de cidadãos comuns, com um registro de eleitor válido, antes e depois das eleições. Todas as entrevistas foram feitas pessoalmente, ou face-a-face. Neste artigo focamos a nossa atenção nos dados coletados após as eleições, com um total de 1150 entrevistados, os quais responderam as questões do *List Experiment*. Cada entrevistado foi escolhido

aleatoriamente, tanto para o grupo de controle quanto para o grupo de tratamento. O roteiro para o grupo de controle foi lido em voz alta, como segue abaixo:

*Eu vou ler para você uma lista de atividades que consta nesta folha e eu quero que você me diga quantas destas atividades você participou nas recentes semanas. Por favor, não me diga quais, apenas quantas. As três atividades são:*

- a) *Assiste noticiários de TV que mencionam um candidato;*
- b) *Participa de algum evento de campanha;*
- c) *Fala sobre política com outras pessoas.*

Aos entrevistados escolhidos para o grupo de tratamento foi lido, em voz alta, o seguinte roteiro o qual inclui um item sensitivo adicional sobre venda de votos:

- a) *Assiste noticiários de TV que mencionam um candidato;*
- b) *Participa de algum evento de campanha;*
- c) *Troca o seu voto por algum presente, favor ou acesso a um tipo de serviço;*
- d) *Fala sobre política com outras pessoas.*

Nós enfatizamos aos entrevistados que a questão sensitiva pergunta se eles trocaram seus votos, não se eles receberam um benefício. Nós acreditamos que esta formulação pode medir melhor a compra de votos, eliminando os tipos de trocas que os entrevistados não consideram dignas do seu voto. Campanhas eleitorais rotineiramente distribuem alguns benefícios simbólicos como chaveiros, adesivos etc. Se contabilizarmos o recebimento de benefícios os quais os eleitores consideram simbólicos como venda de voto, provavelmente estaremos subestimando o efeito da compra de votos no resultado final do pleito.

A Tabela 1 resume os dados deste *List Experiment* e da questão direta. Estimar a proporção de quem vendeu seu voto usando o *List Experiment* é uma tarefa simples. Porque as duas subamostras foram aleatoriamente escolhidas, a média de atividades não sensitivas que os entrevistados afirmaram participar deve ser igual para as duas listas, significando que qualquer diferença de média pode ser atribuída ao fator compra de votos. Uma pergunta indiscreta sobre a compra de votos foi feita mais tarde, no *survey*; no entanto, é importante mencionar que a pergunta é se o entrevistado recebeu a oferta de algum serviço, presente ou outro benefício qualquer em troca dos seus votos, não se eles aceitaram a oferta. A barra inferior para uma resposta positiva para a pergunta sensitiva deveria, pelo menos, estreitar a distância entre as medidas. No entanto, a diferença entre as duas formas de medida é impressionante. A medida com a pergunta sensitiva mostra que apenas 5,5% dos entrevistados receberam um benefício, ao passo que o *List Experiment* revela que 19% venderam seus votos. Os intervalos de confiança são {4.2, 6.8} para a medida

que inclui a pergunta sensitiva e {18.1, 20.7} para a estimativa do método *List Experiment*.

A validade da estimativa do *List Experiment* tem como pressuposto que os entrevistados não mentem a respeito da pergunta sensitiva e que adicionando este item às respostas para o item de controle também não mudam (IMAI, 2011; BLAIR e IMAI, 2012). Muito embora estes pressupostos não sejam testáveis diretamente, Blair e Imai (2012) propõem um teste estatístico que pode ser usado para detectar certas distorções.

### **Usando as Respostas Previstas do *List Experiment***

**Tabela 1** – Resumo das respostas do *List Experiment* e da pergunta direta sobre compra de votos

Valor da resposta	<i>List experiment</i>				Pergunta direta		
	Grupo de controle		Grupo de tratamento		Valor da resposta	Frequência	Proporção
	Frequência	Proporção	Frequência	Proporção			
0	33	5,7%	39	6,8%	Não	1087	94,5%
1	256	44,5	208	36,1	Sim	63	5,5%
2	197	34,3	197	34,3			
3	73	12,7	89	15,5			
4			28	4,9			
Não respondeu	16	2,8%	14	2,4%	Não respondeu	0	0%
Total	575	100	575	100	Total	1150	100

Nota: De acordo com a simples diferença de estimativa média, o *List Experiment* indica que 19.4% (com 95% de intervalo de confiança de {18.1, 20.7} dos entrevistados venderam seus votos, números que são bem mais altos do que 5.5%, com {4.2, 6.8} da pergunta sensitiva direta.

Fonte: Elaboração própria.

Isto gera um valor-p de 0.27, indicando que não há violação dos pressupostos anunciados. Como qualquer outro teste estatístico, este resultado pode ser consequência da falta de poder estatístico. No entanto, devido a este promissor resultado inicial, nós pensamos em desenvolver uma nova metodologia estatística, a qual permite aos pesquisadores usarem as respostas do *List Experiment* como previsões em modelos de regressão.

### **A Metodologia Proposta**

Nesta seção, propomos métodos para usar as respostas do *List Experiment* em resultados de um modelo de regressão. Nós começamos por uma breve revisão de uma análise multivariada do *List Experiment* proposta por Imai (2011) e Blair e Imai

(2012). Então propomos duas novas formas de fazer estimativas. Primeiro, descrevemos um modelo de duas etapas, que utiliza as estimativas da regressão multivariada mencionada acima e tenta melhorar a *performance* do “ingênuo” modelo de duas etapas. Apesar de sua simplicidade, a mais importante limitação dessa abordagem é que ela não pode ser generalizada para além de uma simples regressão linear. O modelo também não é estatisticamente eficiente, seguidamente resultando em grandes erros padrão. Para superar estes problemas, nós propomos um ML geral para cálculo de estimativas que pode ser operacionalizado em uma única etapa. Muito embora este modelo seja mais desafiador operacionalmente, ele se mostra completamente eficiente e é aplicável a uma gama de modelos estatísticos.

### **A Configuração**

Suponhamos que nós temos uma simples amostra aleatória de  $N$  entrevistados de uma população. No mais simples *List Experiment*, cada entrevistado é aleatoriamente escolhido para ambos, o grupo de tratamento,  $T_i = 1$ , ou o grupo de controle,  $T_i = 0$ . Os entrevistados no grupo de controle são convidados **Respostas** a informar o número de itens afirmativos de uma lista de  $J$  itens de controle. Não raras vezes, estes itens de controle não são sensitivos, mas isto não é um pré-requisito. Em contraste, os entrevistados no grupo de tratamento são convidados a informar o número de itens afirmativos da lista  $J+1$  item, contendo os mesmos itens de  $J$  no grupo de controle e mais um item sensitivo. Digamos que  $Y_i$  representa a contagem relatada para este *List Experiment*. Para os entrevistados no grupo de controle,  $Y_i$  representa um número inteiro entre 0 e  $J$ . Para aqueles no grupo de tratamento,  $Y_i$  representa um número inteiro de 0 a  $J+1$ . Por fim, observa-se um vetor de covariáveis  $K$  para cada entrevistador, que é designado por um vetor na coluna  $X_i$  e  $X$ , onde  $X$  é o suporte de  $X_i$ .

O objetivo deste artigo é desenvolver novos métodos de como estimar o resultado de um modelo de regressão, cujas previsões incluem a resposta oculta (não observada) para o item sensitivo de um *List Experiment*. Para formalizar isto, nós seguimos os apontamentos usados em Imai (2011) e Blair e Imai (2012) que denotam a resposta oculta para o  $j$  th item na lista  $Z^*_{ij}$ . Sem a perda de generalidade, nós presumimos que o item sensitivo aparece no final da lista, com  $Z^*_{i, J+1}$  representando a resposta oculta para este item. Na prática, a ordem dos itens no *List Experiment* é com frequência aleatorizada para minimizar algum efeito de ordenamento, e a metodologia proposta é aplicável independentemente de um ordenamento particular dos itens. Finalmente nós usamos  $Y^*_i = \sum_{j=1}^J Z^*_{ij}$  para denotar o número de respostas afirmativas para os itens de controle  $J$ , o que é observado para

os entrevistados no grupo de controle, mas podem ser desconhecidos para aqueles no grupo de tratamento. Assim, o relacionamento entre o resultado observado e estas (possivelmente) variáveis ocultas é dado por

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

Ao longo deste artigo, como mencionado na *The List Experiment*, nós presumimos a inexistência de efeito de design ou mentira; estas suposições são formalmente descritas em Imai (2011) e Blair e Imai (2012).

Sob esta configuração, os pesquisadores podem se interessar pelo uso de respostas ocultas do *List Experiment* como um modo de fazer previsões em modelos de regressão para um resultado variável  $V_i \in \mathcal{V}$ , onde  $V(\cdot)$  é o suporte de  $V_i$ . Por exemplo, um dos mais simples, como este, é a regressão linear que segue abaixo:

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

onde o usual pressuposto de exogeneidade é feito com relação a ambas as variáveis, observáveis e não observáveis, isto é,  $\mathbb{E}(\epsilon_i | X_i, Z_{i,J+1}^*, Y_i^*) = 0$ . Sob este modelo, o principal parâmetro de interesse é o coeficiente para a resposta oculta para o item sensível, isto é,  $\gamma$ . Para além deste modelo de regressão linear, os pesquisadores podem estar interessados em estimar modelos não lineares tais como regressão logística, onde as respostas do *List Experiment* são usadas para fazer previsões. O problema, claro, é que nós não observamos  $Z_{i,J+1}^*$  e  $Y_i^*$  diretamente para muitos entrevistados. O Objetivo deste artigo é mostrar como estimar os resultados destes modelos de regressão com princípios.

### **O Modelo de Regressão Multivariada para o List Experiment: Uma Revisão**

A seguir, nós brevemente vamos revisar o modelo de regressão multivariada desenvolvido por Imai (2011) e Blair e Imai (2012) para analisar o *List Experiment*. Estes modelos formam a base da nossa nova metodologia introduzida nas seções 3.3 e 3.4. Imai (2011) e Blair e Imai (2012) propõe uma classe dos modelos de regressão multivariada que estão representados pelos seguintes submodelos:

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

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

Para  $z = 0, 1$ ,  $y = 0, 1, \dots, J$ , e  $x \in X$ , onde  $\delta$  e  $\psi$  são vetores de parâmetros desconhecidos.

O primeiro submodelo dado na equação (3) modela a probabilidade de uma resposta afirmativa para o item sensitivo, ao passo que o segundo submodelo dado na equação (4) representa um modelo para a resposta aos itens de controle  $J$  dado a resposta para o item sensitivo. Ambos os modelos incluem um vetor de covariáveis  $X_i$ . Um submodelo comum para o item sensitivo é a regressão logística, que é,  $g_{\delta}(x) = \exp(\delta_0 + x^T \delta_1) / \{1 + \exp(\delta_0 + x^T \delta_1)\}$  e  $\delta = (\delta_0, \delta_1)$ . De forma semelhante, um modelo binominal com um link logístico é frequentemente usado

como o submodelo para os itens de controle, isto é,  $h_{\psi}(y | x, z) = \binom{J}{y} p(x, z)^y \{1 - p(x, z)\}^{J-y}$ , onde

$$p(x, z) = \exp(\psi_0 + \psi_1 z + x^T \psi_2) / \{1 + \exp(\psi_0 + \psi_1 z + x^T \psi_2)\}$$

$$\text{e } \psi = (\psi_0, \psi_1, \psi_2).$$

Imai (2011) e Blair e Imai (2012) mostram como se conseguir estimativas consistentes de  $\delta$  e  $\psi$  usando mínimos quadrados e métodos de fazer estimativas ML. Particularmente, eles desenvolvem um algoritmo de maximização de expectativa (EM, em inglês), para o cálculo de estimativas ML completamente eficientes.

### **O Cálculo para Fazer Estimativas de Duas Etapas**

Nós primeiro consideramos o modelo de duas etapas usando a regressão multivariada descrita na seção anterior como o primeiro passo. A principal vantagem desta forma de cálculo de estimativa é a sua simplicidade. No entanto, este modelo pode apenas ser aplicado quando o resultado for linear. Além disso, a simplicidade desta forma de cálculo traz em contrapartida uma ineficiência estatística, como demonstraremos através de uma simulação na seção *Um estudo simulado*. Nós mostramos que o nosso modelo de duas etapas proposto melhora a performance de um mais ingênuo modelo de fazer estimativas.

Nos consideramos o resultado do modelo de regressão linear dado na equação (2). Observe que a variável de tratamento  $T_i$  não está incluída como regressor neste modelo porque a aleatoriedade a faz independente de  $(X_i, Z_{i,J+1}^*, Y_i^*)$ . A inclusão de  $T_i$ , embora possível, não afeta, em grande parte, propriedades estatísticas da nossa metodologia proposta sob as proposições padrão do *List Experiment*. Além disso, é possível não incluir  $Y_i^*$  como um regressor. Aqui, simplesmente o incluímos para ser mais completo. Todos os nossos resultados teóricos se aplicam ao modelo sem esta variável.

### **O ingênuo processo de duas etapas para fazer estimativas**

Nós começamos examinando o ingênuo processo de duas etapas para fazer estimativas, o qual é uma abordagem que os pesquisadores podem considerar. Nesta abordagem, primeiro ajustamos o modelo de regressão multivariada, tal como o descrito na Seção *O Modelo de Regressão Multivariada para o List Experiment: Uma Revisão*. Usando o modelo ajustado, computamos, para cada entrevistado, a probabilidade prevista das respostas afirmativas do item sensível como uma função das características dos entrevistados. Nós denotamos esta quantidade por  $\hat{Z}_{i,J+1}^* = g_{\delta}(X_i)$  (ver equação 3). Semelhantemente, se  $Y_i^*$  é incluído no resultado do modelo de regressão, podemos também estimar a esperada resposta para os itens de controle, isto é,  $\hat{Y}_i^* = \sum_{y=1}^J y \{h_{\psi}(y | X_i, 1) \hat{Z}_{i,J+1}^* + h_{\psi}(y | X_i, 0)(1 - \hat{Z}_{i,J+1}^*)\}$ . Na segunda etapa, nós incluímos estes valores previstos no resultado da regressão. Por exemplo, nós ajustaríamos o modelo de regressão linear dado na equação (2) substituindo  $Z_{i,J+1}^*$  e  $Y_i^*$  com seus valores previstos,  $\hat{Z}_{i,J+1}^*$  e  $\hat{Y}_i^*$ , respectivamente.

Infelizmente, este ingênuo processo de duas etapas sofre de diversos problemas. Primeiro, o processo é distorcido se o resultado da regressão não for linear. Segundo, o usual erro padrão vai ser subestimado porque o modelo não vai levar em consideração as incertezas das estimativas da primeira fase, as quais são normalmente grandes. Terceiro, e talvez mais importante, mesmo se o resultado do modelo de regressão for linear de forma que o processo não seja distorcido, ele se mostra excessivamente ineficiente estatisticamente para ser útil na prática. A razão para isto é que os valores previstos,  $\hat{Z}_{i,J+1}^*$  e  $\hat{Y}_i^*$  estão altamente correlacionados com outros regressores no modelo  $X_i$ . De fato, se uma simples regressão linear é usada na primeira etapa, o modelo não pode ser ajustado devido à perfeita colinearidade. Uma maneira de contornar este problema é excluir algumas covariáveis da segunda etapa da

regressão tratando elas como variáveis instrumentais. No entanto, a hipótese deste tipo de exclusão é frequentemente muito difícil de se justificar.

### **O Novo Processo para Fazer Estimativas de Duas Etapas**

Nós aperfeiçoamos o processo utilizando todas as informações do *List Experiment*. Como foi dito acima, focamos nossa atenção nos resultados das regressões lineares. O velho processo se baseia na expectativa condicional do resultado dado  $X_i$ . No novo processo, nos baseamos em todas as informações do *List Experiment*, isto é,  $(Y_i, T_i)$ , assim como as características dos entrevistados,  $X_i$ . Isto resulta na seguinte equação de regressão:

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

onde nós usamos a relação dada na equação (1), bem como no pressuposto de exogeneidade do modelo de regressão original dado na equação (2), isto é,  $\mathbb{E}(\epsilon_i | Z_{i,J+1}^*, Y_i^*, X_i) = 0$  implicando

$$\mathbb{E}(\epsilon_i | X_i, T_i, Y_i) = \mathbb{E}(\mathbb{E}(\epsilon_i | Z_{i,J+1}^*, Y_i^*, T_i, X_i) | X_i, T_i, Y_i) = \mathbb{E}(\mathbb{E}(\epsilon_i | Z_{i,J+1}^*, Y_i^*, X_i) | X_i, T_i, Y_i) = 0.$$

Além do mais, usando a regra de Bayes, a probabilidade desconhecida na equação (5) pode ser escrita como

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

Usando esta abordagem, o modelo de regressão linear dado na equação (5) pode ser ajustado seguindo um procedimento de duas etapas. Primeiro, nos estimamos os parâmetros desconhecidos  $(\delta, \psi)$  do modelo de regressão multivariada  $g_\delta(x)$  e  $h_\psi(y|x, z)$  dado na equação (3) e (4) usando o método de estimativa descrito em Imai (2011) e Blair e Imai (2012). Substituir  $g_\delta(x)$  e  $h_\psi(y|x, z)$  na equação (5) com as suas estimativas,  $\hat{g}_\delta(x)$  e  $\hat{h}_\psi(y|x, z)$ , e então ajustar os resultados deste modelo de regressão linear. Enfatizamos que a estimativa de probabilidade dada na equação (6) está prontamente disponível a partir do *E-step* do algoritmo EM desenvolvido por Imai (2011) e Blair e Imai (2012). O padrão de erro que incorpora

as incertezas das estimativas na primeira etapa pode ser calculado usando o método de reamostragem *bootstrap*, ou analiticamente no, simples mas entediante, método de momentos *framework* (IMAI, 2011).

### **O Geral e Completo processo ML de fazer estimativas**

Muito embora o processo de duas etapas desenvolvido acima aperfeiçoe o antigo e ingênuo método, ele ainda tem problemas pelo fato de ser aplicável apenas quando os resultados do modelo forem lineares. Além do mais, ele ainda não é completamente eficiente estatisticamente, porque na primeira etapa, ele não usa as informações dos resultados da regressão. Eficiência estatística é uma importante consideração quando analisamos *List Experiments*, porque a técnica de questões indiretas como o *List Experiment* reduz as distorções sacrificando a eficiência estatística.

Para enfrentar esta limitação, nós desenvolvemos o geral e completo processo ML de fazer estimativas. Este método de uma etapa incorpora todas as informações do quadro de probabilidades, portanto ele é completamente eficiente. O método também é geral no fato de que ele pode ser usado para uma gama enorme de modelos de resultados.

Considere o resultado geral do modelo definido a seguir:

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

Por exemplo, se o modelo de regressão linear normal é presumido, nós teremos

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

onde  $\theta = \{\alpha, \beta, \gamma, \zeta, \sigma\}$ . Dado o resultado deste modelo, nós podemos escrever os dados observados através da função

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

Infelizmente, maximizar o logaritmo dos dados observados desta função é difícil devido a sua estrutura complicada e misturada. No entanto, como foi feito no modelo de regressão multivariada para o *List Experiment* descrito na Seção *O Modelo de Regressão Multivariada para o List Experiment : Uma Revisão*, o algoritmo EM pode ser desenvolvido para resolver este problema de otimização.

Para fazer isso, primeiro derivamos uma função completa dos dados observados dada na equação:

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

Em cada iteração, o *M-step* do algoritmo EM maximiza a expectativa condicional da completa função de probabilidade, a qual é dada aqui:

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

Onde o peso  $\tilde{w}(X_i, Y_i, T_i, V_i)$  é dado pela seguinte *E-step*,

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

No algoritmo, este peso é avaliado usando-se o valor do parâmetro  $(\theta, \delta, \psi)$  obtido na iteração anterior. Note a ação deste peso sobre o resultado da variável  $V_i$ , enquanto que o peso usado no modelo de duas etapas dado na equação (6) não o faz. Esta diferença se constitui a fonte do ganho em eficiência deste modelo de apenas uma etapa. O nosso algoritmo ME começa após adequados valores são selecionados para todos os parâmetros, e ele é iterado até a convergência. Os erros padrão são calculados analiticamente computando a média da amostra do produto cruzado da pontuação dos vetores. A expressão detalhada desta assintótica analítica variância é fornecida no apêndice suplementar A.1.

Apesar de suas propriedades atrativas e flexibilidade, o proposto modelo ML tem uma desvantagem com relação ao modelo de duas etapas, que é o fato de aquele ser mais difícil de computar. A função de probabilidade apresenta uma complicada e misturada estrutura, e como o EM algoritmo pode terminar em uma posição máxima. Além disso, acrescentando muitas formas de fazer estimativas pode levar a uma instabilidade computacional. Uma possível solução seria desenvolver uma cadeia Markov de Monte Carlo algoritmo, mas vamos deixar este passo para futuras pesquisas.

## Um Estudo Simulado

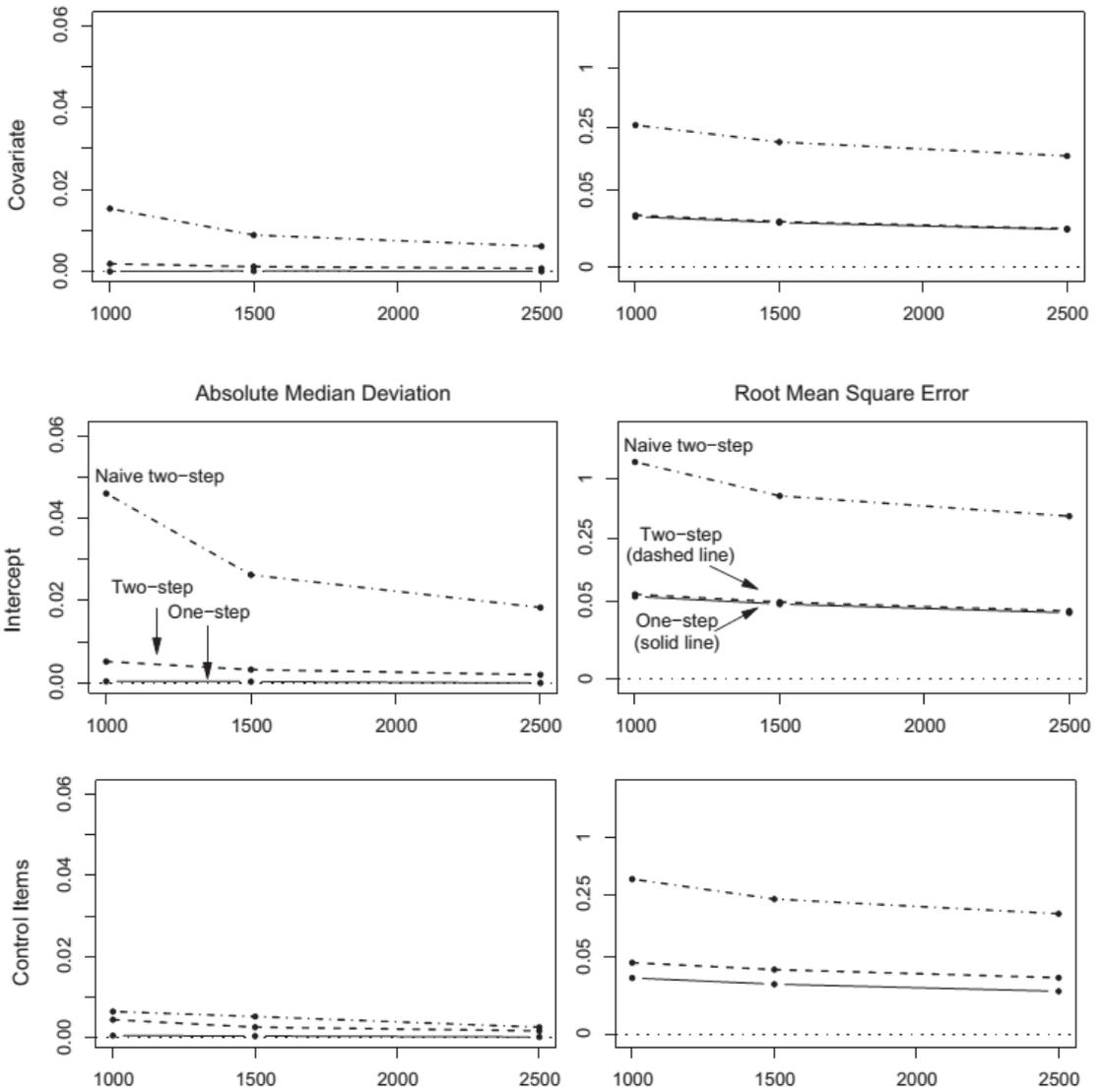
Nós agora conduzimos um estudo simulado para comparar os dois modelos de fazer estimativas acima, bem como o ingênuo modelo de duas etapas, usando o resultado de uma regressão linear (desta forma os três modelos podem ser comparados em pé de igualdade). Nós testamos uma única covariável  $X_i$  de uma distribuição normal uni variável, com média -1 e variância 1. Para o modelo sensitivo  $g_\delta(X_i)$ , nós usamos regressão logística com uma interceptação e  $X_i$ , onde os verdadeiros valores dos coeficientes são 0.5 e 1, respectivamente. Para as respostas dos itens de controle, nós usamos regressão logística binominal  $h_\delta(Y_i|X_i, Z_{i,J+1}^*)$ , onde os verdadeiros valores dos coeficientes para uma interceptação  $X_i$  e a variável oculta  $Z_{i,J+1}^*$  são ambos igual a 1. Finalmente, o resultado  $\hat{f}_\theta(V_i|X_i, Y_i^*, Z_{i,J+1}^*)$  é uma regressão linear, onde os

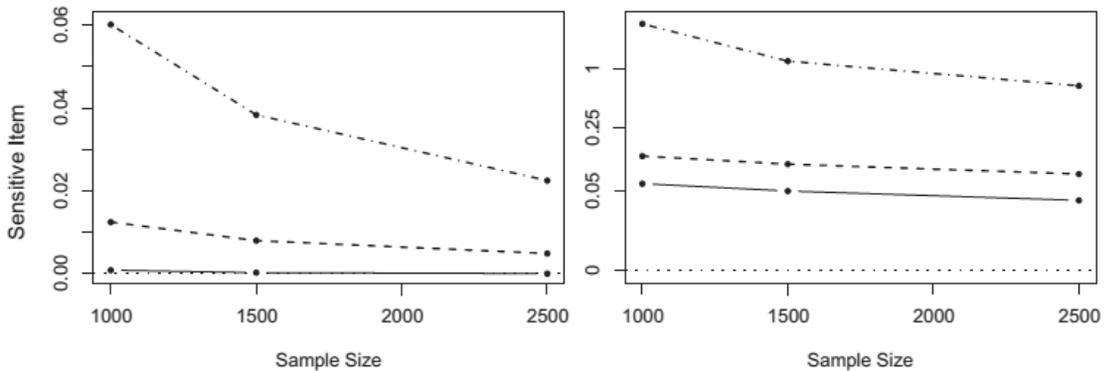
verdadeiros valores para os coeficientes, correspondendo a uma interceptação,  $X_i$ ,  $Y_i^*$  e  $Z_{i,j+1}^*$ , são iguais a (-1, 1, 1, 0.5).

Nós examinamos três diferentes tamanhos de amostras: 1000, 1500 e 2500. Para cada tamanho de amostra, nós, independentemente, simulamos um total de 100,000 grupos de dados utilizando um processo de geração de dados especificado aqui. Nós ajustamos o ingênuo modelo de duas etapas, uma etapa, e duas etapas para cada simulado grupo de dados e, em seguida, calculamos o Desvio Médio Absoluto e o erro médio quadrático (RMSE) através destes grupos de dados simulados. Utilizamos o Desvio Médio Absoluto sem viés como uma medida de erro sistemático para cada modelo porque o modelo de duas etapas é extremamente instável e suas distorções não podem ser estimadas confiavelmente em uma amostra pequena sempre acima de 100,000 grupos de dados simulados.

A Figura 1 mostra os resultados deste estudo simulado. Cada linha representa um parâmetro do resultado do modelo. Observe que RMSE é representado graficamente usando a figura (RMSE+0.01) para uma grande variância do ingênuo modelo de duas etapas, embora os rótulos dos eixos estejam na escala original. As linhas tracejadas mostram os resultados do modelo de duas etapas, enquanto as linhas sólidas representam os do modelo de uma etapa. As linhas com traços e pontos representam o ingênuo modelo de duas etapas, o qual mostra consistentemente a pior *performance*, em termos de Desvio Médio Absoluto e RMSE. Muito embora, como esperado, ambos os nossos modelos propostos recuperam a verdade, assim que o tamanho das amostras aumentam, o modelo de uma etapa supera o de duas etapas em amostras pequenas em ambas as métricas. O que é interessante, no entanto, é que a diferença de desempenho entre os dois modelos é a maior para o coeficiente do item sensitivo, o qual é o item de maior interesse aqui. Quando o tamanho da amostra é 1000 (o qual é um tamanho normal para pesquisas em Ciências Sociais), notamos que o desvio médio e o RMES do modelo de uma etapa são menos da metade dos do modelo de duas etapas. Isto demonstra claramente as vantagens relativas do modelo de uma etapa, o qual não é apenas mais flexível, mas é menos distorcido e mais eficiente.

Figura 1 – Simulação de estudo





Nota: Um estudo simulado comparando o desempenho do modelo de uma etapa (linhas sólidas), o modelo de duas etapas (linhas tracejadas), e o ingênuo modelo de duas etapas (linhas tracejadas com pontos). O tamanho da amostra varia de 1000, 1500, até 2500. Muito embora ambos os modelos propostos superem de maneira significativa o ingênuo modelo de duas etapas, o modelo de uma etapa supera o modelo de duas etapas proposto em termos de média de desvio absoluto e RMSE, especialmente para o coeficiente do item sensível. O RMSE é transformado antes de traçado; o gráfico mostra o tronco ( $RMSE + 0.01$ ), mas os rótulos do eixo y estão na escala original.

Fonte: Elaboração própria.

## Análise Empírica

Nesta seção, nós aplicamos o modelo de uma etapa no *México 2012 Panel Study* descrito na Seção Aplicação Empírica: *The México 2012 Panel Study*. Avaliamos a associação entre a compra de votos e dois resultados centrais na literatura sobre clientelismo: comparecimento eleitoral e aprovação do candidato vencedor Enrique Peña Nieto.

## Efeitos Sobre a Compra de Votos

### Dados e Modelo

Nós começamos introduzindo a nossa medida de comparecimento eleitoral. Como uma alternativa ao auto relato sobre o comparecimento eleitoral, que seguidamente gera números superestimados (BURDEN, 2000), os recenseadores no *México 2012 Panel Study* pediram para os entrevistados mostrarem os seus registros eleitorais e seus comprovantes de votação. Os eleitores que mostraram os seus registros e seus comprovantes de votação foram marcados como votantes efetivos, ao passo que aqueles sem o comprovante de votação foram marcados como não votantes. Alguns eleitores não mostraram seus registros quando perguntados. Nós codificamos estes como não votantes. Enquanto que o auto relato como votante, na amostra, era de 84%, a abordagem alternativa mediu o comparecimento eleitoral em 60.2%,

sendo esta medida muito próxima do relato oficial que foi de 64.5%. Provando a plausibilidade da nossa nova metodologia, providenciamos um modelo básico de comparecimento eleitoral que inclui as variáveis mais frequentemente sugeridas por pesquisadores de campo. Nós incluímos um conjunto de variáveis demográficas padrão: riqueza, idade, gênero e educação (WOLFINGER e ROSENSTONE, 1980; BLAIS, 2000). Cidadãos que são mais engajados em política são normalmente mais prováveis votantes, então nós incluímos uma medida de interesse sobre política assim como uma variável código do recenseador para capturar se o entrevistado colocou ou não propaganda política no lado de fora de sua casa<sup>3</sup>. Pela mesma razão, aqueles cidadãos que estão desiludidos com a política estão mais propensos a ficar em casa, então incluímos uma medida binária para descobrir se os entrevistados acreditavam que as eleições haviam sido limpas ou não (DOMÍNGUEZ e MCCANN, 1995). Todas estas covariáveis são também usadas para o modelo que prevê a compra de votos<sup>4</sup>.

Além destas covariáveis padrão, nós incluímos medidas de forte partidarização para as três principais partes para operacionalizar reivindicações já conhecidas na literatura sobre a compra de votos, descrita na Seção *Compra de Votos, Comparecimento Eleitoral, e Aprovação do Candidato*. Muito embora tal operacionalização não seja uma tarefa trivial, esta medida se constitui uma maneira de avaliar o argumento de Nichter (2008) que chefes buscando um aumento do comparecimento às urnas, oferecem benefícios seletivos para seus apoiadores. Neste caso, a inclusão da medida de forte partidarização permite-nos explorar o argumento de Cox e McCubbins (1986) de que eleitores apoiadores respondem melhor a incentivos seletivos do que eleitores apenas em potencial<sup>5</sup>.

---

<sup>3</sup> Nós observamos que os entrevistados que estejam mostrando propaganda pode estar mostrando suas preferências partidárias ou uma maneira de buscar algum benefício (NICTER e PALMER-RUBIN, 2015).

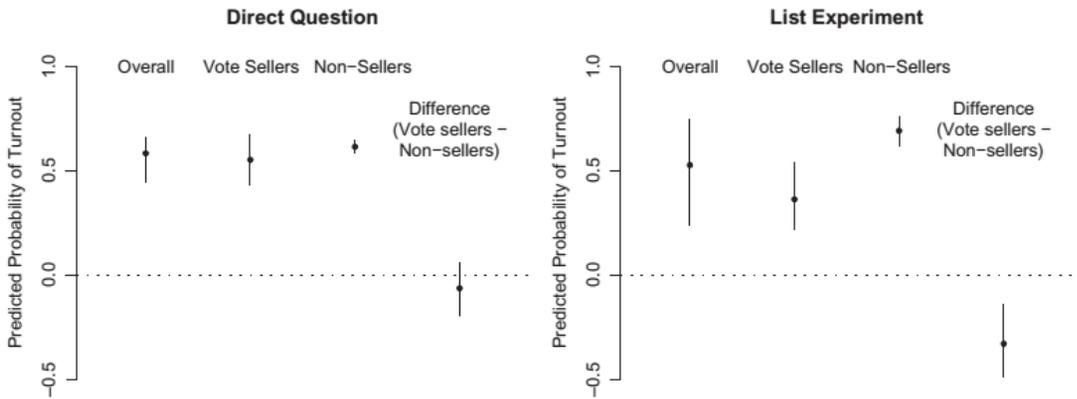
<sup>4</sup> Ao contrário do modelo sobre a aprovação do candidato na Seção *Efeito da Compra de Votos Sobre a Aprovação do Candidato*, nós não incluímos variáveis com indicadores regionais, porque a sua inclusão parece piorar o ajuste do modelo quando a variável é o comparecimento eleitoral.

<sup>5</sup> A questão usada para medir preferência partidária é uma normalmente usada nas pesquisas de opinião no México: “Geralmente, você se considera um PANista, PRIista ou PRDista? Você pensa ser um forte ou fraco PANista, PRIista ou PRDista?” “Os entrevistados que se identificaram fortemente com qualquer dos três maiores partidos foram codificados como apoiadores. Todos os partidos pode se engajar na compra de votos, mas os menores partidos são menos prováveis de fazê-lo efetivamente.

## Resultados Empíricos

Os coeficientes estimados do modelo descrito acima são apresentados no Apêndice Suplementar A.2. Aqui nos focamos na apresentação de resultados substantivos calculando a probabilidade média prevista do modelo ajustado.

Figura 2 – Probabilidade



Nota: A probabilidade de compra de votos e o efeito desta no comparecimento eleitoral produzido pelo *List Experiment*. O painel esquerdo mostra a probabilidade de compra de votos prevista com 95% de intervalo de confiança para a população em geral, e para os apoiadores e não apoiadores dos três principais partidos. Não há diferença estatisticamente significativa entre os apoiadores e não apoiadores dos partidos. O painel da direita mostra o efeito estimado da compra de votos no comparecimento eleitoral com 95% de intervalo de confiança para os mesmos três grupos de eleitores. Entre os apoiadores e não apoiadores, aqueles que venderam seus votos são os que mais provavelmente não irão votar. Novamente, não há diferença estatisticamente significativa entre o comportamento de apoiadores e de não apoiadores.

Fonte: Elaboração própria.

O painel esquerdo, na Figura 2, mostra a probabilidade de compra de votos. Nós ajustamos dois diferentes modelos separadamente: o primeiro para uma sub amostra de fortes inclinações partidárias e o segundo para o restante da amostra. Os nossos modelos fazem uma previsão de que 33,2% dos entrevistados (com um intervalo de confiança de 95% de [24,7, 42,9]) venderam seus votos e que não há diferença estatisticamente significativa entre apoiadores e não apoiadores. Muito embora haja muitas maneiras de se medir o conceito de eleitores certos, os nossos resultados indicam que os partidos no México não diferenciam apoiadores de não apoiadores, na hora de oferecer algum tipo de benefício em troca do voto, resultado este inconsistente com outros conceitos de eleitores certos como Cox e McCubbins (1986) ou a afirmação de Nichter de que os partidos deveriam comprar o

comparecimento nas urnas de seus apoiadores mais leais. No entanto, os nossos resultados são consistentes com outros, os quais enfatizam que os partidos oferecem benefícios para vários e diferentes grupos, ou por erro (STOKES *et al.*, 2013), ou como parte da estratégia eleitoral (MAGALONI, DIAZ-CAYEROS e ESTÉVEZ, 2007; GANS-MORSE, MAZZUCA e NICTER 2014).

O painel direito da Figura 2 apresenta os resultados de um modelo de regressão sobre comparecimento eleitoral. A compra de votos parece diminuir a probabilidade de comparecimento eleitoral em aproximadamente 31,7%, em média, independentemente do grau de apoio dos eleitores a um determinado partido (com 95% de intervalo de confiança de [19,8, 59,3] e [1,6, 47,0] para apoiadores e não apoiadores, respectivamente). Entre as duas sub amostras, de apoiadores e não apoiadores, foi possível observar uma relativamente forte correlação negativa entre a compra de votos e o comparecimento eleitoral.

Vale o alerta de que devido ao tamanho pequeno, as análises das sub amostras são menos robustas. Por exemplo, o modelo ajustado para a sub amostra de apoiadores partidários produz um grande erro padrão para alguns parâmetros no modelo item sensitivo. Isto pode indicar um fraco modelo ajustado devido ao menor tamanho da amostra. Retirar algumas variáveis (por exemplo, indicadores urbanos e de propaganda) não altera substancialmente os resultados, reduzindo a probabilidade estimada de compra de votos em algo próximo de 34,8% a 24,9% (95% de intervalo de confiança de [11,4%, 46,4%]). No entanto, o efeito da compra de votos sobre o comparecimento eleitoral entre os apoiadores partidários sob esta alternativa de análise é reduzido para 5,5% com um intervalo de confiança de 95% de [-38,6%, 20,5%].

Em uma análise superficial, estes resultados parecem derrubar o argumento de que a compra de votos aumenta os índices de comparecimento eleitoral, e ao invés disto, acaba embasando o embaraçoso argumento de que os chefes partidários compram a abstenção eleitoral (COX e KOUSSER, 1981). Embora possível, relatórios de campo não comprovam um massivo esquema de compra de abstenção eleitoral no México. Ao contrário, se os chefes partidários direcionam os benefícios para aqueles cuja probabilidade de ir votar é menor e seus esforços nem sempre são bem sucedidos, então os que vendem seus votos ainda assim podem estar menos propensos a comparecer no dia da eleição, em média, do que aqueles que não venderam seus votos. Esta discrepância pode existir mesmo se os pagamentos aumentarem a probabilidade de comparecimento eleitoral.

Nós observamos que estes simples resultados comportamentais não são discutidos na literatura sobre o assunto. Ao invés disto, estudos já existentes têm

principalmente focado suas atenções em amplas teorias sobre o relacionamento entre compra de votos e comparecimento eleitoral. Ferramentas como as nossas tem o potencial de mover a literatura para frente, não apenas por analisar relacionamentos empíricos, mas por trazer à tona relacionamentos ainda não teorizados. Como a Figura 3 demonstra, apenas usando as respostas do modelo *List Experiment*, ao invés de resposta a questões diretas, nós podemos detectar um efeito da compra de votos no comparecimento eleitoral.

## ***Efeito da Compra de Votos Sobre a Aprovação do Candidato***

### ***Dados e modelo***

Como discutido na Seção *Compra de Votos, Comparecimento Eleitoral, e Aprovação do Candidato*, a compra de votos pode também ser associada com a aprovação do candidato. Nas eleições mexicanas de 2012, a campanha de Peña Nieto foi acusada de usar benefícios seletivos em troca de apoio. Estes pagamentos poderiam ser associados com uma maior aprovação do candidato vencedor se os eleitores tivessem respondido aos benefícios. É possível também que ele tenha escolhido como público alvo os seus apoiadores do partido, aqueles que naturalmente já o tem em uma mais alta estima. No entanto, muito embora nós tenhamos encontrado uma associação entre a compra de votos e a aprovação de Peña Nieto, não foi possível encontrar evidências de o apoio partidário estar fortemente relacionado a compra de votos.

Uma limitação do *List Experiment* no *México 2012 Panel Study* é que ele não identifica qual partido ou candidato comprou os votos do eleitor. As eleições mexicanas de 2012 tiveram a escolha simultânea para vários cargos eletivos, presidente, congresso, bem como nos estados e municípios. Como resultado, os encarregados da tarefa poderiam ter tentado comprar o apoio para todo o partido, não apenas para o candidato a presidente. Além disso, os entrevistados poderiam já ter os seus candidatos em mente quando responderam o *List Experiment*. No entanto, a influência da eleição presidencial, o controle nacional sobre a propaganda eleitoral, e a relativa ausência de votantes de diferentes partidos em uma única cédula nos leva a conjecturar que os eleitores consideraram as eleições presidenciais na hora de responderem as questões do *List Experiment*. Em seguida, na questão direta sobre a compra de votos, indagando aos entrevistados que partido havia tentado comprar seus votos, muitos responderam que receberam pagamento de um partido ou campanha presidencial; nenhum entrevistado mencionou os candidatos a outros cargos eletivos.

Para poder explorar o relacionamento entre a compra de votos e a aprovação do candidato, nós ajustamos um modelo de regressão linear para analisar a opinião dos entrevistados sobre Peña Nieto em uma escala de 1 a 11, onde 1 representa a opinião menos favorável e 11 a mais favorável. Este modelo inclui o mesmo grupo de variáveis que o modelo sobre o comparecimento eleitoral, com alguns elementos a mais. Primeiro, separamos os entrevistados pelas quatro principais regiões geográficas do país (Norte, Central, Sul e a Região Metropolitana da Cidade do México), porque o apoio político e a estratégia de compra de votos podem variar por região geográfica no México. Nós também ajustamos o modelo para captar a identificação do eleitor com os três principais partidos, como um indicador separado. Isto nos permite ver se aqueles identificados com o partido de Peña Nieto são mais propensos a relatar a venda dos seus votos, assim como ajustamos a identificação partidária quando analisamos a opinião sobre cada candidato.

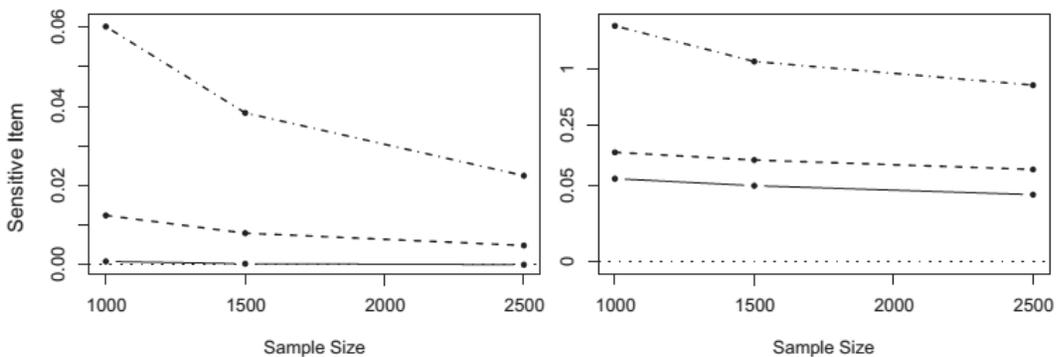
### **Resultados Empíricos**

Nós ajustamos o modelo descrito acima usando a resposta para a pergunta direta, assim como, as respostas previstas no *List Experiment*. Os coeficientes estimados destes modelos são apresentados no Apêndice Suplementar A.3. É interessante notar que os resultados destes dois modelos são substancialmente diferentes. O painel esquerdo na Figura 4 apresenta as estimativas baseadas na questão direta. Neste cenário, nós observamos que a diferença estimada entre os que venderam seus votos e os que não venderam é aproximadamente -0.82 (com 95% de intervalo de confiança de [-1,50, -0,16]), o que sugere que a compra de votos pode diminuir a avaliação positiva para Peña Nieto. No entanto, a conclusão baseada no *List Experiment* é exatamente oposta. No painel da direita, na mesma figura, podemos ver que muito embora o apoio estimado permaneça em um patamar parecido, a diferença entre aqueles que venderam seus votos e aqueles que não o fizeram é positiva e próxima dos 0.71 pontos (com 95% de intervalo de confiança de [-0,04, 1,47]). Isto é, quem vendeu o seu voto parece avaliar Peña Nieto mais positivamente.

As diferentes (e opostas) estimativas geradas pelos dois modelos sugerem as consequências das incorreções cometidas quando se lida e avalia a compra de votos. Se fossemos basear os nossos resultados principalmente na questão direta, nós teríamos concluído que a compra de votos não funciona como os chefes partidários gostariam e, na verdade, pode ter um efeito contrário do esperado, contrariando as afirmações empíricas e teóricas feitas por inúmeros autores, as quais revisamos na *Seção Aplicação Empírica: The México 2012 Panel Study*. Nós não acreditamos que

esta conclusão necessariamente está correta. Ao contrário, os erros de medição na questão direta são os prováveis culpados. O modelo logístico da resposta para a questão direta demonstra que os apoiadores de Peña Nieto podem ter se mostrado menos propensos a admitir que receberam pagamento (o coeficiente para o suporte ao PRI é negativo e significativo no patamar de 0.1), ao passo que as estimativas do item sensível do *List Experiment* mostram que os seus apoiadores estavam tão propensos a vender seus votos quanto os outros (ver o Apêndice Suplementar A.3 para as tabelas de coeficientes completas). Concluimos que os apoiadores de Peña Nieto se negaram a declarar suas atividades abertamente; a batalha legal no que diz respeito a compra de votos que continuou após as eleições pode ter desencorajado os beneficiários de admitir que eles receberam pagamentos.

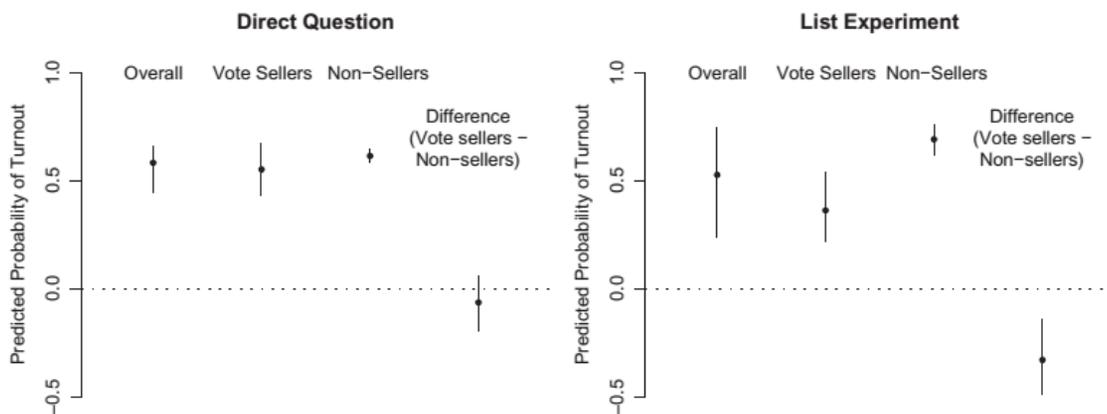
Figura 3 – Comparecimento eleitoral



Nota: Comparecimento eleitoral e o efeito da compra de votos sobre este comparecimento medidos pela pergunta direta e pelo List Experiment. A figura mostra o comparecimento eleitoral entre aqueles que relataram a venda de seus votos e aqueles que não o fizeram, medidos pela questão direta (painel esquerdo) e pelo List Experiment (painel direito). As barras verticais representam 95% de intervalo de confiança. De acordo com a questão direta, aqueles que venderam os seus votos não são menos prováveis de comparecerem as urnas do que aqueles que não venderam os seus votos. No entanto, o List Experiment produz um resultado diferente: a compra de votos tem um impacto negativo de aproximadamente 31,7 % sobre a probabilidade de comparecimento para votar.

Fonte: Elaboração própria.

Figura 4 – Índices de aprovação



Nota: Os índices de aprovação de Peña Nieto (PRI) e a os efeitos estimados da compra de votos oriundos da questão direta e do List Experiment. A figura mostra o índice de aprovação, o qual vai de 1 (mais baixo nível) a 11 (mais alto nível) de Enrique Peña Nieto, o candidato vencedor (do PRI), entre aqueles que relataram que venderam os seus votos e aqueles que não o fizeram, como foram apurados pela questão direta (painel esquerdo) e o List Experiment (painel direito). As barras verticais representam 95% de intervalo de confiança. De acordo com a questão direta, aqueles que venderam os seus votos avaliam o candidato aproximadamente 1 ponto mais baixo do que aqueles que não venderam os seus votos. No entanto, o List Experiment produz um resultado oposto: aqueles que venderam seus votos apresentam um impacto positivo de 0,71 na avaliação positiva do candidato.

Fonte: Elaboração própria.

O *List Experiment* não apenas oferece uma melhor medição da atividade da compra de votos, mas também produz o mais intuitivo resultado que afirma que a compra de votos está associada com uma mais alta aprovação de Peña Nieto. A história de compra de votos por seu partido (CORNELIUS e CRAIG, 1991; GREENE, 2007) e outras evidências vindas do campo de pesquisa durante a campanha para as eleições de 2012 sugerem que a campanha de Peña Nieto tentou mais comprar votos do que seus rivais. A associação entre a compra de votos e a aprovação de Peña Nieto poderia indicar que a sua campanha tentou o pagamento de seus apoiadores ou que a compra de votos funcionou como era esperado. A hipótese de tentativa se mostra menos plausível dado que a preferência partidária com relação ao PRI não está associada com a compra de votos no modelo de item sensitivo. É mais provável que a compra de votos melhorou a aprovação de Peña Nieto, um resultado consistente com Stokes *et al.* (2013) e que nós investigamos mais tarde em um outro trabalho (GREENE, 2014).

## **Observações finais**

O *List Experiment* tem obtido mais popularidade recentemente entre os cientistas sociais como uma metodologia de *survey* no sentido de conseguir respostas mais confiáveis para questões sensíveis. Muito embora métodos estatísticos tenham sido propostos para modelar respostas em itens sensíveis como uma variável dependente, nenhum modelo existente permite o uso das respostas do *List Experiment* como variáveis explicativas em um modelo de regressão. E ainda, os pesquisadores podem querer conhecer os efeitos do comportamento sensível e atitudes sobre outras ações e opiniões. Nós tentamos preencher esta lacuna metodológica propondo uma nova técnica de fazer estimativas. O nosso modelo de uma etapa pode ser aplicado em uma variada gama de regressões, desse modo providenciando uma ferramenta flexível para os pesquisadores.

A abordagem metodológica proposta neste artigo pode ser estendida para outras configurações. Por exemplo, pode-se generalizar este método para outros *designs* do *List Experiment* como o duplo *List Experiments* (DROITCOUR *et al.*, 1991; GLYNN, 2013) ou outros métodos de questões indiretas como o *Endorsement Experiments* (BULLOCK, IMAI e SHAPIRO, 2011) e o método de respostas aleatórias (WARNER, 1965; BLAIR, IMAI e ZHOU, 2014). Estes e outros projetos de pesquisas relacionados estão sendo conduzidos atualmente pelos autores e outros pesquisadores.

Finalmente, na medida em que o *List Experiment* se torna popular, é importante para quem trabalha com pesquisas aplicadas levar muito a sério os pressupostos requeridos. Testes estatísticos para detectar violações desses pressupostos e estratégias de modelos para a sua utilização foram desenvolvidos (BLAIR e IMAI 2012). Outra estratégia é agregar o nível individual dos resultados do *List Experiment* e validá-los usando verdadeiras informações disponíveis (ROSENFELD, IMAI e SHAPIRO, 2014). Nós acreditamos que como em qualquer pesquisa, um extensivo estudo piloto e o uso de diferentes fontes de dados são essenciais para fazer confiáveis análises empíricas.

Usado criteriosamente, o *List Experiment* e outras técnicas de questionamento indireto podem ser poderosas ferramentas para se entender comportamentos sensíveis e atitudes em uma variedade de disciplinas. O método proposto aumenta a utilidade do *List Experiment* por permitir aos pesquisadores utilizarem informações sensíveis de diferentes maneiras, as quais eram impossíveis anteriormente.

■ Kosuke Imai é Professor de Princeton University. E-mail: kimai@princeton.edu.

■ Bethany Park é Estudante de Graduação em Princeton University. E-mail: bapark@princeton.edu.

■ Kenneth Greene é Professor Associado da University of Texas, Austin. E-mail: kgreene@austin.ut.edu.

---

## Referências

- AHLQUIST, John; MAYER, Kenneth; JACKMAN, Simon. Alien abduction and voter impersonation in the 2012 US general election: Evidence from a survey list experiment. *University of Wisconsin, Madison*, Technical Report, 2013.
- BIEMER, Paul; BROWN, Gordon. Model-based estimation of drug-use prevalence using item count data. *Journal of Official Statistics*, v. 21, n. 2, p. 287-308, 2005.
- BLAIR, Graeme; IMAI, Kosuke. Statistical analysis of list experiments. *Political Analysis*, v. 20, n. 1, p. 47-77, Winter 2012.
- BLAIR, Graeme; IMAI, Kosuke; LYALL, Jason. Comparing and combining list and endorsement experiments: Evidence from Afghanistan. *American Journal of Political Science*, v. 58, n. 4, p. 1043-063, fev. 2014.
- BLAIR, Graeme, IMAI, Kosuke; PARK, Bethany. 2014. *List*: Statistical methods for the item count technique and list experiment. Available at the Comprehensive R Archive Network (CRAN). Disponível em: <<http://CRAN.R-project.org/package=list>>. Acesso em: 1 jun. 2014.
- BLAIR, Graeme, IMAI, Kosuke; ZHOU, Yang-Yang. Design and analysis of randomized response technique. Pittsburgh, USA: Princeton University, 2014.
- BLAIS, André. *To vote or not to vote: The merits and limits of rational choice theory*. Pittsburgh, USA: University of Pittsburgh Press, 2000.
- BRUSCO, Valeria; NAZARENO, Marcelo; STOKE, Susan. Vote buying in Argentina. *Latin American Research Review*, v. 39, n. 2, p. 66-88, jun. 2004.
- BULLOCK, Will; IMAI, Kosuke; SHAPIRO, Jacob N. Statistical analysis of endorsement experiments: Measuring support for militant groups in Pakistan. *Political Analysis*, v. 19, n. 4, p. 363-384, Autumn 2011.
- BURDEN, Barry C. Voter turnout and the National Election Studies. *Political Analysis*, v. 8, n. 4, p. 389-398, 2000.
- CALVO, Ernesto; MURILLO, Maria Victoria. Who delivers? Partisan clients in the Argentine electoral market. *American Journal of Political Science*, v. 48, n. 4, p. 742-757, set. 2004.
- CARRERAS, Miguel; IREPOGLU, Yasemin. Trust in elections, vote buying, and turnout in Latin America. *Electoral Studies*, v. 32, n. 4, p. 609-619, dez. 2013.

- CORNELIUS, Wayne. Mobilized voting in the 2000 elections: The changing efficacy of vote buying and coercion in Mexican electoral politics. In: DOMINGUEZ, Jorge; LAWSON, Chappell (Eds.). *Mexico's Pivotal Democratic Election: Candidates, votes and the presidential campaign of 2000*. Stanford, USA: Stanford University Press, 2004. p. 47-65.
- CORNELIUS, Wayne; CRAIG, Ann. *The Mexican political system in transition*. La Jolla, USA: Center for U.S.-Mexican Studies, 1991.
- CORSTANGE, Daniel. Sensitive questions, truthful answers? Modeling the list experiment with LISTIT. *Political Analysis*, v. 17, n. 1, p. 45-63, 2009.
- \_\_\_\_\_. Vote-trafficking in Lebanon. *International Journal of Middle East Studies*, v. 44, n. 3, p. 483-505, ago. 2012.
- COX, Gary W.; KOUSSER, J. Morgan. Turnout and rural corruption: New York as a test case. *American Journal of Political Science*, v. 25, n. 4, p. 646-663, nov. 1981.
- COX, Gary W.; MCCUBBINS, Mathew D. Electoral politics as a redistributive game. *Journal of Politics*, v. 48, n. 2, p. 370-389, maio 1986.
- DIAZ-CAYEROS, Alberto; ESTÉVEZ, Federico; MAGALONI, Beatriz. Welfare benefits, canvassing, and campaign handouts. In: J. DOMINGUEZ, Jorge; LAWSON, Chappell; MORENO, Alejandro (Eds.). *Consolidating Mexico's democracy: The 2006 presidential campaign in comparative perspective*. Baltimore, USA: Johns Hopkins University Press, 2009. p. 229-245.
- DIXIT, Avinash; LONDREGAN, John. The determinants of success of special interests in redistributive politics. *Journal of Politics*, v. 58, n. 4, p. 1132-1155, nov. 1996.
- DOMÍNGUEZ, Jorge; MCCANN, James. Shaping Mexico's electoral arena: The construction of partisan cleavages in the 1988 and 1991 national elections. *American Political Science Review*, v. 89, n. 1, p. 34-48, mar. 1995.
- DROITCOUR, Judith *et al.* The item count technique as a method of indirect questioning: A review of its development and a case study application. In: BIEMER, Paul P. *et al.* (Eds.). *Measurement errors in surveys*. New York: John Wiley & Sons, 1991. p. 185-210
- GANS-MORSE, Jordan; MAZZUCA, Sebastián; NICTER, Simeon. Varieties of clientelism: Machine politics during elections. *American Journal of Political Science*, v. 58, n. 2, p. 415-432, abr. 2014.
- GLYNN, Adam N. What can we learn with statistical truth serum? Design and analysis of the list experiment. *Public Opinion Quarterly*, v. 77, n. S1, p. 159-72, 2013.
- GONZALEZ-OCANTOS, Ezequiel *et al.* Vote buying and social desirability bias: Experimental evidence from Nicaragua. *American Journal of Political Science*, v. 56, n. 1, p. 202-17, jan. 2012.
- GREENE, Kenneth F. *Why dominant parties lose: Mexico's democratization in comparative perspective*. New York: Cambridge University Press, 2007.
- \_\_\_\_\_. Back from the dead: Vote-selling zombies and the return of Mexico's PRI. *Latin American Studies Association Congress*, Chicago, 2014.
- GREENE, Kenneth F. *et al.* *The Mexico 2012 Panel Study*. 2012. Disponível em: <<http://kgreene.webhost.utexas.edu>>.
- HOLBROOK, Allyson L.; KROSNICK, Jon A. Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly*, v. 74, n. 1, p. 37-67, 2010.
- IMAI, Kosuke. Multivariate regression analysis for the item count technique. *Journal of the American Statistical Association*, v. 106, n. 494, p. 407-16, jun. 2011.

- Imai, Kosuke, 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, James G.; CRAIG, Stephen C.; WALD, Kenneth D. Religion and presidential politics in Florida: A list experiment. *Social Science Quarterly*, v. 85, n. 2, p. 281-93, jun. 2004.
- KRAMON, Eric. Vote-buying and political behavior: Estimating and explaining vote-buying's effect on turnout in Kenya. *Afrobarometer*, Working Papers, p. 1-31, 2009.
- KUKLINSKI, James H.; COBB, Michael D.; GILENS, Martin. Racial attitudes and the "New South". *Journal of Politics*, v. 59, n. 2, p. 323-49, maio 1997.
- LINDBECK, Assar; WEIBULL, Jörgen W. Balanced-budget redistribution as the outcome of political competition. *Public Choice*, v. 52, n. 3, p. 273-297, 1987.
- MAGALONI, Beatriz; DIAZ-CAYEROS, Alberto; ESTÉVEZ, Federico. Clientelism and Portfolio Diversification: A Model of Electoral Investments with Applications to Mexico. In: KITSCHELT, Herbert; WILKINSON, Steven. *Patrons, Clients, and Policies: Patterns of Democratic Accountability and Political Competition*. New York, USA: Cambridge University Press, 2007. p. 182-204.
- NICHTER, Simeon. Vote buying or turnout buying? Machine politics and the secret ballot. *American Political Science Review*, v. 102, n. 1, p. 19-31, fev. 2008.
- NICHTER, Simeon; PALMER-RUBIN, Brian. Clientelism, declared support, and Mexico's 2012 campaign. In: DOMINGUEZ, Jorge *et al.* (Eds.). *Mexico's evolving democracy: A comparative study of the 2012 elections*. Baltimore, USA: Johns Hopkins University Press, 2015. p. 220-26.
- PENFOLD-BECERRA, Michael. Clientelism and social funds: Evidence from Chavez's Misiones. *Latin American Politics and Society*, v. 49, n. 4, p. 63-84, dez. 2007.
- ROSENFELD, Bryn; IMAI, Kosuke; SHAPIRO, Jacob. An empirical validation study of popular survey methodologies for sensitive questions. 2014. Disponível em: <<http://imai.princeton.edu/research/validate.html>>. Acesso em: 1 jun. 2014.
- STOKES, Susan. Perverse accountability: A formal model of machine politics with evidence from Argentina. *American Political Science Review*, v. 99, n. 3, p. 315-325, 2005.
- STOKES, Susan *et al.* *Brokers, voters, and clientelism: The puzzle of distributive politics*. Cambridge Studies in Comparative Politics. Cambridge, UK: Cambridge University Press, 2013.
- STREB, Matthew J. *et al.* Social desirability effects and support for a female American president. *Public Opinion Quarterly*, v. 72, n. 1, p. 76-89, 2008.
- TOURANGEAU, Roger; YAN, Ting. Sensitive questions in surveys. *Psychological Bulletin*, v. 133, n. 5, p. 859-83, 2007.
- WARNER, Stanley L. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, v. 60, n. 309, p. 63-69, mar. 1965.
- WIMBUSH, James C.; DALTON, Dan R. Base rate for employee theft: Convergence of multiple methods. *Journal of Applied Psychology*, v. 82, n. 5, p. 756-63, 1997.
- WOLFINGER, Raymond; ROSENSTONE, Steven. *Who votes?* New Haven, USA: Yale University Press, 1980.

*Publicado originalmente como:*

IMAI, Kosuke; PARK, Bethany; GREENE, F. Kenneth. Using the Predicted Responses from List Experiments as Explanatory Variables in Regression Models. *Political Analysis*, v. 23, n. 2, p. 180-196. 2015.

---

*Texto recebido em 06 de abril de 2015.  
Aprovado em 16 de abril de 2015.*

