

**List Experiments for Estimating Vote Fraud in US Elections:
The 32 Percent of Republicans Abducted By Aliens Can Be Wrong**

L.J Zigerell
Assistant Professor
Illinois State University
404 Schroeder Hall
Normal, IL 61790
ljzigerell@ilstu.edu
@LJZigerell

Abstract. List experiments reported in Ahlquist et al. (2014) were used to support the claim that vote fraud was not common in the 2012 US election. However, these list experiments produced confidence intervals at least 16 percentage points wide and thus did not have sufficient power to detect the low-but-substantively-important levels of vote fraud that might be expected in the United States. Moreover, reanalysis of the list experiment data produced an estimate that 12% of registered voters were contacted by political candidates or activists with an offer to exchange money or gifts for a vote, but it is difficult to take this estimate seriously, given that another list experiment in the same survey produced an estimate that 10% of registered voters had been recently abducted by aliens. The data reported in Ahlquist et al. (2014) thus do not provide nontrivial information about the amount of vote fraud in the 2012 election.

Acknowledgements. The author thanks John Ahlquist for posting the Ahlquist et al. (2015) data and for answering questions about Ahlquist et al. (2014).

List Experiments for Estimating Vote Fraud in US Elections: The 32 Percent of Republicans Abducted By Aliens Can Be Wrong¹

Movements to require more stringent laws to prevent vote fraud in US elections have been described by some as an effort to secure the integrity of the ballot and by others as an effort to disenfranchise poor, elderly, and black voters. Criminal convictions indicate that the amount of vote fraud is nonzero: between October 2002 and September 2005, 92 persons were convicted of vote buying, ineligible voting, multiple voting, registration fraud, voter intimidation, or ballot forgery (Lipton and Urbina 2007). However, such data from criminal convictions provide no evidence that vote fraud is widespread.

Critics of stricter voter laws propose that the lack of evidence of more widespread vote fraud is because vote fraud is rare, but proponents of stricter voter laws propose that the lack of close oversight means that the government lacks the ability to catch most instances of voter fraud. To address these competing claims, Ahlquist et al. (2014) reported results from two survey experiments indicating that "strict voter ID requirements address a problem that was certainly not common in the 2012 U.S. election" (p. 460). However, the list experiment method that Ahlquist et al. used to detect voter fraud could detect only relatively large amounts of vote fraud and is thus unable to provide meaningful information about the at most small-but-impactful levels of vote fraud in the United States.

¹ Acknowledgements. Thanks to John S. Ahlquist for publicly posting the Ahlquist et al. (2015) data and for answering questions about the surveys. Such transparency should be highly commended.

List experiments on vote fraud cannot detect low levels of vote fraud

The list experiment is a common method to estimate the prevalence of socially undesirable behaviors and beliefs (e.g., Miller 1984, Kuklinski et al. 1997, Kane et al. 2004). The logic of the list experiment is that a respondent will be more likely to admit to a socially undesirable behavior or belief if the respondent can mask admission of the behavior or belief.

List experiments reported on in Ahlquist et al. (2014) divided respondents into two groups: members of one group received a list of four items and instructions to indicate how many items described activities that the respondent was involved in during a given time period; members of the other group received the same instructions and list, but with a fifth item of interest. Respondents then reported a number indicating how many – but not which – items applied to the respondent. The presumption of the list experiment is that respondents who report a number less than the maximum have masked whether any of the items apply to the respondent. Researchers can then subtract the mean number of items reported in the four-item group from the mean number of items reported in the five-item group to estimate the percent of respondents to which the fifth item applies.

List experiments have detected evidence of vote trafficking or intimidation in several countries at conventional levels of statistical significance. Point estimates from list experiments have indicated that 15% of a sample of Russian citizens felt that their vote in the 2011 State Duma elections would affect their job security, benefits, and/or income (Frye et al. 2014), 24% of a sample of registered voters in Nicaragua received a gift or favor from candidates or activists during the 2008 campaign (Gonzalez-Ocantos et al. 2012), 35% of a sample of Turks had their decision for whom to vote influenced by something of value

being offered to themselves or their family (Çarkoğlu and Aytaç 2015), and 55% of a sample of Lebanese had their decision to vote or for whom to vote influenced by something of value being offered to themselves or their relatives (Corstange 2012).

Each of the aforementioned countries scored poorly on Transparency International's Corruptions Perceptions Index: for 2014 data, on a scale of 0 to 100 where higher numbers indicate less corruption, respective scores for Russia, Lebanon, Nicaragua, and Turkey were 27, 27, 28, and 45. The standard deviation of the corruption scale was roughly 20 points, so the United States score of 74 was more than one standard deviation higher than the score for Turkey and more than two standard deviations higher than the scores for Russia, Lebanon, and Nicaragua. These scores reflect perceptions of a country's public sector corruption and not corruption among citizens or political parties, but – to the extent that private and public corruption are correlated – the scores provide at least suggestive evidence that the United States would not have levels of vote trafficking or intimidation similar to countries in which list experiments have successfully produced evidence of vote trafficking or intimidation.

Table 1 presents estimates of the prevalence of the four behaviors investigated in the 3,000-person list experiments from September 2013 described in Ahlquist et al. (2014), based on data from Ahlquist et al. (2015).² Ahlquist et al. (2014) alluded to the statistical

² See Ahlquist et al. (2014) for more detail on the two list experiments: a 1,000-person list experiment from December 2012, and a 3,000-person list experiment from September 2013. That article contains the wording for the list experiments that did not concern attempted vote buying. The wording for the 3,000-person attempted vote buying experiment was: "Here are some things that political candidates or activists may have done in your area during the election this past November. HOW MANY of these things happened to you?" [1] Political candidates or activists put up campaign posters or signs in your neighborhood or city." [2] Political candidates or activists visited your home." [3] Political candidates or activists placed campaign advertisements on television or radio." [4] Political

power of the list experiments: "We cannot reject the null that the amount of voter impersonation is 0% but nor can we reject the null that the amount of fraudulent voting is 1%" (p. 473). However, regarding the statistical power of the 3,000-person list experiment to detect attempted vote buying, it is not possible to reject the null that the percentage of Americans offered something for their vote is 0%, or 5%, or 10%, or 15%. It is not clear whether the beliefs of any serious observer of US elections would be challenged by a [-7%, 18%] confidence interval for the prevalence of attempted vote buying in the United States.

[Table 1 about here]

For the 3,000-person list experiments, confidence intervals for voter impersonation have a width similar to confidence intervals for vote buying, but the weighted upper bound for voter impersonation is substantially lower, at 1%. However, the corresponding point estimate is an impossible *negative* 10 percent, so caution is warranted if interpreting the 1% upper bound estimate as rejecting the possibility that the prevalence of voter impersonation is not 2% or higher; after all, the weighted point estimate for the 1,000-person list experiment was 0%, and centering at 0% the weighted confidence interval from the 3,000-person list experiment for voter impersonation produces an upper bound of 11%.

Disaggregated analyses produced evidence of vote buying, alien abduction, and impossibilities

Reflecting the idea that voter impersonation might be more common in particular subgroups, Ahlquist et al. (2014: 467ff) reported results from the 1,000-person list

candidates or activists personally threatened you." [Item of interest 5] Political candidates or activists offered you money or a gift for your vote." The content of the items was similar for the 1,000-person list experiment.

experiment for groups such as whites, Democrats, and voters in contested states; however, Ahlquist et al. did not report disaggregated results for the larger, 3,000-person list experiment, which had more power to detect vote fraud among disaggregated groups.

It makes sense for vote buyers to target persons with votes to sell, so it is reasonable to test for vote buying among registered voters. Restricting the list experiment to the 2,643 persons in the 3,000-person sample who reported being registered to vote, the list experiment estimated that the percentage contacted by political candidates or activists with an offer to exchange money or gifts for a vote was 12% in the weighted data ($p=0.018$) and 9% in the unweighted data ($p=0.058$), with respective 95% confidence intervals of [2%, 22%] and [0%, 18%].

However, before the aforementioned evidence is used to justify stricter punishments for vote solicitation, consider that another list experiment in the same survey produced estimates for the percentage of registered voters abducted by aliens of 10% in the weighted data ($p=0.017$) and 7% in the unweighted data ($p=0.051$). Moreover, analyses disaggregated by partisanship produced estimates for the percentage of Republicans abducted by aliens of 32% in the weighted data ($p=0.006$) and 15% in the unweighted data ($p=0.047$).

Moreover, analyses disaggregated by race and ethnicity produced physically impossible estimates. Based on the 300 Hispanics in the 3,000-person sample, the percentage of Hispanics involved in voter impersonation during the November 2012 election was *negative* 55% in the weighted data ($p=0.004$) and *negative* 26% in the unweighted data ($p=0.054$), which might reflect Hispanics insincerely deflating reports for the vote impersonation list as a method by which to send a clear signal of non-participation

in vote impersonation (see Zigerell 2011). However, negative estimates were not present for Hispanics in the vote buying list experiment, which produced weighted and unweighted estimates of 37% ($p=0.106$) and 13% ($p=0.389$), so the negative and statistically significant estimates for the voter impersonation list experiment might only or mainly reflect randomization error that would be expected on occasion when conducting disaggregated analyses.

Conclusion

Measurement error makes it difficult to use direct questions to accurately measure the prevalence of rare events (Ansolabehere et al. 2015); social desirability makes it difficult to accurately measure phenomena with a list experiment (Zigerell 2011); and the wide confidence intervals of a list experiment make it difficult to *precisely* measure phenomena with a list experiment.

Ahlquist et al. (2014) reported on two list experiments regarding the prevalence of voter impersonation in the United States, but the confidence intervals for these list experiments were 38 and 23 percentage points wide; these list experiments thus do not provide meaningful information about the prevalence of vote fraud in the United States, given that these confidence intervals centered at zero contain all reasonable beliefs about the prevalence of vote fraud in the United States.

It is also difficult to place much trust in point estimates from these list experiments: weighted point estimates from the 1,000-person and 3,000-person list experiments were respectively 0% and -10% for voter impersonation and -4% and 5% for attempted vote buying; these 10 and 9 percentage-point fluctuations between list experiments are almost

certainly larger than the percentage of Americans who commit voter impersonation or who are propositioned by partisans for their votes.³

Results from the 3,000-person list experiment did reveal evidence at conventional levels of statistical significance that a nontrivial percentage of registered US voters participated in vote buying during the 2012 election, but skepticism is warranted for these results, given that list experiments from the same survey produced improbable and impossible results at conventional levels of statistical significance.

Ahlquist et al. admitted to selecting the alien abduction item for rhetorical purposes (2014, footnote 29), possibly alluding to their conclusion that "about as many people admit to alien abduction as admit to voter impersonation" (p. 473). But there is a different interpretation for the alien abduction item: if a measuring tool cannot be used to reject the claim that 15% of Americans were recently abducted by aliens, then that tool is not precise enough to have much value for estimating the prevalence of vote fraud in the United States.

³ Note that the 1,000-person vote buying list experiment lacked a response option for 0 items for the four-item and five-item groups.

References

- Ahlquist, John S., Kenneth R. Mayer, and Simon Jackman. 2014. "Alien Abduction and Voter Impersonation in the 2012 US General Election: Evidence from a Survey List Experiment." *Election Law Journal* 13(4): 460-475.
- Ahlquist, John; Mayer, Kenneth; Jackman, Simon. 2015. Replication Data for: "Alien Abduction and Voter Impersonation in the 2012 U.S. General Election: Evidence from a Survey List Experiment." <http://dx.doi.org/10.7910/DVN/DGH00R>, Harvard Dataverse, V1 [UNF:6:H4A7EKMTcworEZRXsldnpg==]
- Ansolabehere, Stephen, Samantha Luks, and Brian F. Schaffner. 2015. "The Perils of Cherry Picking Low Frequency Events in Large Sample Surveys." *Electoral Studies*. Available online before print.
- Çarkoğlu, Ali, and S. Erdem Aytaç. 2014. "Who Gets Targeted for Vote-Buying? Evidence from an Augmented List Experiment in Turkey." *European Political Science Review*. Available online before print.
- Corstange, Daniel. 2012. "Vote Trafficking in Lebanon." *International Journal of Middle East Studies* 44(3): 483-505.
- Frye, Timothy, Ora John Reuter, and David Szakonyi. 2014. "Political Machines at Work: Voter Mobilization and Electoral Subversion in the Workplace." *World Politics* 66(2): 195-228.
- Gonzalez-Ocantos, Ezequiel, Chad Kiewiet De Jonge, Carlos Meléndez, Javier Osorio, and David W. Nickerson. 2012. "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua." *American Journal of Political Science* 56(1): 202-217.
- Kane, James G., Stephen C. Craig, and Kenneth D. Wald. 2004. "Religion and Presidential Politics in Florida: A List Experiment." *Social Science Quarterly* 85(2): 281-293.
- Kuklinski, James H., Michael D. Cobb, and Martin Gilens. 1997. "Racial Attitudes and the 'New South'." *Journal of Politics* 59(2): 323-349.
- Lipton, Eric, and Ian Urbina. 2007. "In 5-Year Effort, Scant Evidence of Voter Fraud." *New York Times*. April 12.
- Miller, Judith Droitcour. 1984. A New Survey Technique for Studying Deviant Behavior. Unpublished dissertation. George Washington University.
- Transparency International. 2015. "Corruption Perceptions Index 2014: Results." Retrieved July 25, 2015, from: <http://www.transparency.org/cpi2014/results>.
- Zigerell, L. J. 2011. "You Wouldn't Like Me When I'm Angry: List Experiment Misreporting." *Social Science Quarterly* 92(2): 552-562.

Table 1			
Results for the 3,000-Person List Experiment from September 2013			
	Point Estimate	P-value	95% Confidence Interval
<i>Unweighted results</i>			
"I cast a ballot under a name that was not my own."	-1%	0.766	[-9%, 7%]
"Political candidates or activists offered you money or a gift for your vote."	7%	0.092	[-1%, 16%]
"I read or wrote a text (SMS) message while driving."	22%	<0.001	[15%, 30%]
"I was abducted by extraterrestrials (aliens from another planet)."	6%	0.082	[0%, 13%]
<i>Weighted results</i>			
"I cast a ballot under a name that was not my own."	-10%	0.088	[-22%, 1%]
"Political candidates or activists offered you money or a gift for your vote."	5%	0.387	[-7%, 18%]
"I read or wrote a text (SMS) message while driving."	23%	<0.001	[11%, 36%]
"I was abducted by extraterrestrials (aliens from another planet)."	6%	0.247	[-4%, 16%]

Data source: Ahlquist et al. (2015).