Quarterly Journal of Political Science, 2007, 2: 325–343

The Exaggerated Effects of Advertising on Turnout: The Dangers of Self-Reports*

Lynn Vavreck

University of California, Los Angeles, 4289 Bunche Hall, Box 951472, UCLA, Los Angeles, CA 90095-1472, USA

ABSTRACT

Political Scientists routinely rely on self-reports when investigating the effects of political stimuli on behavior. An example of this is found in the American politics work addressing whether campaign advertising mobilizes voters. Findings appear to vary by methodology and are based on varying degrees of self-reports; yet, little attention is paid to the furtive complications that arise when self-reports are used as both dependent and independent variables. In this paper, I demonstrate and attempt to account for the correlated yet unobservable errors that drive self-reports of advertising exposure and political behavior. The results are from a randomized survey experiment involving approximately 1500 respondents. Before the 2002 elections, I showed a professionally developed, non-partisan, get-out-the-vote advertisement to a random subset of a randomly drawn national sample via televisions in their own homes. The analysis shows a great divide between the true effect (using assigned treatment and validated vote) and results using respondent recall of these activities.

* This project was funded by the Center for Investigation and Research on Civic Learning and Education (CIRCLE) at the University of Maryland and Yale University's Institution of Social and Policy Studies. The Dean of Social Sciences at UCLA provided additional support. I thank the Ad Council and Polimetrix, Inc. for donating their professional services. I appreciate the advice of Don Green, Doug Rivers, and John Zaller, each were critical to the completion of the project. Additional thanks go to Steve Ansolabehere, Alan Gerber, Jon Krosnick, Greg Huber, Markus Prior, Mike Thies, and John Geer who provided useful feedback on various versions of the manuscript. I am particularly grateful to Jennifer A. Blake who looked up every respondent in this study on state voter files to ascertain vote history. Finally, I thank the editors for their suggestions and patience. This project was approved by the UCLA Institutional Review Board Committee on Human Subjects, #G02-10-060-01.

ISSN 1554-0626; DOI 10.1561/100.00006005

How does campaign advertising affect political preferences and behavior? The importance of this question to research on public opinion and voting behavior is central to an understanding of how citizens respond to campaigns. Often, however, we obtain answers to this question by using a dependent and at least one independent variable derived from self-reports of respondents' behavior. This approach is problematic because of the potentially correlated yet often unobservable errors that drive self-reporting on both measures. This errors-in-variables complication provides a significant challenge to inference, since it is difficult to know the direction of the bias without knowing something about the covariances among the substantive variables and their errors (Madalla 1992, p. 470). In this paper, I use a survey experiment to show the dangers of relying on self-reports of independent and dependent variables simultaneously. The effect of advertising is over-estimated by more than 600% when self-reports of both advertising exposure and turnout in elections are used (a true 1 point effect is estimated at more than 7 points). The use of one or more self-reports dramatically alters the conclusions one draws from this analysis.

The same data are used in all of the analyses presented here; what changes are the operationalizations of advertising exposure and turnout in the election. The paper proceeds as follows: first, I show that the estimated effect of reported exposure to ads on reported turnout (analysis 1) is more than an order of magnitude larger than the effect of actual exposure to the ad on reported turnout (analysis 2). Then I show that the true effect of seeing the ad on validated turnout is 14 times smaller (analysis 3) than when both self-reports are used. By sorting the data on treatment condition, I go on to show that respondents in the control condition who misreport exposure to the advertisement are driving the exaggerated result. Finally, I demonstrate and account for the correlated errors between these two self-reports and bring the survey results in line with the truth.

The study of political advertising effectiveness has focused mainly on whether ads affect political participation. Despite initial experimental work (Ansolabehere and Iyengar 1995), most research in this area is conducted observationally. Results from observational data in some cases suggest campaign advertising has small to absent effects (Finkel and Geer 1998, Johnston *et al.* 2004) and in other cases, evidence suggests large effects (Freedman and Goldstein 1999, Wattenberg and Brians 1999, Freedman *et al.* 2004, Hillygus 2005). In contrast, experimental work routinely suggests controlled exposure to advertising can affect behavior in modest manners (Ansolabehere and Iyengar 1995). Oddly, only one attempt has been made to reconcile these divergent findings into an accumulated sense of knowledge (Ansolabehere *et al.* 1999) by comparing the effects of actual tone of advertising and respondents' recall of this tone. If errors associated with recall of campaign exposure and turnout are correlated it is not surprising that observational work on advertising effectiveness yields results that cover a wide range (0–14 percentage points), while experimental work routinely shows modest effects (1–2 percentage points).

It comes as no surprise to survey researchers that self-reports of behavior tend to be over-reports. Scholars have tried creatively to control for over-reports by altering the context of the questionnaire or making the interview less personal for respondents

(Parry and Crossley 1950, Sudman and Bradbrun 1974, Locander *et al.* 1976, Rogers 1976, Dillman 1978, Katosh and Traugott 1981, Abramson and Clagget 1986, Silver *et al.* 1986). Despite these efforts, people continue to exaggerate when reporting behavior in surveys. Zaller (2002, p. 311) reports that while roughly 30% of Americans report watching the national news every night of the week, Nielsen ratings estimate the combined daily audience for the major broadcast news networks, plus CNN, to be roughly 20 million people. In a country of about 225 million adults, a 30% incidence rate yields 67 million viewers per night — more than three times as many as Nielsen reports. Similar patterns are found for news magazines as well. *Newsweek* and *Time* reach 5.5% and 9% of the population each week, respectively, but surveys routinely overestimate subscription rates by an additional 5–6 percentage points. These results are robust to various survey vendors, sampling methods, and interviewing mediums including the Internet (Krosnick and Rivers 2005, p. 15). The measures on which we most rely and the behaviors in which we are most interested may have correlated errors that result in biased estimates of effects.

In order to shed light on whether the errors associated with recall measures of various political behaviors are correlated, a method needs to chart the real values and the reported values by respondents. One method that asks respondents to report on themselves, can generalize to a large population, and allows the researcher control over actual treatment is the field experiment. A field test aimed at uncovering the weakness in self-reports of behavior could proceed in many ways. One of the most affordable ways to move into the field is to use web-based survey technology. Web-based field experiments have a number of appealing features: respondents do not interact with an interviewer and they are treated on their own televisions, in their own living rooms, with their own family members around them, and in the context of whatever other political information exists in their life at the time (and whatever else might vie for their attention). The effects of treatments that are uncovered from this kind of fieldwork are the marginal effects of each additional treatment in the real-world context of a respondent's political information and environment.

To execute this field test I partnered with the Ad Council, whose award winning public service campaigns include Smokey the Bear, McGruff the Crime Dog, and the Drinking and Driving Can Kill a Friendship campaign. To boost turnout in the 2002 midterm elections, the Ad Council joined with the Federal Voting Assistance Program, and WestWayne Communications of Atlanta, to create "Decision Guy." Decision Guy is a fictional character at the center of their voting campaign. By annoyingly interfering with the real-life decisions young people encounter every day Decision Guy offers a comical portrayal of what can happen when young people let others make decisions for them.

The study is designed to measure the effect of exposure to one 30-second "Decision Guy" ad on voter turnout among young people. Knowledge Networks (KN), a survey research firm using Web-TV to conduct surveys on randomly selected samples, administered the experiment. A total of 3076 citizens aged 18–34 were invited to participate in the experiment. KN used random digit dialing to select households for inclusion in their panel. At the time of this study, they provided selected households with an interactive

television device to connect them to the Internet.[1] Participants in this experiment were randomly selected from the nation as a whole, excluding those who lived in the nation's top 30 media markets.[2] The sample was randomly divided into treatment and control groups. Each group was contacted twice — once before the 2002 midterm elections and once after.

The first wave survey was administered approximately 2 weeks before the 2002 midterm elections, from October 23 to November 5. The experimental group saw three ad clips made by the Ad Council, while the control group saw none.[3] The second wave of the survey was fielded immediately following the election to panel members who had successfully completed the first wave and remained in the field for two weeks. There were no videos shown in the second wave, both groups received the same follow up survey.

Respondents assigned to the treatment group saw 1 minute of video (three advertisements) during the beginning portion of the first wave survey. The first two ads were about responsible car ownership and encouraged people to check their tire pressure routinely and get regular tune ups. Both ads linked these services to cleaner air and were made by the Ad Council. The final ad was the 30-second Ad Council Decision Guy spot called "Fast Food." Fast Food is shot in black and white at a drive through restaurant speakerphone. In the ad, a young man pulls up to the drive through speaker in his car and begins to order. In the middle of placing the order, Decision Guy jumps out from the bushes behind the speaker and starts changing the order. Although the customer objects, Decision Guy does not give up and keeps insisting the customer wants a "#2 with a chocolate shake" instead of a "#5 with lemonade." The tag-line reads: "Stinks when people make decisions for you, huh? That is what happens when you do not vote."

A distinctive feature of this project is that it contains an actual experimental treatment and a survey of the same subjects. This allows the analysis to be done in two different manners: the experimental way, which uses actual treatment and actual turnout as reported by the Secretaries of States; and the observational way, which relies on respondents' recall of whether they saw the ad and whether they voted. Comparing these two sets of results will shed light on whether recall measures are adequate proxies for actual experience and behavior.

In an effort to get the most accurate assessment of recall, the Web-TV technology is leveraged to its fullest capabilities. Previous studies typically use written descriptions of the ads in survey questions, allow respondents to describe ads they recall seeing in open-end formats, or inquire about television shows the respondent watched to infer ad

---

[1] Information about acceptance, response, and attrition rates for the entire KN panel are presented in Dennis (2001) and Krosnick and Chang (2001). The average length of time in the KN panel for the invited respondents in this study was 18 months; for those who opted into the study, the average length of time in the panel was 18.5 months. Information about completion rates is presented in Table A.1 in Appendix A.

[2] Because the Ad Council ad appeared on broadcast television in the nation's top 30 media markets, respondents who lived in these areas were excluded from the sampling frame. This is an attempt to limit the chances of respondents seeing the advertisement on television as well as in the experiment.

[3] In an ideal design, the control group would also be exposed to advertising, although about something other than politics. This design was prohibitively expensive.

Table 1. Recall of advertising images in treatment and control groups

	Wave 1		Wave 2	
	Control	Treatment	Control	Treatment
Turnout image 1	4.72 (615)	85.3 (709)	16.0 (614)	72.7 (707)
Turnout image 2	3.73 (616)	78.4 (708)	7.46 (617)	65.7 (706)
EPA image 1	4.70 (617)	80.1 (709)	25.0 (615)	75.4 (706)
EPA image 2	10.5 (618)	79.8 (708)	21.2 (617)	71.3 (706)
Decoy image 1	*	*	3.40 (617)	7.21 (707)
Decoy image 2	*	*	0.810 (614)	3.96 (707)

Cell entries are percentages of respondents who recall seeing the image in each row (N's in parentheses).

exposure. Because the Web-TV technology is capable of showing still images to respondents, frames from the actual ads (and some fake ads) can be shown and respondents can answer questions about whether they recognize the images. The ability to show the respondent an actual image from the ad proves quite effective.

At the end of the first and second wave surveys, all respondents saw four still images from the ads, two from "Fast Food" and one each from the other ads with questions about recalling the images. In the post-election wave, respondents also saw two decoy images, which no one in the sample could have seen. Table 1 presents the results of the recall analysis.

Between 78% and 85% of the people in the treatment group remember seeing the images in the advertisements immediately after seeing them (at the end of the first wave survey). These findings are higher than those reported by Ansolabehere *et al.* (1999, p. 901) who report about 50% of people in the treatment condition recall ads they saw. When interviewed 2–3 weeks later there is some memory loss, but not much. Between 65% and 75% of the treatment group remember seeing the images.[4] For each of the still images from the Decision Guy ad, roughly 12% of the respondents forgot they had seen the images over the length of the panel, which was about 1 month. Decoy images of children playing at the beach and a woman climbing a volcano show fewer than 8% of the respondents mistakenly report familiarity with these images.[5] Interestingly, respondents in the treatment group were twice as likely to make this mistake as respondents in the control group. Generally speaking, most people do a good job of recognizing still images they have seen in advertisements. This method, in which three-quarters of the people correctly remember what they saw nearly 1 month later, compares favorably to asking

[4] A possible explanation for the rise in control group recall in the second wave is that some people in the control group may remember seeing the still images in the first wave of the survey. If this is happening, the effect would be an attenuation of the treatment effect, as people who report seeing the ad actually only remember the image from wave 1. Since I know who actually saw the ad, however, I estimate that 5%–10% of the respondents are mistakenly reporting recall in this manner.

[5] Ansolabehere *et al.* (1999, p. 901–2) report 4% of respondents in the control group falsely state they had seen a political ad when asked to list the ads they recall seeing.

respondents to list the ads they recall seeing, a method by which half the people cannot remember what they have seen 30 minutes after seeing it (Ansolabehere *et al.* 1999, p. 901).

ANALYSIS 1: SELF-REPORTS OF BOTH EXPOSURE AND TURNOUT

Observational researchers typically ask people whether they recall seeing ads or images and whether they voted on election day — an analysis with two self-reports. This type of analysis results in a robust, positive effect of advertising on turnout. The first section of Table 2 reveals the percent of respondents who report voting in the election controlling for whether they report seeing images from the turnout ads.

People who recall the ad's images voted in the election more often than those who do not remember the advertisement (by 7.32 percentage points). Using conventional levels of significance, the null hypothesis of independence is rejected. The conclusion from analyzing the data in this way is that reported exposure to one additional PSA about voting increases turnout among young people in midterm elections by more than 7 points.[6] It is reasonable to infer that increasing the frequency of exposure to Get Out

Table 2. The effects of reported advertising exposure on reported turnout by condition

	Definitely did not see	Not sure	Definitely saw	Total
Entire sample				
Did not vote	70.6	61.6	63.2	65.0
Voted	29.4	38.4	36.8	35.0
Total	100 (394)	100 (315)	100 (612)	100 (1321)
	χ^2 is 7.8	$p \leq 0.02$		
Experimental group				
Did not vote	71.3	65.1	63.6	64.8
Voted	28.4	34.9	36.4	35.2
Total	100 (87)	100 (106)	100 (514)	100 (707)
	χ^2 is 1.91	$p \leq 0.384$		
Control group				
Did not vote	70.4	59.8	61.2	65.3
Voted	29.6	40.2	38.8	34.7
Total	100 (307)	100 (209)	100 (98)	100 (614)
	χ^2 is 6.97	$p \leq 0.031$		

[6] If I substitute recall of the other PSAs for recall of the turnout advertisements, the effect is not distinguishable from zero. This is true for both treatment and control groups. This suggests that people are actually remembering some of the content of the ads as well as the images from the ad.

the Vote (GOTV) ads would boost turnout even higher. Conclusion: advertising affects turnout.

ANALYSIS 2: ACTUAL EXPOSURE INSTEAD OF SELF-REPORT (WITH REPORTED TURNOUT)

What happens to the effectiveness of advertising when actual assignment to the treatment condition is used in place of respondent recall of the ad? This is the kind of analysis scholars engaged in lab experiments perform, and in existing literature, lab work generates results of smaller magnitude than analyses using two self-reports. These data demonstrate this pattern nicely. By changing only the method of analysis and not the underlying data, the effect of advertising diminishes to 0.53 percentage points (from 7.32) and the null hypothesis of independence cannot be rejected (see Table A.3 in Appendix A). Using the same data, the overall conclusion is fundamentally altered when self-reports of exposure are replaced with actual treatment. The estimate based on two self-reports (exposure and turnout) is 1200% greater than the estimate based on a single self-report (turnout) and leads us to an opposite conclusion from the initial analysis with two self-reports. Conclusion: advertising does not affect turnout.

ANALYSIS 3: ACTUAL EXPOSURE AND TURNOUT INSTEAD OF BOTH SELF-REPORTS

How close does the analysis with one self-report come to estimating the true effect of the treatment? In the interest of fully elucidating the dangers of relying on self-reports, turnout records from secretaries of states offices were appended to these data.[7] Respondents in this sample reported voting in the election of 2002 at a rate of 35%. When validated voting information is used, the actual turnout rate among young people in the survey is 12%.[8]

Substituting actual turnout for reported turnout (still using actual assignment to treatment) moves the estimate of ad effectiveness from 0.53 to nearly one point. If the analysis of these data is done observationally, using two self-reports, an erroneous inference will be made suggesting that advertising markedly increases turnout. This survey result exaggerates the effect by more than 600%. If the analysis is done only relying on voters' reports of their turnout in the election, the correct inference will be

---

[7] This work was done through cooperation between KN and Polimetrix, Inc. Validated turnout information was available for two-thirds of the sample. Of the two-thirds of respondents living in states with current voter files, nearly half of the individuals were found on their states' voter rolls as registered voters. This number is consistent with U.S. Census Data on registration by age group (http://www.census.gov/population/www/socdemo/voting).

[8] The Current Population Study and the Center for Investigation and Research about Civic Education (CIRCLE) ascertain political participation in a similar manner and find over-reporting of about the same magnitude for this demographic group.

made, but the estimate of ad effectiveness is attenuated to half its true size. Conclusion: exposure to a single GOTV advertisement does not increase turnout.

ANALYSIS 4: SELF-REPORTS OF EXPOSURE AND TURNOUT BY TREATMENT CONDITION

The seven point effect generated using ad recall instead of actual treatment may not be that surprising if those who pay more attention and remember the images from ads are more likely to be affected by them. Memory may be a necessary part of the mechanism at work. It is possible that in observational work the effects of advertisements are being conflated with whether a respondent remembers the advertisement. These two separate and separable concepts can be more fully explored by comparing the effect of memory of the ad in the treatment group to the effect of "memory of the ad" in the control group. For example, even though most people (see Table 1) correctly remember whether they saw an advertisement, the effect of remembering the ad should be negligible for those respondents who in fact did not see it (the control group). To the extent that people who did not see the ad report recall, they are not being influenced by the advertisement, but by something else.[9] The bottom two panels of Table 2 present the analysis of advertising recall on reported turnout for the treatment and the control groups, separately.

In the treatment group a familiar pattern emerges. Respondents who recall seeing the ad vote in larger numbers than those who do not recall the advertisement, by about 7.64 percentage points. The cross-tab for the control group ought to reveal an absence of this relationship, since these people *were never exposed to the advertisement*. The relationship between advertising recall and turnout, however, is not only present in the control group, it is larger (9.14 points) — and quite precise despite the drop in cases. Among people who were *never exposed to the advertisement*, those who report exposure to the ad vote in greater numbers than those who deny exposure.

This finding complicates the story about what is actually being measured by a self-report of advertising exposure in surveys. It seems it is *not* the case that recall measures memories of the treatment instead of actual treatment, because some people who report remembering the ads were not treated at all (roughly 17%). In truth, the recall measure is imprecise for two reasons: it not only misses people who were actually exposed but do not remember the treatment (12%), it also includes people who were never exposed but mistakenly report exposure (17%). It is this latter group that clarifies the absence of a reasonable mechanism at work in the investigation of advertising effectiveness using self-reports. The substantial positive result of advertising recall on self-reported turnout is an artifact of the survey method, not a function of memory or accessibility. These data suggest that some respondents want to appear to be the kind of

[9] There is nothing explicitly or implicitly political in the still images shown to respondents, therefore, merely seeing the still image of a young man at a drive-through restaurant window should not be enough to increase either one's actual political participation or one's reporting of such activities.

person who pays attention to politics or votes in elections and perhaps even subscribes to *Newsweek* and *Time* magazines, or watches the nightly network news, even when they have not.[10]

In 1999, Ansolabehere, Iyengar, and Simon critiqued the use of advertising recall as endogenous to turnout in elections. The dangers, however, are worse than they anticipated. Recall of political ads and turnout in elections are related, as the treatment group results demonstrate; but, recall of political ads and turnout, even when one has not seen the ad, are also related, as the control group results show. There are correlated and unobservable errors driving self-reporting of these two variables that yield significantly different results than the actual effects. The 7.32 point increase in turnout is being *equally driven by the control group as well as the treatment group*. These results suggest that the treatment actually has *no effect* on turnout, not a 7.32 point effect. In a typical survey, we do not get to peak at the process that generates the data. These unique data provide clues that the biases associated with over-reporting are notable. The dangers of relying on self-reports are furtive, but legitimate.[11]

ANALYSIS 5: ARE THE ERRORS IN SELF-REPORTS CORRELATED?

In a stark illustration of social desirability bias, among young people who reported voting in this election, only 32% actually did. Clearly, people feel pressure to say they voted in elections, even when they have not. Do these same people, however, feel pressure to over-report their exposure to visual stimuli? Eleven percent of the control group over reported both exposure to the ads and voting in the election.

The literature is relatively clear in terms of useful covariates for turnout exaggerations; for example, race, region, age, education, and gender have all been shown to relate to misreports of participation. In Table 3, I report the results of a seemingly unrelated probit analysis for over-reports of both dependent variables.

In addition to the demographic variables mentioned above, I included campaign interest in the over-report model for turnout. In the over-report of exposure model, I added attention to the media, campaign interest, and attention to popular television shows such as *ER*. These data show that the older, more educated, and more politically interested are more likely to over-report political participation. This is consistent with previous

---

[10] One possible explanation for this behavior is social desirability bias, which contends that respondents' answers are conditioned by social norms about what is "good" or "right." Another plausible mechanism is the idea that respondents want to please the interviewer by giving the answers that affirm what the interviewer is asking ("oh, yes, I saw that," or "yes, I voted").

[11] Observational data are rarely analyzed without controls. In the online Appendix A, I present results for the self-reported treatment effect using typical controls: age, education, perceptions of party differences, and campaign interest (see Table A.4). The control group continues to exhibit precise effects from recall of the ads while treatment group effects are noisy. As controls are added, the positive result in the entire sample disappears; however, it remains robust in the control group. If more respondents mistakenly reported seeing these ads, falsely positive results would emerge in the full sample, even with controls.

Table 3. Predictions of over-reports of advertising exposure and turnout

Variable	Coefficient model 1	S.E.	Coefficient model 2	S.E.
Over-report exposure				
Age (18–30)	−0.004	0.02	−0.008	0.02
Male	0.153	0.144	0.122	0.155
Education (4 levels)	−0.067	0.08	−0.098	0.086
Campaign interest (0–3)	0.106	0.104	0.066	0.111
Media attention (0–3)	0.233	0.163	0.133	0.179
Watch popular TV shows	−0.12	0.095	−0.128	0.101
Duty (5 levels)	*	*	−0.184	0.092
Efficacy (5 levels)	*	*	0.031	0.08
Partisan strength (0-1)	*	*	0.106	0.189
Constant	0.374	0.545	1.09	0.69
Over-report turnout				
Age (18–30)	0.049	0.024	0.058	0.026
Male	0.093	0.168	0.051	0.188
Education (4 levels)	0.265	0.095	0.176	0.104
Campaign interest (0–3)	0.601	0.128	0.561	0.143
Non-white	0.143	0.211	0.086	0.247
South	0.142	0.169	0.121	0.186
Duty (5 levels)	*	*	−0.285	0.117
Efficacy (5 levels)	*	*	0.183	0.095
Partisan strength (0–1)	*	*	0.311	0.206
Constant	−3.56	0.64	−3.45	0.826
	$N = 322$		$N = 291$	

Cell entries are probit coefficients and standard errors generated using seemingly unrelated probit models.

Model 1: Correlation between errors in two equations: 0.183 Breusch–Pagan test statistics = $2.88 \sim \chi^2(1),\ p \leq 0.09$.

Model 2: Correlation between errors in two equations: 0.134 Breusch–Pagan test statistics = $1.26 \sim \chi^2(1),\ p \leq 0.263$.

findings and is explained by the notion that older, more educated citizens appreciate the expectation of civic duty and feel some pressure to conform to it, as do those interested in politics. The model of exposure over-reporting performs less well. Paying more attention to the media may increase a person's probability of exaggerating their exposure slightly, but none of the other variables are good predictors. The question remains: do the errors in these models contain systematic elements that are correlated with one another? The correlation between the residuals from the seemingly unrelated probit analysis is 0.18. A test that the errors are truly independent of one another suggests

rejecting the null hypothesis of independence with a slightly increased risk of Type I error ($\sim\chi^2(1) = 2.88$, $p \leq 0.09$).[12]

There may be other sets of variables that we do not typically use in campaign research that can sever the correlation between the errors. For example, in 1986, Silver, Anderson, and Abramson detailed a set of variables that predict over-reporting on turnout extremely well. These variables include measures of political efficacy, civic duty, partisan strength, and interest in politics. Increases in these things lead to increases in exaggerating political participation, presumably because more dutiful citizens, and people who believe they can influence government, are more likely to feel pressure to participate in politics — even if they have not really done so. With the exception of political interest, most researchers do not include these variables in analyses of campaign effects, public opinion, or vote choice.[13] When these variables are added the results of the seemingly unrelated probit analysis change measurably, particularly in the turnout model.

The results of these new analyses indicate that the more people believe it is their duty to become informed about politics, the more likely they are to exaggerate their participation. Further, the more people believe they can influence government decisions, the more likely they are to exaggerate their voting record. If they are stronger partisans, older, more educated, or more interested in campaigns — they are more likely to over-report political participation. These results are presented in the last two columns of Table 3. It is *not* the least educated, younger, less interested voters who pretend they have voted; it is the people who care about whether they actually vote who misreport their commitment when they do not make it to the polls. Even the exposure model fares better with the inclusion of Silver, Anderson, and Abramson's variables. An increased sense of civic duty leads to over-reports on exposure to advertisements. The inclusion of these variables drives the correlation between the errors down and the test of independence can no longer be easily rejected ($\sim\chi^2(1) = 1.26$, $p \leq 0.26$). By controlling for these different covariates, I am able to sever the correlation between the errors in these two over-reporting models.

One last analysis remains in order to show that observational tests using self-reports on both the independent and dependent sides of the equation can yield accurate results. I re-estimate the original observational model including the control variables that broke the correlation between the errors in both over-reporting models. When these covariates (civic duty, efficacy, and partisan strength) are added to the regression of self-reports of exposure to the campaign ad on self-reported turnout, the coefficient on reported treatment is indistinguishable from zero for the entire sample, and for *both the treatment and control groups.* As Table 4 makes clear, the observational results and the experimental results have converged.

Allowing respondents to report on themselves is easy, but it comes with analytic, computational, and inferential costs for which we rarely account. The correlated,

---

[12] Likelihood-ratio test of $\rho = 0$.

[13] Although see Finkel and Geer (1998) who include interest, partisan strength, and efficacy (but not civic duty). This may be why Finkel and Geer's findings on attack advertising demobilization are attenuated compared to other observational findings.

Table 4. The effects of reported ad exposure on reported turnout with additional controls

Variable	Full	Treatment	Control
Recall exposure? (−1, 0, 1)	0.009 (0.014)	−0.024 (0.025)	0.023 (0.024)
Education	0.068 (0.014)	0.075 (0.02)	0.063 (0.021)
Age (18–30)	0.006 (0.004)	0.004 (0.005)	0.008 (0.005)
Party differences	0.017 (0.016)	−0.0002 (0.021)	0.045 (0.024)
Campaign interest	0.153 (0.015)	0.147 (0.021)	0.163 (0.022)
Civic duty	−0.04 (0.015)	−0.05 (0.021)	−0.024 (0.022)
Efficacy	0.006 (0.012)	0.003 (0.017)	0.009 (0.017)
Partisan strength	0.149 (0.031)	0.186 (0.043)	0.102 (0.044)
Constant	−0.182 (0.107)	−0.058 (0.15)	−0.335 (0.155)
R^2	0.265	0.257	0.285
N	1163	623	540

unobservable errors in self-reports lead to results from survey data that are vastly different (by 600%) from the real processes that generate the data. With the right controls, the systematic portion of the correlation between these reporting errors can be accounted for and the bias can be reduced. Finding the right controls, however, is a challenge when the investigator cannot compare estimation results to the truth. In this case, identifying controls that eliminate the bias in the regression results is possible because I can compare the self-reported treatment effect to the true treatment effect. In a truly observational setting, this is impossible.

Exposure to one 30-second get-out-the-vote ad may modestly increase turnout among young people. While this design is not powerful enough to pinpoint the effect with certainty, it is not hard to imagine a design that might uncover effects of this magnitude. If these data are an indication, the size of the effect, whether interpreted as large or small, is likely to be much smaller than reported by survey data. The large advertising effects recently demonstrated with observational data are likely an artifact of the survey method and the modeling challenges associated with the observational work.

REFERENCES

Abramson, P. R. and W. Claggett. 1986. "Race-Related Differences in Self-Reported and Validated Turnout." *Journal of Politics* 48: 412–422.

Ansolabehere, S. D., S. Iyengar and A. Simon. 1999. "Replicating Experiments Using Aggregate and Survey Data: The Case of Negative Advertising and Turnout." *American Political Science Review* 93(4): 901–909.

Ansolabehere, S. D. and S. Iyengar. 1995. *Going Negative: How Attack Ads Shrink and Polarize the Electorate*. New York: Free Press.

Dennis, J. M. 2001. "Are Internet Panels Creating Professional Respondents? The Benefits from Online Panels Far Outweigh the Potential for Panel Effects." *Marketing Research*, Summer: 34–38.

Dillman, D., 1978. *Mail and Telephone Surveys*. New York: Wiley.

Finkel, S. E. and J. G. Geer. 1998. "Spot Check: Casting Doubt on the Demobilizing Effect of Attack Advertising." *American Journal of Political Science* 42(2): 573–595.
Freedman, P. and K. Goldstein. 1999. "Measuring Media Exposure and the Effects of Negative Campaign Ads." *American Journal of Political Science* 43(4): 1189–1208.
Freedman, P., M. Franz and K. Goldstein. 2004. "Campaign Advertising and Democratic Citizenship." *American Journal of Political Science* 48(4): 723.
Hillygus, D. S. 2005. "Campaign Effects and the Dynamics of Turnout Intention in Election 2000," *Journal of Politics* 67(1): 50–68.
Johnston, R., M. Hagen and K. Hall Jamieson. 2004. *Dynamics of the 2000 Presidential Campaign: Evidence from the Annenberg Survey*. New York: Cambridge University Press.
Katosh, J. P. and M. W. Traugott. 1981. "The Consequences of Validated and Self-Reported Voting Measures. *Public Opinion Quarterly*, 45: 519.
Krosnick, J. and D. Rivers. 2005. "Comparing the Results of Probability and Non-Probability Sample Surveys." Paper Presented at 2005 American Association of Public Opinion Research conference, Miami, Florida.
Krosnick, J. and L. Chang. 2001. "A Comparison of Random Digit Dialing Telephone Survey Methodology as Implemented by Knowledge Networks and Harris Interactive." Unpublished Manuscript.
Locander, W., S. Sudman and N. Bradburn. 1976. "An Investigation of Interview Method, Threat and Response Distortion." *Journal of the American Statistical Association* 71: 269–275.
Madalla, G. S. 1992. *Introduction to Econometrics*, 2nd edn. New York: Macmillan.
Parry, H. J. and H. M. Crossley. 1950. "Validity of Responses to Survey Questions." *Public Opinion Quarterly* 14: 61.
Rogers, T. F. 1976. "Interviews by Telephone and in Person: Quality of Responses and Field Performance." *Public Opinion Quarterly* 40: 51.
Silver, B. D., B. A. Anderson and P. R. Abramsom. 1986. "Who Overreports Voting?" *American Political Science Review* 80(2): 613.
Sudman, S. and N. M. Bradburn. 1974. *Response Effects in Surveys: A Review and Syntehsis*. Chicago: Aldine.
Wattenberg, M. P. and Craig L. Brians. 1999. "Negative Campaign Advertising: Demobilizer or Mobilizer?" *American Political Science Review* 93(4): 891–899.
Zaller, J. R. 2002. "The Statistical Power of Election Studies to Detect Media Exposure Effects in Political Campaigns." *Electoral Studies* 21: 297–329.

APPENDIX

Table A.1. Assignment of persons to experiment and control conditions

Group	Number fielded	Number completed	Rate (%)
Wave 1			
Control	1318	688	52
Experiment	1758	891	51
Total	3076	1579	51
Wave 2			
Control	681	619	91
Experiment	761[a]	710	93
Total	1442	1329	92

Although 891 people were successfully assigned to the experiment group, only 761 were recontacted in wave 2 of the survey. In the control group, 681 of 688 were recontacted. Of the 137 people who dropped out of the experiment, 14 left the Knowledge Networks panel in general, and were thus no longer available to be contacted for the follow up survey. Seven people were in the control group and 7 in the treatment group. Knowledge Networks has studied this phenomenon and can provide evidence to show that people who leave the general panel do so randomly. The other 123 people who were not recontacted in wave 2 experienced technical problems downloading the video portion of the experiment to their Web-TV unit in wave 1, and consequently never viewed the video. All of these people were in the experimental group. It would be an obvious violation of randomization to assign these people to the control group. Since they did not get the treatment, however, they cannot be analyzed as part of the experiment group either. There is nothing systematic about these 123 people that would bias the analysis of the experiment in any way. They are missing from the experimental group at random due to a technical difficulty that had nothing to do with their ability to use the Web-TV box or their familiarity with playing videos. Eliminating them from the experiment does not affect the analysis in any systematic way. I tested the randomness of their dropping out using demographic and validated vote information.

Table A.2. Distributions of variable means and standard deviations

Variable	Full sample	Treatment	Control	All invitations	Treatment invitations	Control invitations
Treatment	0.53 (0.50)	1 (0)	0 (0)			
Validate vote	0.12 (0.32)	0.11 (0.32)	0.12 (0.32)			
Campaign interest	1.44 (1.25)	1.37 (1.01)	1.52 (1.03)			
Campaign attention	1.25 (0.87)	1.22 (0.87)	1.29 (0.88)			
Party differences	1.96 (0.88)	1.94 (0.91)	1.99 (0.85)			
Validated registration	0.46 (0.50)	0.45 (0.50)	0.46 (0.50)			
Education	4.06 (1.5)	4.09 (1.51)	4.01 (1.48)	3.97 (1.5)	4.00 (1.47)	3.94 (1.44)
Gender	0.47 (0.50)	0.44 (0.50)	0.50 (0.50)	0.50 (0.50)	0.48 (0.50)	0.51 (0.50)
Age	25.10 (3.64)	25.20 (3.57)	24.99 (3.72)	24.8 (3.68)	24.8 (3.61)	24.8 (3.78)
South	0.36 (0.48)	0.33 (0.47)	0.38 (0.49)	0.36 (0.48)	0.35 (0.48)	0.39 (0.49)
Efficacy	2.94 (1.06)	2.92 (1.06)	2.97 (1.07)			
Duty	2.62 (0.91)	2.61 (0.89)	2.62 (0.93)			
Party strength	0.22 (0.41)	0.21 (0.41)	0.22 (0.42)			
Married	0.35 (0.48)	0.35 (0.48)	0.34 (0.47)	0.34 (0.48)	0.34 (0.47)	0.36 (0.48)
Attention to popular TV Shows	2.31 (0.78)	2.30 (0.79)	2.32 (0.76)			
Attention to media	0.64 (0.48)	0.63 (0.48)	0.66 (0.48)			

(*Continued*)

Table A.2. (*Continued*)

Variable	Full sample	Treatment	Control	All invitations	Treatment invitations	Control invitations
First wave						
GOTV Image 1	0.11 (0.92)	0.75 (0.63)	−0.63 (0.57)			
GOTV Image 2	0.02 (0.92)	0.66 (0.69)	−0.71 (0.53)			
Get Green 1	0.18 (0.86)	0.68 (0.68)	−0.39 (0.67)			
Get Green 2	0.13 (0.87)	0.69 (0.66)	−0.52 (0.59)			
Second wave						
GOTV Image 1	0.17 (0.86)	0.60 (0.70)	−0.34 (0.74)			
GOTV Image 2	0.03 (0.86)	0.47 (0.79)	−0.49 (0.63)			
Get Green 1	0.22 (0.83)	0.64 (0.67)	−0.11 (0.78)			
Get Green 2	0.29 (0.82)	0.58 (0.71)	−0.19 (0.76)			
Self-report turnout	0.35 (0.48)	0.35 (0.47)	0.35 (48)			
Volcano	−0.56 (0.60)	−0.51 (0.63)	−0.62 (0.55)			
Kids on beach	−0.75 (0.48)	−0.73 (0.53)	−0.78 (0.43)			
Total *N*	1328	709	619	3076	1758	1318

Table A.3. The effects of actual treatment on reported turnout, entire sample

	Control	Treatment
Did not vote	65.27	64.74
Voted	34.74	35.26
Total	100 (619)	100 (709)

$\chi^2 = 0.04,\ p \leq 0.8.$

Table A.4. Self-reported exposure on self-reported turnout by condition

Variable	Full	Treatment	Control
Bivariate Model			
Recall exposure? (−1, 0, 1)	0.034 (0.010)	0.034 (0.025)	0.058 (0.025)
N	1321	707	614
With demographics			
Recall exposure? (−1, 0, 1)	0.020 (0.010)	0.002 (0.025)	0.060 (0.030)
Education (0–9)	0.076 (0.009)	0.078 (0.012)	0.076 (0.010)
Age (18–30)	0.01 (0.004)	0.009 (0.005)	0.01 (0.005)
N	1321	707	614
With demographics & other controls			
Recall exposure? (−1, 0, 1)	0.016 (0.013)	−0.020 (0.020)	0.039 (0.020)
Education (0–9)	0.045 (0.008)	0.050 (0.010)	0.040 (0.010)
Age (18–30)	0.007 (0.003)	0.005 (0.004)	0.01 (0.004)
Party differences (0–3)	0.045 (0.014)	0.034 (0.020)	0.064 (0.020)
Campaign interest (0–3)	0.174 (0.010)	0.170 (0.010)	0.180 (0.010)
N	1303	699	604

Cell entries are linear regression coefficients with standard errors in parentheses.

Table A.5. The effects of actual treatment on validated turnout

	Control	Treatment
Did not vote	69.80	68.89
Voted	30.20	31.11
Total	100 (202)	100 (225)

$\chi^2 = 0.04,\ p \leq 0.8.$

List of question wordings for survey experiment

1. Registration — Sometimes things happen and people do not get around to registering to vote. Would records show that you are …

   — Currently registered to vote
   — Not registered to vote
   — Not sure

2. Party differences — Do you think there is much difference between the Democratic and Republican parties on major policy issues?

   — A lot of difference
   — Some difference
   — Not much difference
   — No difference

3. Attention — Generally speaking, how much attention would you say you pay to political campaigns?

   — A lot
   — Some
   — Not much
   — None

4. Interest — How interested are you in political elections in general?

   — Very interested
   — Somewhat interested
   — Not much interested
   — Not at all interested

5. Television — In general, how often do you watch the following television shows?
   The West Wing Friends ER CSI Bachelor

   — Every week
   — Occasionally
   — Never

6. Image recall — Now we are going to show you some images from different advertisements, please tell us if these are images you have seen before.

   — Definitely saw it
   — Not sure
   — No, have not seen this before

7. Recall turnout	People are often busy and unable to vote in midterm elections. Did you vote in the midterm elections held on November 5? — Definitely voted — Did not vote
8. Duty	Which, if any, of the following factors played a role in helping you decide whether to vote? [A general sense of civic duty] — Played a role — Did not play a role
9. Efficacy	Do you agree or disagree with the following statements? [Someone like me cannot really influence government decisions]. — Strongly agree — Agree — Neither agree nor disagree — Disagree — Strongly disagree