

THE DANGERS OF SELF-REPORTS OF POLITICAL BEHAVIOR: OBSERVATIONAL V. EXPERIMENTAL EVIDENCE

Lynn Vavreck,
University of California at Los Angeles¹

4289 Bunche Hall
Box 951472
UCLA
Los Angeles CA 90095-1472

310-825-4855
lvavreck@ucla.edu

¹ 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. Additional support came from UCLA's Dean of Social Sciences. I thank the Ad Council and Polimetrix for generously donating their professional products and services; and Knowledge Networks for administering the experiment. I am grateful further for the heightened attention to this work by Donald P. Green, Douglas Rivers, and John Zaller, each of whom were critical to the implementation and completion of the project. Additional thanks go to Steve Ansolabehere, Alan Gerber, Jon Krosnick, Greg Huber, Markus Prior, and John Geer who provided excellent feedback. I am particularly grateful to Jennifer A. Blake who diligently and carefully looked up every respondent in this study on state voter files to ascertain vote history. This project was approved by the UCLA Institutional Review Board Committee on Human Subjects, # G02-10-060-01.

Abstract:

Political Scientists routinely rely on self-reports when investigating effects of political stimuli on behavior. An example of this is the work in American politics addressing whether campaign advertising mobilizes voters. Findings 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 account for the correlated yet unobservable errors that drive self-reports of campaign exposure and political behavior. I present results of a unique, randomized survey experiment involving approximately 1,500 respondents. Before the 2002 elections, I showed a professionally developed, non-partisan, get-out-the-vote advertisement to a random sub-set of a randomly drawn national sample via televisions in their own homes. The analysis shows a great divide between actual results (using known treatment and validated vote) and results using respondent recall of these activities.

Why do people vote for incumbent Members of Congress? Does attack advertising demobilize the electorate? Do the media affect political preferences and priorities? These are some of the major questions in American politics driving research through the last several decades. The answers to these questions are compelling and provocative, and add to a general understanding of American political behavior. In many cases, however, the answers are derived from a dependent and at least one independent variable made up of self reports of respondents' experiences.

What are the costs of relying on more than one self-report of important political experiences? The main problems are the potentially correlated yet often unobservable errors that could drive both the reporting on the independent variable and the dependent variable. This kind of errors-in-variables situation is challenging because it is difficult to know even the direction of the bias without knowing something about the covariances among the variables measured with error and their errors (Maddala 1992, p. 470).² Despite this potential bias, self reports of numerous political behaviors are often used in the same explanatory model.

In this paper, I provide evidence from a uniquely designed survey experiment that shows the dangers of relying on self reports of both independent and dependent variables, separately and together. When self reports of exposure and turnout are used, the effect of campaign advertising treatments is to increase turnout by nearly eight percentage points (with a 2.5 point standard error). When actual measures of advertising exposure and turnout are used, the effects diminish to one point (with a 1.5 point standard error). The use of self reports dramatically alters the conclusions one draws from this analysis. The bias, however, is systematic and can be accounted for using unconventional controls.

I. The Problem: Self-reports may affect the Results

² Maddala warns that it is not just errors in the independent variables that create estimation problems.

“What is important to note,” he writes, “is that one can get either underestimation or overestimation when errors of observation are mutually correlated and also correlated with the systematic parts of the model (1992, p. 470).”

Political Scientists routinely ask people to report on themselves and use these measures to illustrate important political mechanisms (for example Cain, Ferejohn, and Fiorina 1987; Verba, Brady, and Scholzman 1996; Rosenstone and Hansen 2002; Finkel and Geer 1998; Freedman and Goldstein 1999; Hagen, Johnston, and Jamieson 2003; Hillygus and Jackman 2003). This is common in areas ranging from studies of basic political participation, to the broad effects of political communication, to specific studies of the effectiveness of district work on Congressional re-election. One topic that has received a lot of attention lately is the study of campaign effects. This is no doubt spurred on mainly by the entrance of impressive large scale data collections such as John Geer's campaign advertising content analysis (Finkel and Geer 1998), Ken Goldstein's Campaign Media Analysis Group tracking data (Freedman and Goldstein 1999), the Annenberg National Election Study (Hagen, Johnston, and Jamieson 2003), and the Knowledge Networks Election Panel (Hillygus and Jackman 2004). These industrious and creative data projects have undoubtedly advanced the study of campaigns and elections. But even with these highly sophisticated and expensive data collection efforts, the question of what kinds of campaign treatments respondents are exposed to is answered mainly with self-reports.³

The most widely read recent debate among scholars of elections is the dispute over whether attack advertising causes voters to stay home on Election Day. Ansolabehere, Iyengar, Simon, and Valentino (1994), Ansolabehere and Iyengar (1995), Finkel and Geer (1998), Freedman and Goldstein (1999), Kahn and Kenney (1999), Wattenberg and Briens (1999), Ansolabehere, Iyengar, and Simon (1999), Vavreck (2000) and many others (see Lau, Sigelman, Heldman and Babbitt (1999) for a meta-analysis of this line of inquiry) have used various measures of exposure to campaign advertising to estimate effects on self reported political participation. In Ansolabehere, Iyengar, Simon, and Valentino (1994) and Ansolabehere and Iyengar (1995) exposure to treatment is strictly controlled in a lab experimental setting, but turnout is reported by the experimental subjects. In the observational studies, both exposure to campaign

³ Even Geer's detailed content analysis and Goldstein's ad-buy data do not reveal any information about whether individuals were actually exposed to these advertisements.

advertising and turnout are measured through respondents' recall. Finkel and Geer (1998) use a measure of how many campaign media the respondent used. The measure is an index comprised of NES questions about whether a person paid any attention to news about the campaign . . . on television, in a newspaper, on the radio, or in a magazine. Freedman and Goldstein (1999) ask respondents about their television viewing habits: "how much television do you watch during [X] part of the day?" Kahn and Kenney (1999) measure exposure to the candidates using the NES questions about whether a respondent saw the candidate on television, read about him or her in a newspaper, and similarly for radio and magazines. Finally, Wattenberg and Briens (1999) use NES open-ended questions asking the respondent whether he or she recalls seeing any presidential campaign advertisements on television, and what they remembered about the ad.

Findings vary systematically by method, yet only one attempt has been made to reconcile these divergent findings into an accumulated sense of knowledge (Ansolabehere, Iyengar, and Simon 1999) by comparing the effects of actual tone of advertising and respondents' recall of advertising tone.⁴ Ansolabehere, Iyengar, and Simon's instrumental variables approach to reconciliation of experimental and observational results was a solid first step toward learning about the biases in self reports of political behavior and exposure. The next logical step is to determine whether self reports of more than one variable are confounding results and what can be done to account for these biases.

II. The Problem: Self Reports are Over Reports

To what extents are self-reports on media habits exaggerations? And, to what extent are they correlated with exaggerations about intended turnout? Because actual subscription and viewership rates of various mass media outlets are publicly available, we can get a good sense of whether people are over-reporting their media habits by

⁴ Ansolbehere, Iyengar, and Simon (1999) account for the problems with recall data by instrumenting for self-reports of exposure to campaign advertising. They also address the simultaneity of exposure recalls and intention to vote.

comparing self reported subscriptions and viewership ratings with the truth, as reported by the media outlets. While roughly 30 percent 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 200 million people, a 30 percent viewership rate yields 60 million viewers per night – three times as many as Nielsen reports (Zaller 2002, p.311). Magazine subscription rates show a similar trend. *Newsweek* reaches 5.5 percent of the population each week, but surveys routinely overestimate its subscription rates by an additional 5 percentage points. *Time* magazine, which reaches 8.8 percent of the population, is overestimated by an even greater amount – nearly 6 points. These results are robust to various survey vendors, sampling methods, and interviewing mediums (Krosnick and Rivers 2005, p. 15).

Over-reports of behavior are commonplace in observational research, and creative attempts to control them range from altering the context of the questionnaire to making the interview less personal for the respondent (Silver, Anderson, Abramson 1986; Sudman and Bradburn 1974; Katosh and Traugott 1981; Parry and Crossley 1950; Dillman 1978; Rogers 1976; Locander, Sudman, and Bradburn 1976; Abramson and Clagget 1986; Bishop, Oldenick and Tuchfarber 1984). Survey researchers offer many solutions that change questions or the questionnaire with hopes of easing the bias in surveys. A lot of attention has been paid to the survey instrument itself.

To compliment work on the survey instrument, economists have given a lot of attention to the context in which the respondent answers survey questions – a hypothetical context, a “what if” context that may bear on the kinds of answers one gives. Hypothetical bias may exist when values that are elicited in a hypothetical context, like a survey, differ from those elicited in a real context, like a market. Results from economics are one-sided: people overstate their true valuation for items in hypothetical settings. Bishop and Heberlein (1986) show this clearly by obtaining hypothetical measures and actual cash values for hunting permits. They find that respondents’ willingness to pay for the permits was much higher in the hypothetical context as compared to the cash market. The average hypothetical values exceeded the real ones by 33 percent in one experiment

and 79 percent in another. As the value of the items for sale becomes less obvious, the biases grow sharply. Neill, Cummings, Ganderton, Harrison, and McGuckin (1994) find that hypothetical bias for a reprint of a medieval map can be as high as 2600 percent. They conclude that the large biases are due to a lack of real economic commitment in the hypothetical setting.⁵

Economists have an advantage here – dollars are easily measured, while it is not always so with political behavior. We have become accustomed to compromises in terms of measuring the things we really care about because they are difficult to measure. The measures on which we most rely and the behaviors in which we are most interested may have correlated errors for which we do not account. Instead of asking someone how much they would pay for a pint of strawberries, and then showing up a few weeks later selling strawberries at their front door to confirm their report, we ask people whether they purchased strawberries in the last month and how much they paid.⁶ Figuring out what kind of person over reports their answers to both questions (if it is systematically possible) is the next step in correcting biases from self reports.

III. The Road to a Solution: A Field Test using a Survey Experiment

The Field: Laboratory Control with External Validity

In order to shed light on whether the errors associated with recall measures of political behaviors are correlated, a method is needed that will chart both the real values of variables and their reported values by respondents. What is the appropriate method for doing so? The simple and appealing essence of laboratory work is the complete control over treatment and the easy interpretation of the results. Experiments require the researcher to make no assumptions about the underlying behavioral model. Results are

⁵ This is consistent with self-reports of valuable behaviors like visits to the dentist, which people over-report dramatically (Campbell, Converse, and Roberts 1976).

⁶ Another famous economic study of hypothetical bias involves the price subjects would pay for a pint of strawberries (Harrison and Rutstrom 2006).

generated using non-parametric analyses that allow the researcher to study effects without relying on many assumptions. The weakness is that researchers have forced compliance with the treatment (in this case, not allowing exposure to vary with relevant covariates) and created such an unusual social experience that results may not generalize beyond the lab.

Observational studies are strong where experiments are vulnerable, but the opposite is also true. Survey results easily generalize to larger populations, but the precision with which survey instruments measure important concepts or behaviors is weak. An additional weakness is that observational work nearly always involves analyses with controls, since treatments are rarely randomly assigned. The assumption of parametric form, along with the decisions about which and how many variables to include as controls is a drawback of observational research. Coefficients on variables of interest often change as controls are added, but it is difficult to know whether the changing magnitudes are due to actual behavioral patterns or the presence of endogenous regressors or variables measured with error.

It would be informative to have a real world calibration at the individual level of observational work on political behavior. A method that asks respondents to report on themselves, can generalize to a large population, and allows the researcher control over actual treatment would meet the criteria for calibration.⁷ Economists and a few political scientists have leveraged field experiments to do these things. Field experiments give scholars the power to randomly treat very large random samples of the population outside the artificial lab setting and to observe the effects outside the lab as well. A field study also provides the mechanism for testing whether results in the lab are different from results outside the lab merely because of the lab setting. 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-

⁷ Ideally, exposure to treatments would be available to people in a randomly assigned treatment group, but the respondents could decide whether to experience the treatment based on whatever criteria they would use to make the same decision in a real political setting.

based interviews can be conducted for no more money than phone surveys, yet can include visual content that was previously only possible with in-person interviewing at a much higher cost.

Web-based field work also provides better external validity than the lab since subjects have no actual contact with the experimenter and there is no chance for the experimenter to inadvertently cue the subjects to desired results. Moreover, subjects 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 respondent's political information and environment. In field tests using the web, researchers can draw random samples from the entire population of American voters and randomly assign some to a treatment group.⁸

Study Design

This project was executed in cooperation 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 partnered 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 – getting their hair cut, choosing which sweater to wear, ordering at a fast food restaurant – "Decision Guy" offers a comical portrayal of what can happen when one lets other people make decisions.

⁸ This can be done using lists or simply by random household assignment.

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, a survey research firm using Web-TV to conduct random sample surveys, was employed to administer the survey experiment. A total of 3,076 citizens age 18-34 were identified for participation in the experiment. Knowledge Networks uses random digit dialing to select households and provides selected households with an interactive television device to connect them to the Internet.⁹ Participants in this experiment were randomly selected from the nation as a whole.¹⁰ The sample was randomly divided into experimental and control groups. Each group was contacted twice – once before the 2002 midterm elections and once after. In Table 1, I present the sample sizes and completion rates for each group in each wave, respectively.

(Table 1 here)

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.¹¹ The second wave of the survey was fielded to panel members who had successfully completed the first wave. There were no videos shown in the second wave, both groups received the same follow up survey.

IV. Effects

⁹ Information about acceptance, response, and attrition rates for the entire Knowledge Networks panel available upon request from the author. See Dennis (2001), Krosnick and Chang (2001).

¹⁰ Because the Ad Council ad appeared in all of the nation’s top 30 media markets, I excluded respondents who lived in these areas from the sample. I did not want participants in the study to see the advertisement on television as well as in the experiment.

¹¹ In an ideal design, the control group would also be exposed to advertising, although about something other than politics. Doing this added considerable expense to the experiment, however, and could not be done at this time.

Recall Measurement

One distinctive feature of this project is its unique design. The project contains an actual experimental treatment and a survey of the same subjects. This allows me to analyze the data 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, even when controls are added.

In an effort to get the most accurate assessment of recall, I leverage the Web-TV technology 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 watches to infer ad exposure. Because the Web-TV technology is capable of showing still images to respondents, I use frames from the actual ads (and some fake ads) and ask respondents whether they recall seeing the images. The ability to show the respondent an actual image from the ad proves quite effective.

The Treatment

Respondents assigned to the experiment group saw one minute of video advertising during the beginning portion of the first wave survey. In the introduction to the survey, respondents were told that I was interested in their use of different forms of media, and their interest in different types of programming by media. In general, they also knew that the survey had some connection to politics and policy.

Respondents in the treatment group first saw a 15 second ad made by the Ad Council for their "Get Green" campaign. The ad is called "Tires" and shows a man filling up his car tires with air. The ad informs people that keeping tires properly inflated will help save the environment and also save them cash. The second 15 second ad was from the same campaign and is called "Tuned." In this ad, a man reminds people that

keeping their cars properly tuned saves on repairs, helps save the environment, and saves them cash. The third ad was the 30 second Ad Council “Decision Guy” spot. It is 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’s what happens when you don’t vote.”

After viewing these three ads, respondents in the experimental group finished the rest of the first wave survey, which was identical to the survey people in the control group received. At the end of the first wave survey, all respondents saw four still images from the ads, two from Fast Food and one each from the Get Green ads with questions about recalling the images. This is the initial test of recall.

After the election, all respondents were re-contacted and administered a post-election survey. This survey was the same for both the experiment and control group. No video was shown in this wave of the survey, but respondents were shown the same four still images from the ads that they saw in the first wave of the survey. This is the second test of recall, weeks later. Respondents were also shown two decoy images, not

from advertising campaigns, that no one in the sample could have seen. Table 2 presents the results of the recall analysis.

(Table 2 about here)

Between 78 and 85 percent 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, Iyengar and Simon (1999, p. 901) who report about 50 percent of people in the treatment condition recalling the ads they saw.¹² When interviewed 2 to 3 weeks later there is some memory loss, but not much. Between 65 and 72 percent of the treatment group remember seeing the images.¹³ For each of the still images from the turnout ad, roughly 12 percent of the respondents forgot they had seen the images over the length of the panel, which was about 1 month. Decoy images (2) of children playing at the beach and a woman climbing a volcano were shown to all respondents in the second wave of the survey to test for baseline false positive results. Fewer than 8 percent of the respondents mistakenly reported familiarity with these images. Interestingly, respondents in the treatment group were twice as likely to make this mistake as respondents in the control group.¹⁴

Generally speaking, most people do a good job of recognizing still images they have seen before in advertisements. This method, when compared to asking respondents to list the ads they recall seeing, does a better job of accurately relaying information about advertising exposure.

¹² Their measure of recall is to ask respondents to list the ads they could remember seeing (30 minutes after seeing them). Only half the participants who saw a campaign ad recalled seeing one. Respondents saw other, non-political types of ads, too.

¹³ The reason control group recall goes up in the second wave is that some people in the control group remember seeing the images in the first wave of the survey during the recall measures.

¹⁴ Ansolabehere, Iyengar, and Simon (1999, p. 901-2) report 4 percent of respondents in the control group falsely state they had seen a political ad when asked to list the ads they recall seeing.

Results: Actual Exposure on Reported Turnout (representative of lab experiments)

In Table 3, I present the percent of respondents who report voting in the 2002 elections controlling for whether the person was assigned to the treatment or control group. This is the kind of analysis typically used in laboratory. These initial results show a modest half-point ($p \leq .8$) increase in reported turnout among people in the treatment condition. The magnitude of this effect is small – and worse, the number of cases makes it impossible to know whether this half-point increase is different from zero. This result is moderately consistent with laboratory results showing modest 2.5 point effects on turnout from advertising (Ansolabehere and Iyengar 1995).

(Table 3 Here)

Results: Reported Exposure on Reported Turnout (representative of observational studies)

The robust advertising effects found in some observational work (for example Freedman, Franz, and Goldstein 1999), however, come from using self-reports on both the right and left side of the equation. What happens to the modest results reported in Table 3 if self reports of exposure to the treatment are used instead of actual assignment to the treatment condition? In the first section of Table 4, I present the percent of respondents who report voting in the election controlling for whether they report seeing images from the turnout ads.

(Table 4 here)

People who say they recall the ad's images voted in the election more often than those who do not remember the advertisement. The difference, 7.32 percentage points, is both substantively and statistically significant. It is a 14 times greater than the half point increase that results from the analysis of actual exposure to the ad on reported turnout and is comparable to recent and previously documented advertising effects using observational work.

Recalling the advertisement boosts the effectiveness of the treatment substantially.¹⁵ This is a plausible result since those who pay more attention and remember the images from the ad may be more likely to be affected by it. If you remember the ad, you are more likely to vote in the election. The change in advertising effectiveness resulting from the use of memory instead of actual exposure suggests that in observational work, the effects of treatments are being conflated with whether a respondent remembers the treatment. These are two separate and separable concepts.

Because this survey is embedded in an experiment, I can further explore the relationship between recalling the advertisement and voting. For example, even though most people (72 percent, see Table 1) correctly remember whether they saw the 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 “recall” seeing it, they are not being influenced by the advertisement, but by something else.¹⁶ The bottom part of Table 4 presents the analysis of advertising recall on reported turnout for the treatment and the control groups, respectively.

In the treatment group, I find a familiar pattern. Respondents who recall seeing the ad vote in larger numbers than those who do not recall the advertisement, by about 7.6 percentage points. The difference is not statistically significant, but this reflects the total number of cases being cut in half.

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 – and statistically significant despite the drop in cases. Among people who were

¹⁵ Even answering that you are “not sure” whether you have seen the image boosts turnout substantially. Among this set of people, 35 percent were actually in the treatment group.

¹⁶ 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-thru restaurant window should not be enough to increase either one’s actual political participation or one’s reporting of such activities.

never exposed to the advertisement, those who reported exposure to the ad voted in greater numbers than those who denied exposure.

This finding complicates the story about what is actually being measured by a self report of advertising exposure. 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 percent). Recall of the treatment is not entirely driven by the treatment itself. In truth, the recall measure is imprecise for two distinct reasons: it not only misses people who were actually exposed but do not remember the treatment (12 percent), it also includes people who were never exposed but mistakenly report exposure (17 percent).

This finding sheds doubt on the use of self reported measures of behavior. The positive result is an artifact of the survey method, not unlike the over valuation economists find in their survey experiments. Just as a high valuation seems to mean very little to subjects because they know they will never actually have to pay it, over-reports of exposure and behavior come easily to voters. These data suggest that many people want to appear to be the kind of 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).¹⁷

In 1999, Ansolabehere, Iyengar, and Simon critiqued the use of advertising recall as endogenous to turnout in elections. It seems, however, that the dangers 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 8 point increase in turnout is

¹⁷ In addition to economists' theory of hypothetical bias, another possible explanation for this behavior is social desirability bias, by which 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").

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 an 8 point effect. In a typical survey, we do not get to “peak” at the mechanisms that generate 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.

In most cases survey data are not analyzed using bivariate cross tabs. Observational researchers typically introduce controls through the use of parametric models. The goal is to add variables to the model that affect the dependent variable and are also correlated with the treatment. Since most observational work is non-experimental in design, respondents self-select into treatments. By adding controls to the models, the true effect of the treatment may be revealed, in theory. The problem is that there are many reasons that the coefficients on treatments may change, and without the benefit of knowing who was actually treated, a survey researcher is left to assume that changing coefficients are due to improvements in model specification, not other confounding factors such as collinearity or endogeneity.

In Table 5, I present the results of regression analyses of these observational data using typical controls. The first row details the simple bivariate analysis of self-reported exposure on self-reported turnout: the full sample analysis shows a treatment effect of nearly 7 percentage points, but it is entirely driven by people in the control group (who show almost a 12 point effect from recall).¹⁸

(Table 5 here)

The addition of demographic controls changes very little. The full sample effect remains, although it is slightly smaller. Effects are absent once more in the actual treatment group and largely present (12 points) in the actual control group. Education

¹⁸ Self reports of exposure are measured in three categories: definitely saw it, not sure, definitely did not see it. For this reason, it is necessary to multiply the coefficient by 2 in order to get the effect of moving from non-exposure to exposure.

and age have the expected effects on turnout in the election – each increases participation by substantial amounts. The constancy of the results is encouraging since neither age nor education can be endogenous to exposure; and it is unlikely that one of these two demographic variables is driving positive recalls of treatment.

Adding substantive control variables, such as interest in the campaign or whether there are any important differences between the two major parties, causes more change. The effect of self reported exposure in the full sample disappears and in the actual treatment group, the coefficient is negative, although not distinguishable from zero. The full sample results behave as anticipated: adding theoretically appropriate controls yields a result that looks like the true effect – a null effect from treatment. As before, however, the actual control group effect remains positive and significant, although it decreases in magnitude to 8 points. If there were more respondents who mistakenly reported seeing these ads, falsely positive results would emerge.

Results: Actual Exposure on Actual Turnout

To explore the possibility of correlated errors between people who say they were treated and people who say they voted in the election, individual level data on actual turnout are needed. I hired Polimetrix, Inc., a survey research firm based in Palo Alto, California to gather turnout data from the Secretaries of States offices in which my survey respondents live. Data of this type are matters of public record, but some states make it more difficult than others to gain access to their voter rolls. I was able to obtain state records to ascertain the participation of 93 percent of the survey respondents.¹⁹

¹⁹ Six states in this experiment proved difficult for Polimetrix to access: Wisconsin, Oregon, Maine, Washington, Vermont, and Massachusetts. In 2003, Polimetrix had a license to use data provided by Aristotle, Inc., a Washington DC based vendor of vote history information. In subsequent years, Polimetrix built their own nationwide voter file. Aristotle's collection of vote history encompassed 93 percent of my sample's geography. There is no reason to believe that the unavailable states are in any way systematically related to the random assignment of individuals to the treatment condition, thus the effect of this missing information should not systematically affect the results. Of the 93 percent with vote history, 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. The Census reports 38

Polimetrix worked directly with Knowledge Networks to append turnout information to my experimental data.

Nationwide turnout among 18-34 year olds in the 2002 midterm election was nearly 22 percent. The range of participation spanned a high of 43 percent among young people in Minnesota to a low of 9 percent participation in Arizona. Respondents in this sample reported voting in the election of 2002 at a rate of 35 percent. When validated voting information is used, the actual turnout rate among respondents in the survey is 12 percent. Many of the respondents who reported voting in the election did not vote.²⁰

What does the use of validated voting data show about the actual effect of the treatment on turnout? The effect of the ad is bolstered by the use of real voting information, moving from a half-point effect to nearly 1 point. Although the effect is doubled, the drop in cases again makes it difficult to precisely estimate the effect.²¹ I present these results in Table 6.

(Table 6 here)

The Solution

In a stark illustration of social desirability or hypothetical bias, among young people who reported voting in the election, only 32 percent actually did. Clearly, people

percent of 18-24 year olds are registered and 55 percent of 25 to 44 year olds are registered. Registration rates go up to 70 percent in subsequent age groups

(<http://www.census.gov/population/www/socdemo/voting>).

²⁰ These over-reports are not unusual. The Current Population Study ascertains political participation in a similar manner and finds over-reporting of about the same magnitude for this demographic group. The Center for Investigation and Research about Civic Education (CIRCLE) also finds similar patterns of over-reporting among young people.

²¹ The drop in cases is mainly the result of voter registration. Of the 1,328 respondents in the study, slightly less than half of them were actually registered to vote.

feel pressure to say they voted in elections, even when they have not. Roughly a quarter of the entire sample mistakenly reported voting. Do these same people, however, feel pressure to over-report their exposure to campaign stimuli? Eleven percent of the control group over reported both exposure to the ads and voting in the election.

What factors explain over-reporting of turnout and exposure? The literature is relatively clear in terms of useful covariates for turnout exaggerations; for example, things like race, region, age, education, and gender have all been shown to relate to whether people misreport their level of participation. Do these same factors explain why people over-report their exposure to certain images? In Table 7, I report the results of a seemingly unrelated regression (SUR) analysis for over-reports of both dependent variables.

(Table 7 here)

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. These data show that the more educated you are, and the more interested you are in campaigns, the more likely you are to over-report your political participation. This is consistent with previous findings and is explained by the notion that 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.

These simple models are able to explain 11 percent of the variation in over-reporting of turnout and 4 percent of the variation in over-reporting of exposure to advertisements. This leaves a significant portion of the variation unexplained. 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 SUR analysis above is nearly .10. 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 ($\sim X^2(1) = 2.8, p \leq .09$).²²

There is something systematic in the error terms of each model, something that I have not accounted for; and, the systematic components of each model's error are related to one another. Including political variables like level of interest in the campaign and attention to the media did not break the correlation between these errors.

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 discovered a set of variables that predicts over-reporting on turnout extremely well. These variables include efficacy, 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.²³ When these variables are added the results of the SUR change measurably, particularly in the turnout model.

(Table 8 here)

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. 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 – the more likely they are to over-

²² Breusch-Pagan test of independence.

²³ 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.

report political participation. It is *not* the least well 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. These variables explain 17 percent of the variation in over-reporting on turnout.

Even the exposure model fares better with the inclusion of Silver, Anderson, and Abramson's variables. An increased sense of duty and increasing age lead to over-reports on exposure to advertisements. The inclusion of these variables drives the correlation between the errors down to .07, and the test of independence can no longer be easily rejected ($\sim\chi^2(1) = 1.3, p \leq .26$). 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 (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*. I present these results in Table 9.

(Table 9 here)

In Table 5, the recall measure of exposure was insignificant for the entire sample, but the effect from recall of exposure in the control group remained positive and significant at 8 points. In Table 9, even the control group shows an effect statistically indistinguishable from zero. Finally, the observational results and the experimental results have converged.

V. Conclusion

Relying on multiple self-reports of behavior can generate misleading results if correlated, unobservable errors are present in the data. This correlation can be minimized

by controlling for covariates that break the systematic correlation between reporting errors in explanatory variables and the outcomes in which we are interested. These correlated, initially unobservable errors may lead to results from survey data that are vastly different (700 percent exaggerated) from the real processes that generated them.

Allowing respondents to report on themselves is easy, but it comes with analytic, computational, and inferential costs for which we rarely account. These results show that even when parameterization and covariates were introduced to the models using observational data on self-reports, the large positive effects of self-reported exposure to campaign advertising on turnout were driven by people in the control group who were exaggerating their exposure. Normal science progress in the area of campaign effects has moved markedly over the last several decades. Even as this knowledge accumulates, we must take time and care to attend to the difficult but consequential complications that arise from studying attitudes, opinions, and behaviors without directly observing them.

Table 1. Assignment of Persons to Experiment and Control Conditions

Group	Number Fielded	Number Completed	Rate
<i>Wave 1</i>			
Control	1318	688	52%
Experiment	1758	891	51%
Total	3076	1579	51%
<i>Wave 2</i>			
Control	681	619	91%
Experiment	761 ²⁴	710	93%
Total	1442	1329	92%

²⁴ 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 seven 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 2. Recall of Advertising Images in Treatment and Control Groups

	Wave 1		Wave 2	
	Control	Treatment	Control	Treatment
First Ad Image	4.06	84.39	15.96	72.70
Second Ad	3.52	78.16	7.46	65.72
Image				
Decoy Image	*	*	3.4	7.2
Total	688	891	619	710

Cell entries are percentages of respondents in each category

Table 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 Square = .04 p<=.8

Table 4. The Effects of Respondent Recall of Exposure on Reported Turnout

	Definitely Did Not See	Not Sure	Definitely Saw	Total
Entire Sample				
Did Not Vote	70.56	61.59	63.24	65.03
Voted	29.44	38.41	36.76	34.97
Total	100 (394)	100 (315)	100 (612)	100 (1321)
	Chi Square 7.8		p<=.02	
Experimental Group				
Did Not Vote	71.26	65.09	63.62	64.78
Voted	28.74	34.91	36.38	35.22
Total	100 (87)	100 (106)	100 (514)	100 (707)
	Chi Square 1.9		p<=.40	
Control Group				
Did Not Vote	70.36	59.81	61.22	65.31
Voted	29.64	40.19	38.78	34.69
Total	100 (307)	100 (209)	100 (98)	100 (614)
	Chi Square 6.9		p <= .03	

**Table 5. Regression Analyses by Treatment Condition
(Self-Reported Exposure on Self-Reported Turnout)**

Bivariate Model	Coefficients:		
	Full	Treat-ment	Control
Recall Exposure? (-1, 0, 1)	.033* (.010)	.033 (.025)	.058* (.025)
N	1321	707	614
With Demographics			
Variable	Full	Treat.	Cont.
Recall Exposure? (-1, 0, 1)	.020* (.010)	.002 (.025)	.060* (.030)
Education (0-9)	.080* (.009)	.070* (.012)	.080* (.010)
Age (18-30)	.009* (.004)	.009 (.005)	.009 (.005)
N	1321	707	614
With Demographics & Other Controls			
Variable	Full	Treat.	Cont.
Recall Exposure? (-1, 0, 1)	.015 (.013)	-.020 (.020)	.040* (.020)
Education (0-9)	.044* (.008)	.050* (.010)	.040* (.010)
Age (18-30)	.007* (.003)	.005 (.004)	.009* (.004)
Party Differences (0-3)	.045* (.014)	.034 (.020)	.060* (.020)
Campaign Interest (0-3)	.170* (.010)	.170* (.010)	.180* (.010)
N	1303	699	604

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

** indicates statistical significance at the $p \leq .05$ level*

Table 6. The Effects of Actual Treatment on Validated Turnout

	Control	Treatment
Did Not Vote	70.12	69.17
Voted	29.88	30.83
Total	100 (251)	100 (266)

Chi Square = .05 p<= .8

Table 7. Predicting Who Will Over-Report Exposure and Turnout

OVER-REPORT EXPOSURE		
Variable	Coefficient	S.E.
25 +	-.08	.06
Male	.06	.06
Education	-.02	.02
Campaign Interest (0-3)	.04	.04
Media Attention (0-3)	.10	.06
Watch Popular TV Shows	-.04	.04
Constant	.67	.13*
N = 323	R ² = .04	
OVER-REPORT TURNOUT		
Variable	Coefficient	S.E.
25+	.07	.05
Male	.02	.04
Education	.04	.01*
Campaign Attention (0-3)	.14	.03*
Race	-.03	.05
South	.04	.04
Constant	-.24	.09*
N = 323	R ² = .11	

Cell entries are linear regression coefficients generated using SUR.

** Indicates statistical significance at the $p \leq .05$ level*

Correlation between errors in two equations: .09
 Breusch-Pagan test statistic = 2.8 \sim Chi²(1), $p \leq .09$

Table 8. Predicting Who Will Over-Report Exposure and Turnout using Psychological Controls

OVER-REPORT EXPOSURE		
Variable	Coefficient	S.E.
25 +	-.10	.06
Male	.04	.06
Education	-.03	.02
Campaign Interest (0-3)	.02	.04
Media Attention (0-3)	.05	.06
Watch Popular TV Shows	-.04	.04
Duty	.07*	.03
Efficacy	.02	.03
Partisan Strength	.05	.07
Constant	.94*	.21
N = 292	R ² = .06	
OVER-REPORT TURNOUT		
Variable	Coefficient	S.E.
25+	.08	.05
Male	.01	.04
Education	.03	.01
Campaign Interest (0-3)	.10*	.03
Race	-.04	.05
South	.03	.04
Duty	.07*	.03
Efficacy	.05*	.03
Partisan Strength	.12*	.05
Constant	-.12	.15
N = 292	R ² = .17	

Cell entries are linear regression coefficients generated using SUR.

** Indicates statistical significance at the $p \leq .05$ level*

Correlation between errors in two equations: .07
 Breusch-Pagan test statistic = 1.3 ($\sim X^2(1)$, $p \leq .26$)

Table 9. Regression Analyses by Treatment Condition using Self-Reported Exposure with Additional Controls

Variable	Full	Treat- ment	Control
Recall Exposure? (-1, 0, 1)	.01 (.01)	-.04 (.03)	.03 (.03)
Education (0-9)	.04* (.01)	.05* (.01)	.03* (.01)
Age (18-30)	.01* (.004)	.01 (.006)	.01 (.006)
Party Differences (0-3)	.03 (.02)	.001 (.03)	.08* (.03)
Campaign Interest (0-3)	.14* (.02)	.14* (.02)	.15* (.03)
Civic Duty	-.03 (.01)	-.04 (.03)	-.02 (.03)
Efficacy	.02 (.01)	.03 (.02)	.01 (.02)
Partisan Strength	.15* (.04)	.2* (.05)	.08 (.05)
Constant	-.35* (.13)	-.30* (.03)	-.43* (.20)
R ²	.26	.26	.26
N	1303	699	604

Appendix A

Table A1. Distributions of Variables

Variable	Mean	Standard Deviation	Min	Max
Treatment	.53	.49	0	1
Validate Vote	.11	.31	0	1
Campaign Interest	1.2	.74	0	2
Campaign Attention	1.2	.87	0	3
Party Differences	2.0	.88	0	3
Validated Registration	.46	.50	0	1
Education	4.1	1.5	1	9
Gender	.53	.49	0	1
Age	25.1	3.6	18	30
South	.44	.49	0	1
Efficacy	3.0	1.5	1	5
Duty	2.6	.9	1	5
Party Strength	.24	.42	0	1
Married	.37	.48	0	1
Attention to Popular TV Shows	1.3	1.0	-2	2
Attention to Media	.64	.48	0	1
First Wave				
GOTV Image 1	.09	.91	-1	1
GOTV Image 2	.02	.92	-1	1
Get Green 1	.12	.86	-1	1
Get Green 2	.29	.82	-1	1
Second Wave				
GOTV Image 1	.17	.86	-1	1
GOTV Image 2	.03	.86	-1	1
Get Green 1	.22	.83	-1	1
Get Green 2	.29	.81	-1	1
Self Report Turnout	.35	.48	-1	1
Volcano	-.56	.60	-1	1
Kids on Beach	-.75	.48	-1	1

Table A2. List of Question Wordings for Survey Experiment

1. Registration	<p>Sometimes things happen and people don't get around to registering to vote. Would records show that you are ...</p> <p>Currently Registered to Vote Not Registered to Vote Not Sure</p>
2. Party Differences	<p>Do you think there is much difference between the Democratic and Republican parties on major policy issues?</p> <p>A lot of Difference Some Difference Not Much Difference No Difference</p>
3. Attention	<p>Generally speaking, how much attention would you say you pay to political campaigns?</p> <p>A lot Some Not Much None</p>
4. Interest	<p>How interested are you in political elections in general?</p> <p>Very Interested Somewhat Interested Not Much Interested Not at all Interested</p>
5. Television	<p>In general, how often do you watch the following television shows?</p> <p>The West Wing Friends ER CSI Bachelor</p> <p>Every Week Occasionally Never</p>
6. Image Recall	<p>Now we are going to show you some images from different advertisements, please tell us if these are images you've seen before.</p> <p>Definitely Saw It</p>

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 can't really influence government decisions].

Strongly Agree
Agree
Neither agree nor disagree
Disagree
Strongly Disagree

- Abramson, Paul R., and William Claggett. 1986. "Race-related differences in self-reported and validated turnout," *Journal of Politics* 48: 412-22.
- Ansolabehere, Stephen D., Shanto Iyengar, Adam Simon, and Nicholas Valentino. 1994. "Does negative advertising demobilize the electorate?," *American Political Science Review* , 1994, 829-38.
- Ansolabehere, Stephen D., Shanto Iyengar, and Adam Simon. 1999. "Replicating Experiments Using Aggregate and Survey Data: The Case of Negative Advertising and Turnout," *American Political Science Review* 93(4):901-9.
- Ansolabehere, Stephen D. and Shanto Iyengar. 1995. *Going Negative: How attack Ads Shrink and Polarize the Electorate*. New York: Free Press.
- Bartels, Larry M. 1993. "Messages Received: The Political Impact of Media Exposure." *American Political Science Review* 87 (2): 267.
- Bishop, Richard and Thomas Heberlein. 1986. "Does Contingent Valuation Work?" in R. Cummings, D. Brookshire & W. Schulze, Eds. *Valuing Environmental Goods: A State of the Arts Assessment of the Contingent Valuation Method*, Totowa, NJ: Rowman & Allenheld.
- Bishop, George F., Robert W. Oldenick, and Alfred Tuchfarber. 1984. "What must my interest in politics be if I just told you 'I don't know'?" *Public Opinion Quarterly* 48:510-19.
- Brennan Center for Justice. 2000. "2000 Presidential Race First in Modern History where Political Parties Spend More on TV Ads than Candidates," web page: http://www.brennancenter.org/presscenter/releases_2000. Release dated December 11, 2000. Accessed August 8, 2003.
- Campbell A, Converse PE, Rogers WL. 1976. "The quality of American life. Perceptions, Evaluations, and Satisfactions." New York: Russell Sage Foundation, 106.
- Cain, Bruce E., John A. Ferejohn, and Morris P. Fiorina. 1987. *The Personal Vote: Constituency Service and Electoral Independence*. Cambridge: Harvard University Press.
- Dennis, J. Michael. 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, Don. 1978. *Mail and Telephone Surveys*. New York: Wiley.

- Finkel, Steven E. and John G. Geer. 1998. "Spot Check: Casting Doubt on the Demobilizing Effect of Attack Advertising," *American Journal of Political Science*, 42 (2): 573-595.
- Freedman, Paul and Ken Goldstein. 1999. "Measuring Media Exposure and the Effects of Negative Campaign Ads," *American Journal of Political Science*, 43 (4): 1189-1208.
- Freedman, Paul, Michael Franz, and Ken Goldstein. 2004. "Campaign Advertising and Democratic Citizenship," *American Journal of Political Science* 48:4, 723.
- Harrison, Glenn W. and E. Elisabet Rutström. 2006. "Experimental Evidence on the Existence of Hypothetical Bias in Value Elicitation Methods," in Charles Plott and V.L. Smith, Eds. *Handbook of Experimental Economics Results*, Burlington, MA: Elsevier.
- Hillygus, D. Sunshine and Simon D. Jackman. 2003. "Voter Decision Making in Election 2000: Campaign Effects, Partisan Activation, and the Clinton Legacy," *American Journal of Political Science*, 47:4, 583-596.
- Johnston, Richard, Michael Hagen, and Kathleen Hall Jamieson. 2004. *Dynamics of the 2000 Presidential Campaign: Evidence from the Annenberg Survey*. New York: Cambridge University Press.
- Kahn, Kim Fridkin and Patrick J. Kenney. 1999. "Do Negative Campaigns Mobilize or Suppress Turnout? Clarifying the Relationship between Negativity and Perception," *American Political Science Review* 93 (4) 877-89.
- Katosh, John P. and Michael W. Traugott. 1981. "The Consequences of Validated and Self-Reported Voting Measures." *Public Opinion Quarterly*, 45:519.
- Krosnick, Jonathan and Douglas 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, Jonathan and L. Chang. 2001. "A Comparison of Random Digit Dialing Telephone Survey Methodology as Implemented by Knowledge Networks and Harris Interactive." Unpublished Manuscript.
- Lau, Richard R. Lee Sigelman, Caroline Heldman, and Paul Babbitt. 1999. "The Effects of Negative Political Advertisements: A Meta-Analytic Assessment," *American Political Science Review* 93:4, 851.

- Locander, William, Seymour Sudman, and Norman Bradburn. 1976. "An Investigation of Interview Method, Threat and Response Distortion," *Journal of the American Statistical Association* 71:269-75.
- Madalla, G.S. 1992. *Introduction to Econometrics*, Second Edition. New York: Macmillan.
- Neill, Helen R., Ronald G. Cummings, Philip T. Ganderton, Glenn W. Harrison, and Thomas McGuckin. "Hypothetical Surveys and Real Economic Commitments," *Land Economics* 70:2, 145-154.
- Parry, Hugh J. and Helen M. Crossley. 1950. "Validity of Responses to Survey Questions," *Public Opinion Quarterly*, 14: 61.
- Rogers, Theresa F. 1976. "Interviews by telephone and in person: Quality of responses and field performance," *Public Opinion Quarterly* 40:51.
- Rosenstone, Stephen J. and John Mark Hansen. 2002. *Mobilization, Participation, and Democracy in America*. New York: Longman.
- Silver, Brian D., Barbara A. Anderson, and Paul R. Abramsom. 1986. "Who Overreports Voting?" *American Political Science Review* 80:2, 613.
- Sudman, Seymour, and Normal M. Bradburn. 1974. *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine.
- Vavreck, Lynn. 2000. "How does it all 'Turnout?' Exposure to Attack Advertising, Campaign Interest, and Attention in American Presidential Elections," in Larry M. Bartels and Lynn Vavreck, eds. *Campaign Reform: Insights and Evidence*, Ann Arbor: University of Michigan Press.
- Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady. 1996. *Voice and Equality: Civic Volunteerism in American Politics*. Cambridge: Harvard University Press.
- Wattenberg, Martin P. and Craig L. Brians. 1999. "Negative Campaign Advertising: Demobilizer or Mobilizer?" *American Political Science Review* 93 (4): 891-99.
- Zaller, John R. 2002. "The Statistical Power of Election Studies to Detect Media Exposure Effects in Political Campaigns," *Electoral Studies* 21: 297-329.