Experiments to Reduce the Over-reporting of Voting: A Pipeline to the Truth

Michael J. Hanmer  
Antoine J. Banks  
University of Maryland  

Ismail K. White  
The Ohio State University

Abstract

Voting is a fundamental part of any democratic society. But survey based measures of voting are problematic because a substantial proportion of nonvoters report that they voted. This over-reporting has consequences for our understanding of voting as well as the behaviors and attitudes associated with voting. Relying on the “bogus pipeline” approach, we investigate whether altering the wording of the turnout question can cause respondents to provide more accurate responses. We attempt to reduce over-reporting simply by changing the wording of the vote question by highlighting to the respondent that: 1) we (survey administrators) know some people who say they voted did not; and 2) that we can in-fact find out, via public records, whether or not they voted. We examine these questions through a survey on U.S. voting age citizens after the 2010 midterm elections. Our evidence shows that these questions improved accuracy of the reports by reducing over-reporting of turnout.

For helpful comments we thank Fred Conrad, Jenna Fulton, Christina Heshmatpour, Karen Kaufmann, Vince Hutchings, Roger Tourangeau, Candace Turrito, Nick Valentino, and Eric Wish and his staff at the Center for Substance Abuse Research (CESAR) at the University of Maryland. This project was funded by Time Sharing Experiments for the Social Sciences (TESS). We also thank the TESS PIs, reviewers, and staff. Banks and Hanmer also received research support from the Dean’s Research Initiative in the College of Behavioral and Social Sciences and the Department of Government and Politics at the University of Maryland.
1 Introduction

Research examining voting behavior via surveys has a severe limitation—some respondents report that they voted when they did not actually vote, a phenomenon referred to as vote over-reporting (see, e.g., Clausen 1968; Traugott and Katosh 1979).\(^1\) Using voter validation data from election officials, Belli, Traugott, and Beckmann (2001) found that the percentage of over-reporters in the American National Election Studies (ANES) typically ranged from 8 to 14 percent. High rates of over-reporting raise serious concerns about the validity of our measure of voting—one of the most basic political behaviors. The purpose of our paper is to offer new survey questions designed to reduce the over-reporting of voting.

Besides finding that vote over-reporting greatly distorts the accuracy of survey estimates of voter turnout, researchers have also found that vote over-reporting is not distributed randomly across the population (see, e.g. Hill and Hurley 1984; Belli et al. 2001; Silver, Anderson, and Abramson 1986; Bernstein, Chadha, and Montjoy 2001; Cassel 2003; Karp and Brockington 2005; Deufel and Kedar 2010; Ansolabehere and Hersh 2012). The unequal distribution of over-reporting has led some researchers to contemplate the value of using survey measures of voter turnout to understand who votes and why (see, e.g., Silver et al. 1986; Bernstein et al. 2001). For example, several studies find substantial differences in the rate of over-reporting among some groups of nonvoters, ranging from 11% to over 50% (Silver et al. 1986; Bernstein et al. 2001). These differences produced biased estimates of the effects key demographic and attitudinal variables have on voting. For example, Bernstein et al. (2001) find that over-reporting exaggerates the estimated effects of education and partisan strength on turnout and attenuates the

\(^1\) Reporting that one did not vote when he/she did (under-reporting) is another source of error. However, the rate of underreporting is so miniscule that researchers give it little, if any, attention (Belli, Traugott, and Beckmann 2001). Additional problems can arise through survey and item non-response (Burden 2000, 2003; McDonald 2003; Martinez 2003).
effect of race and region. As a result, the implications of over-reporting are consequential, as some “distortions can be severe enough to cause researchers to miss support for some hypotheses or falsely accept support for others” (Bernstein et al. 2001, 41). In other words, vote over-reporting creates biased estimates that challenge some of the standard theoretical expectations and empirical regularities in the literature on voting (Silver et al. 1986; Bernstein et al. 2001; Cassel 2003; Deufel and Kedar 2010; Ansolabehere and Hirsch 2012).

Over-reporting on the vote question also causes problems for other related political variables because it operates as a filter question for subsequent behavioral and attitudinal measures. For example, candidate choice questions follow the vote question. Respondents who answer “yes” to the vote question are then asked which candidate they voted for. Thus, inaccurate responses on the voting question can contaminate subsequent vote choice questions (Wright 1993). Moreover, election administration questions, which have become more common since the controversial 2000 presidential election between George W. Bush and Al Gore (see, e.g. Alvarez, Hall, and Llewellyn 2008), are usually administered only to people who reported that they voted. When actual nonvoters answer election administration questions, the consequences of vote over-reporting might extend to real-world policy debates over how to administer proper and fair elections. That is, vote over-reporters may also skew other survey measures such as: confidence that one’s vote was recorded as intended; the length of time one must wait in line to vote; and the helpfulness of poll workers. The collection of misinformation on these questions could potentially generate poor policy decisions about the type of voting machines used, the allocation of voting machines, and poll worker training. Thus, the benefits of

---

2 For a discussion of bias in vote choice among those who reported voting see Carsey and Jackson (2001) and the citations therein.
producing a more accurate measure of voting extend beyond the act of voting to a wide range of scholarly and policy pursuits.

Working from a new theoretical perspective that integrates the literature on voting and the reporting of sensitive behaviors, we propose a new set of measures designed to provide a more accurate assessment of who voted and who did not. Drawing from the “bogus pipeline” approach (see e.g. Jones and Sigall 1971), commonly used in public health and psychology, we investigate whether or not changes in the wording of the turnout question can cause respondents to provide more accurate responses. We discuss the pipeline approach in more detail later but the key feature involves convincing respondents that the researchers can determine whether they are telling the truth.3 Using a survey experiment on a nationally representative sample of voting age citizens, we alter the vote question in two ways. In our treatment conditions we either inform respondents that: 1) we (survey administrators) can in-fact find out, via public records, whether or not they voted; or 2) we know some people who say they voted did not. We also supplemented the survey experiment with a voter validation study to evaluate the survey reports against official records. To the best of our knowledge, we are the first to apply the pipeline approach in political science and on a nationally representative sample. We are also among the first to perform a vote validation study on a nationally representative sample using electronic databases (see Ansolabehere and Hersh 2012). To foreshadow our results, we find that informing respondents that we will check their answers against public records or giving them the more subtle prompt acknowledging the problem of over-reporting significantly reduced over-reporting and improved the accuracy of the responses.

3 The convention in the literature is to refer to the approach as the “bogus pipeline” since much of the work relies on bogus information meant to deceive respondents. Though we use this naming convention here, strictly speaking we use the “actual pipeline” approach since we do not use deception in our research design.
2 Theoretical Framework

2.1 Over-reporting in Political Science

Political scientists have long been concerned with the over-reporting of voting. Since the early 1960s, researchers have identified validity problems associated with measuring voter turnout with self-reported measures (e.g. Clausen 1968). In response to these concerns researchers experimented with the wording of the vote question. Unfortunately, the earliest attempts to reduce over-reporting via changes in question wording failed (see Presser 1990; Abelson, Loftus, and Greenwald 1992; and Belli, Traugott, and Rosenstone 1994). One of the main struggles for researchers has been devising a turnout question that reduces the social desirability pressures that lead to over-reporting (see, e.g. Duff, Hanmer, Park, and White 2007). That is, individuals recognize the importance of voting to citizenship and representative democracy, so when confronted with a turnout question in a survey context, they feel pressure to say they voted even when they did not. Recent research focused on reducing social desirability pressures and/or reducing confusion due to memory failure has been successful in reducing over-reporting (Belli, Traugott, Young, and McGonagle 1999; Belli, Moore, and VanHoewyk 2006; Duff et al. 2007). Several studies have shown that providing answer choices that offer respondents more socially accepted ways to admit nonvoting reduces over-reporting (Belli et al. 2006; Duff et al. 2007). However, despite these improvements, we are still far from accurately measuring turnout with surveys.

---

4 Researchers have also used list experiments to obtain better estimates of sensitive behavior (see e.g. Corstange 2009; Holbrook and Krosnick 2010). Unfortunately, due to the nature of the data generated from list experiments, the ability to draw inferences about who engaged in the behavior of interest, and thus the ability to answer a variety of related research questions, is severely limited (Corstange 2009).
Before discussing how we might move closer to a more accurate vote measure, we offer a new perspective on the various mechanisms that lead to the over-reporting of voting. The pressure to over-report voting has generally been thought of as an external pressure, but we argue that the pressure might be internal as well. Our view derives from the belief that voting is best described as an expressive act. That is, voters obtain some benefit from being a voter (Schuessler 2000). Under this view, people with strong attachments to democracy and/or their political party take pride in being voters and generally think of themselves as voters, even when, from time to time, they do not vote. As Schuessler (2000) argues, citizens view being a voter as an identity held in their own minds as well as an identity they wish to convey to others.

Early studies of vote over-reporting approached the problem as one driven by external pressure—the pressure of the respondent to signal to the interviewer that he/she conformed to the norm of voting (e.g. Silver et al. 1986; Bernstein et al. 2001). For example, Bernstein et al. (2001) argue that pressure to vote leads to feelings of guilt among those who did not vote; they contend that nonvoting respondents “are unwilling to confess to the interviewer…as they wish to avoid the shame they would feel were they to let others know of their failure” (Bernstein et al. 2001, 24-25). Though we do not dispute the presence of external pressure, we contend there is more to the story.

Since expressive motivations can function without being displayed publically (Schuessler 2000), we expect that over-reporting should not be limited to situations in which there is an interviewer. Even when answering a self-administered survey, the internal pressure to report voting may be sufficiently strong to lead individuals to report that they voted when they did not. In fact, Ansolabahere and Hersh (2012) found that over 50% of validated non-voters in the 2008 Cooperative Congressional Election Study (CCES) (a self-administered Internet survey) reported
that they voted. Drawing on identity theory (see e.g. Stryker and Serpe 1994) and insights from survey methodology (see e.g. Schwarz 1996), Brenner (2011) discusses how internal pressures might reduce over-reporting of church attendance. He argues that respondents seeking to preserve their self-conception might interpret a question about a specific behavior to be about a broader identity, such as being someone who values religious attendance. We think this type of pressure should also apply to voting. In addition to interpreting the turnout question as one about behavior in the election in question, respondents may view the question as being a broader one about whether they think of themselves as voters. As a result, respondents may want to think of themselves as civically virtuous people who value participating in democracy’s most fundamental act. Given that most nonvoters in a given election actually voted in other recent elections (see e.g. Berinsky 2005), this phenomenon might be quite prevalent.

We also contend that the logic of expressive voting suggests another mechanism. The act of voting is the primary means by which individuals obtain expressive benefits. For those who think of themselves as voters, failing to vote in an election prevents them from fully attaching themselves to the outcome. For these individuals, participating in the survey gives them an opportunity to claim some of that attachment; that is, by falsely saying that they voted they can express their identity as voters, or as voters for a particular candidate or party. As noted earlier, this sort of expression need not be done in the presence of an interviewer. With the rise of Internet panels consisting of individuals who opt-in to particular surveys, this sort of behavior might become increasingly common. In sum, we contend that the pressure to over-report voting consists both of internal and external pressure. We now turn our attention to an approach that combats this pressure.

Bernstein et al. (2001) also contemplate the possibility that respondents rationalize and interpret the turnout question as being a question about preferences among the candidates.
2.2 The Pipeline Approach

We look outside of political science to find ways to design measures that will improve our ability to reduce over-reporting. Researchers in the fields of public health, psychology, and sociology have for decades dealt with similar issues when studying individual responses to sensitive questions, such as those that have to do with sexual attitudes and behavior and drug use. These researchers have established a set of techniques designed to reduce the influence of social desirability pressures in the context of interviews and surveys dealing with sensitive topics.

Of particular relevance is the “bogus pipeline” approach (see Clark and Tifft 1966; Jones and Sigall 1971; Wish, Sushinsky, Yacoubian, and Fitzgerald 2000; Tourangeau and Yan 2007). Jones and Sigall (1971, 349) begin their seminal article on the bogus pipeline by describing researchers who rely on self-reports as having “fantasies about discovering a direct pipeline to the soul (or some nearby location)” to overcome concerns about the accuracy of self-reported attitudes and behavior. Building on this colorful language, Roese and Jamieson (1993, 363) describe the bogus pipeline method as one “that pierces through the strategic facades and bypasses the concealing words” that survey respondents provide. In studies of sensitive behaviors, such as drug use or smoking, when respondents could be convinced that the researcher would learn the truth about their behavior, such as through a lie detector or a breath sample, respondents were more likely to tell the truth (Tourangeau and Yan 2007). That is, this line of research demonstrated that respondents prefer to come clean on questions subject to social desirability pressures rather than being perceived as, or thinking of themselves as, both a liar and one who engages in a socially undesirable behavior. Researchers have found success with this approach both when they cannot actually determine the truth but deceive the respondents into thinking they can determine the truth, hence the “bogus” label, and when they can actually
determine the truth, known as the “actual pipeline” approach. Importantly, being assigned to a pipeline treatment does not increase nonresponse bias (Tourangeau and Yan 2007).

The pipeline approach is ideally suited to dealing with the vote over-report problem in surveys. If researchers can convince respondents that they can actually verify if the respondent voted this should reduce both the internal and external pressure respondents feel to over-report voting. Since survey responses regarding voting can be checked against official records, researchers can employ the “actual pipeline” approach to reduce the pressures to over-report. That is, researchers can inform respondents in the context of the turnout question that the researcher can and will examine official vote records to verify whether or not they voted. This approach should effectively reduce over-reporting because it offers the non-voting respondent a choice to be seen and to think of themselves in one of two ways, either as a non-voter or a non-voter and a liar, and given that in nearly all situations being a liar represents a greater norm violation than being a non-voter we expect this approach will encourage respondents to tell the truth about whether or not they voted.

2.3 Hypotheses

Given this we offer a set of more specific expectations for how we believe the pipeline approach might be applied to measuring voter turnout and ultimately reducing vote over-reporting. We hypothesize that presenting individuals with the possibility of being discovered to be a liar and thinking of themselves as a liar should reduce nonvoters’ likelihood of over-reporting. More specifically, we believe that by informing respondents, in the context of the turnout question, that we will examine official vote records (which we do) to determine whether or not they voted will reduce over-reporting compared to the traditional question (ANES).
Not all survey researchers are able to validate the vote and some might not want to deceive respondents by telling them they will check the records, when they have no intention of actually checking them. As a consequence, we also manipulate the turnout question in a subtle way—aimed at subtly inducing respondents to think carefully about the truthfulness of their response. That is, we attempt to reduce over-reporting simply by highlighting in the question stem that we know some people who say they voted did not vote. Therefore we test two hypotheses:

H₁: *Telling respondents that we will check their voting records should reduce vote over-reporting relative to the control group (those who receive the standard ANES question) and thus increase the accuracy of vote responses.*

H₂: *Subtly getting respondents to think carefully about the truthfulness of their response should reduce vote over-reporting relative to the control group (those who receive the standard ANES question) and thus increase the accuracy of vote responses.*

3 Experimental Design

Our experimental design randomly assigns subjects to one of three conditions: control condition (standard ANES vote question); actual pipeline condition (checking against public records); and subtle condition (some say they voted when they did not). Using the survey responses and validation results we compare the pipeline condition and the subtle condition to the control condition, to determine the extent to which our vote questions reduce over-reporting and thus improve accuracy relative to the traditional ANES vote question. We present the three different versions of the turnout question below.
ANES Turnout Question (Control Condition)

In talking to people about elections, we often find that a lot of people were not able to vote because they weren’t registered, they were sick, or they just didn’t have time. Which of the following statements best describes you?

Actual Pipeline Turnout Question (Treatment Condition 1)

In talking to people about elections, we often find that a lot of people were not able to vote because they weren’t registered, they were sick, or they just didn’t have time. By looking at public records kept by election officials we can get an accurate report of who actually voted in November, and in previous elections. Of course, these public records do not say who you voted for. Part of our study will involve checking these records against the survey reports. Which of the following statements best describes you?

Subtle Turnout Question (Treatment Condition 2)

In talking to people about elections, we often find that a lot of people were not able to vote because they weren’t registered, they were sick, or they just didn’t have time. We also sometimes find that people who say they voted actually did not vote. Which of the following statements best describes you?

We used the following answer choices for each of the questions: 1) I did not vote (in the election this November); 2) I thought about voting this time but didn’t; 3) I usually vote but didn’t this time; and 4) I am sure I voted.

The experiment was conducted on a representative sample of U.S. citizens of voting age through Knowledge Networks, an Internet survey company, from November 3, 2010 to November 18, 2010. Respondents are matched to the national population on gender, age, race, education, census region, and Internet access. The total sample size was 2,517 U.S. citizens age 18 or older. There was good variation on age, gender, partisanship, and education. The random assignment of subjects to conditions was successful: there were no significant differences across cells of the design in the proportion of respondents across socio-demographic or partisan variables (see Appendix Table 1). As a result, any differences in the post-stimulus dependent measures can be attributed to the manipulation and not to other factors (Kinder and Palfrey 1996).
Catalist provided the vote validation data. Catalist is a private firm that provides voter registration and turnout data from information collected from the states. Ansolabehere and Hersh (2012) provide an extensive review of Catalist’s process and through a clever test demonstrated the high quality of Catalist’s work. In addition to strong performance in Ansolabehere and Hersh’s (2012) test, Catalist came in second place in the 2011 MITRE Multi-Cultural Name Matching Challenge beating a number of leading firms, including IBM. Thus, we are confident that Catalist provides high quality data. Moreover, Catalist did not have access to information on treatment assignment and simply performed the matching using standard background information collected by Knowledge Networks; because subjects are randomly assigned to conditions, any errors in the validation process would be randomly distributed across the three conditions.

4 Results

Before getting to our main results, we address the concern that the pipeline treatment might offend respondents—causing them to drop out of the survey or refuse to answer the turnout question. We do not find any evidence that the pipeline treatment increased either survey nonresponse or item nonresponse. This is consistent with Tourangeau and (Yan 2007).

We now turn to testing our hypotheses. We begin our analysis of over-reporting by focusing on validated nonvoters. As Silver et al. (1986) pointed out, this group is the only group at risk of over-reporting. Moreover, using the percentage of actual nonvoters who reported

---

6 It cleans the data through a process that relies on additional information it regularly collects.
7 Both the number of respondents who dropped out of the survey and the number who refused to answer the voting question were comparable across all 3 conditions. With respect to drop outs, 1 individual in the control condition dropped out (0.12%) compared to 3 in the pipeline treatment (0.35%) and 4 in the subtle treatment (0.48%). And with respect to item non-response 1 respondent refused to answer the control question (0.12%), 1 refused to answer the actual pipeline question (0.12%), and 0 refused to answer the subtle question (0%).
voting as the base, allows us to avoid problems with the marginal distribution of voters and nonvoters in the sample (Silver et al. 1986). Figure 1 shows the percent of overreporters among validated nonvoters across our three conditions. The results are simple: the mean proportion of cases in the respective category across conditions weighted using the standard Knowledge Networks weight. As the figure shows, both the pipeline condition and the subtle condition substantially reduced the rate of over-reporting. While 24.8% of validated nonvoters overreported in the control condition (ANES vote question), just over 20% did so in the subtle condition (some say they voted when they did not), and only 17.2% did so in the pipeline condition (we will check the records). The 7.6 percentage point decrease in the rate of over-reporting in the pipeline condition is statistically significant (p = 0.014, 2 tailed) and substantively significant, representing a 31% decline in over-reporting. The decrease in the rate of over-reporting is especially noteworthy given that few studies have reduced over-reporting via question wording changes. Though the difference between the control and the subtle condition is substantively meaningful (a 4.8 point decrease), it is not statistically significant at conventional levels (p = 0.123, 2 tailed).

Important: our results for the pipeline condition hold even when we limit the analysis just to those who are actively registered to vote. Berent, Krosnick, and Lupia (2011) raise concerns with the quality of voter registration records noting that much of the difficulty with voter validation involves locating individuals on the registration records. When we look just at those individuals that Catalist could identify as registered, we find that 20.5% of the validated nonvoters in the control condition over-reported while only 12.9% of the validated nonvoters in the pipeline condition over-reported voting—a difference of 7.6 percentage points (p = 0.076, 2
tailed). In other words, our results are not influenced by the assumption that the failure to locate an individual’s registration record indicates that the individual did not vote.

Moving now to the second portion of H1 and of H2, Figure 2 shows the accuracy with which each of the turnout measures captured the respondents’ actual turnout. For example, in the control condition, 85.5% of the respondents gave answers that matched up with what we found in the voting records. Contrasting this with the subtle condition where 87.1% of responses matched up with the voting records, we see that there is a 1.6 percentage point increase in the accuracy of the responses in the subtle condition compared to those reported the control condition. Although this 1.6 percentage point increase in accuracy is suggestive of the fact that the subtle measure may be more accurate than the control, the difference is not statistically significant at conventional levels of statistical significance (p = 0.323, 2 tailed). The pipeline condition, on the other hand, performs significantly better than the traditional ANES question. Comparing the pipeline question to the traditional ANES question, we see that accuracy increases by 4.1 percentage points. In the pipeline condition, 89.6% of responses matched up with what we found in the voting records. The difference between the pipeline and control conditions is statistically significant by conventional standards (p=.012, 2 tailed).

The results above show strong support for H1 and suggestive support for H2. We now turn to the percentage of over-reporters by condition among all respondents. Figure 3 presents these results. The figure shows that the subtle condition and the pipeline condition reduce over-reporting relative to the control condition. That is, 8% of respondents over-reported in the pipeline condition while 9.4% did so in the subtle condition and 11.8% did so in the control condition. The 3.8 percentage point differences between the pipeline condition and the control condition is statistically significant at p = 0.008 (2 tailed) while the 2.4 percentage point
difference between the subtle condition is not statistically significant at conventional levels (p = 0.104 2 tailed). Overall, the reduction in over-reporting in the at-risk group (validated nonvoters) was large enough to meaningfully reduce the rate of over-reporting across the entire sample, especially in the pipeline condition.

Although our primary goal was to assess performance of the new measures when asking about the most recent election, we worded the pipeline question to signal that the records contained information on previous elections as well. As we noted above, the main addition to the turnout question in the pipeline condition is the following sentence: “By looking at public records kept by election officials we can get an accurate report of who actually voted in November, and in previous elections.” Since over-reporting is also an issue for recall of turnout from past election cycles (Weir 1975), we hypothesized that our new question would reduce over-reporting when asking about the 2008 election. We asked all respondents the same question about voting in the 2008 presidential election. That is, we simply asked everyone the standard Knowledge Networks core question on previous voting: “Did you happen to vote in the 2008 presidential election?” giving respondents the chance to say either “Yes” or “No.” Given the historic nature of the 2008 presidential election, we also expected that the pressure to over-report would be particularly high. That is, Obama supporters will have wanted to associate with having played a role in his election and his detractors will have wanted to associate with having tried to prevent his victory.

Figure 4 shows the rate of over-reporting by condition among those who were age 20 or older in 2010, and thus eligible to vote in 2008, and who were validated as nonvoters in 2008 (i.e. the at risk group). As expected, the overall rate of over-reporting was higher than in 2010, with 46.1% of validated nonvoters in the control group reporting that they voted. Also as
expected, those in the pipeline condition were substantially less likely to over-report (a difference of 7.1 points) than those in the control condition, though the result will be viewed by most as statistically insignificant (p = 0.106, 2 tailed). The fact that the pipeline question spills over to recall of more distant elections is encouraging and calls for additional research with larger sample sizes to improve precision and repeated use of the pipeline approach when asking about previous election cycles.

5 Conclusion

How might we reduce vote over-reporting? We offer a new (methodological?) approach that reduces people’s internal pressure to say they voted when they actually did not. To combat this problem, we developed new vote questions that derive from a long line of research in psychology - the pipeline approach. Our study is the first application of the pipeline approach in political science and the first application via a representative survey of adult American citizens.

The findings show that by either informing respondents that we will check their answers against public records or giving them a more subtle prompt acknowledging the problem of over-reporting significantly reduces vote over-reporting in surveys. Our results place the questions among the small group that have been successful in reducing over-reporting. Importantly, all of our respondents were given the “excuse” answer choices which have been shown to reduce over-reporting so our results; thus, the reductions in over-reporting we found were beyond whatever effects derive from these answer choices. While we certainly have not solved the over-reporting problem, our new questions will provide researchers interested in voting behavior with the opportunity for more accurate estimates of turnout and consequently more accurate tests of their theoretical expectations.
As a practical matter, these questions are easy to implement and need not significantly impact the cost of administering the survey instrument. Although the vote validation suggested in the pipeline condition can be a costly endeavor, one of the pipeline approach’s strengths is that the researcher does not actually have to do the validation. However, for researchers concerned about deception, randomly sampling respondents to be validated is consistent with the language we use and will give each respondent an equal chance of having their vote validated and reduce financial costs.

With the financial cost of validation coming down with the sophistication of electronic databases we encourage researchers to add validation efforts to their survey based projects. We recognize that this is not always possible and that doing so can cause significant delays between the administration of the survey and the submission of the research for peer review, as the validation process depends on the availability of the voting records from official sources. Thus, we also encourage scholars to continue designing survey questions to reduce vote over-reporting.
References


Figure 1. Percentage of Over-reporters Among Validated Nonvoters in 2010 by Experimental Conditions

![Bar chart showing percentage of over-reporters among validated nonvoters. The conditions are Control, Subtle Treatment, and Pipeline Treatment (Check Records). The percentages are 24.81%, 20.05%, and 17.20% respectively.]

Notes:

N = 1,028
Control vs. Subtle: p-value (2 tailed) = 0.123
Control vs. Pipeline: p-value (2 tailed) = 0.014
Figure 2. Overall Accuracy in 2010 Across the Sample by Experimental Conditions

<table>
<thead>
<tr>
<th>Experimental Condition</th>
<th>% Accurate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>85.53%</td>
</tr>
<tr>
<td>Subtle Treatment</td>
<td>87.14%</td>
</tr>
<tr>
<td>Pipeline Treatment (Check Records)</td>
<td>89.61%</td>
</tr>
</tbody>
</table>

Notes:

N = 2,515  
Control vs. Subtle: p-value (2 tailed) = 0.323  
Control vs. Pipeline: p-value (2 tailed) = 0.012
Figure 3. Percentage of Over-reporters in 2010 Across the Sample by Experimental Conditions

Notes:

N = 2,515
Control vs. Subtle: p-value (2 tailed) = 0.104
Control vs. Pipeline: p-value (2 tailed) = 0.008
Figure 4. Percentage of Over-reporters Among Validated Nonvoters in 2008 by Experimental Conditions

Notes:

N = 737
Excludes those who were 18 or 19 years old in 2010.
Control vs. Subtle: p-value (2 tailed) = 0.646
Control vs. Pipeline: p-value (2 tailed) = 0.106
Appendix Table 1. Distribution of Characteristics Across Experimental Conditions

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Condition A</th>
<th>Condition B</th>
<th>Condition C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>20-25</td>
<td>25-30</td>
<td>30-35</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Education</td>
<td>High School</td>
<td>College</td>
<td>University</td>
</tr>
<tr>
<td>Income</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
</tbody>
</table>