Voter Turnout and the National Election Studies

Barry C. Burden
Harvard University, Department of Government, Littauer Center 228
North Yard, Cambridge, MA 02138-3001
e-mail: burden@fas.harvard.edu, http://www.fas.harvard.edu/~burden

Though the overreporting of voter turnout in the National Election Study (NES) is widely known, this article shows that the problem has become increasingly severe. The gap between NES and official estimates of presidential election turnout has more than doubled in a nearly linear fashion, from 11 points in 1952 to 24 points in 1996. This occurred because official voter turnout fell steadily from 1960 onward, while NES turnout did not. In contrast, the bias in House election turnout is always smaller and has increased only marginally. Using simple bivariate statistics, I find that worsening presidential turnout estimates are the result mostly of declining response rates rather than instrumentation, question wording changes, or other factors. As more peripheral voters have eluded interviewers in recent years, the sample became more saturated with self-reported voters, thus inflating reported turnout.

1 Introduction

VOTER PARTICIPATION IS an appropriately vital area of political science research. Absolute and relative voter turnout levels have important normative consequences for the political system and affect how campaigns are run, which candidates are elected, and how responsive the leaders are to their constituents. Turnout is the most common and probably the most important component of an individual's participation in the political process. Researchers rely on sample surveys to question citizens about their participation in recent elections. Indeed, turnout is one of a few items included in every National Election Study (NES) since its inception half a century ago.[1] Despite their promise, surveys such as the NES routinely overestimate national voter turnout by worrisome amounts.[2]

Author's note: An earlier version of this paper was presented at the 1999 meeting of the American Political Science Association, Atlanta, GA. NES data were provided by the ICPSR, which bears no responsibility for my analyses or interpretations. I thank John Brehm, Kit Kenny, Gary King, Mike McDonald, and Herb Weisberg for their comments.

[1] NES is the contemporary name for the series of surveys begun at the University of Michigan in the late 1940s. Before 1978, it was referred to as the Center for Political Studies (CPS), Survey Research Center (SRC), or just Michigan study. For simplicity, I use the term NES to refer to the entire series of CPS/SRC/Michigan/NES surveys. I focus on the NES instead of other surveys for two reasons. First, the NES is used more than any other survey for studying U.S. voter turnout and it remains the best source of American political survey data. The NES web site currently lists over 3000 items in "The NES Bibliography." Second, no other continuing survey has generated a time series as long as the NES. For example, the General Social Survey (GSS) turnout question covers only half as many presidential elections as the NES does.

[2] Frequent misreporting led frustrated NES researchers to implement "voter validation" techniques in several of its studies though it was unsuccessful for at least three reasons. First, it does not comprise a long enough time series

Copyright 2000 by the Society for Political Methodology

The overreport bias has already been documented and researchers have done well in identifying the sources of the problem in particular surveys. High turnout estimates are due to such things as the undersampling of nonvoters, the mobilizing effects of preelection surveys on respondents, and infamous misreports (Anderson and Silver 1986; Belli et al. 1999; Campbell et al. 1960; Clausen 1969; Silver et al. 1986; Traugott and Katosh 1979). What has gone undocumented is the *growing severity* of this bias.

In 1952, the NES overestimated official voter turnout in the presidential election by 11 percentage points. This is a serious error, but it has more than doubled since then. The 1996 NES turnout estimate was almost 24 points higher than official voter turnout. Indeed, the bias grew rather steadily over this 45-year time period with just a few exceptions. However, I demonstrate in this article that the turnout bias is not ubiquitous; it has worsened noticeably only in presidential elections. The modest amount of bias in congressional turnout has not changed much over time, particularly in midterm elections. I find that the worsening NES estimates are due largely to declining response rates. Other factors that explain the overreport such as question wording and instrumentation have not changed, but successful interview rates have.

2 Official Turnout and the NES

It is common knowledge that the NES estimates are substantially higher than official government estimates of voter turnout reported by the media just after each election.[3] In a search for explanations, researchers focus on the fact that turnout is assessed by asking respondents for *self-reports* of their own election behavior, allowing many nonvoters to report that they voted. However large the role of this particular factor, the fact that a continuing bias exists at all should raise concerns for researchers using NES data.

A common defense is that the bias, while problematic cross-sectionally, has not changed over time. For example, any faulty inferences made about turnout using the 1964 NES will be similar in the 1984 NES, guaranteeing temporal comparability despite uncertain substantive conclusions. Election-specific influences such as candidate characteristics will not persist across surveys, the argument goes, but more general relationships such as that between education and voting should be affected in the same way by the bias regardless of which election year is analyzed.

Beginning with the earliest analyses, researchers find a consistent NES bias over time. Chronologically, the bias has been reported as 12 (Campbell et al. 1960), 12 to 13 (Clausen 1969), 14 to 15 (Shaffer 1981), 15 (Abramson and Aldrich 1982; Teixeira 1987), and 10 to 12 (Teixeira 1992) points. Further, self-reported turnout tracks official turnout closely. Shaffer (1981, p. 71) finds that "the magnitude of turnout inflation has changed little over time" and that "the decline in reported turnout . . . reflects the actual turnout decline." Abramson and Aldrich (1982, p. 503) argue that "in spite of their systematic overreporting of turnout,"

for longitudinal analysis. Second, less than a third of published studies rely on available validated data (Bianco and Poggione 1997). Finally, many researchers find that replacing self-reported voter turnout with validated voter turnout in multivariate analyses does not change the substantive conclusions reached (e.g., Hill and Hurley 1984; Katosh and Traugott 1981; Sigelman 1982).

[3] Official voter turnout is defined as the number of votes cast in a race divided by the Voting Age Population (VAP), as estimated by the Census Bureau. The VAP is problematic because it includes people like aliens and felons who are sometimes denied suffrage and excludes people like citizens living abroad and people uncounted by the Census who are eligible to vote (Burnham 1986; Clausen 1969; McDonald and Popkin 1999; Rusk and Stucker 1984). I have treated this subject more extensively elsewhere (Burden 1999) and found that adjustments to the VAP did little to remedy the NES bias.

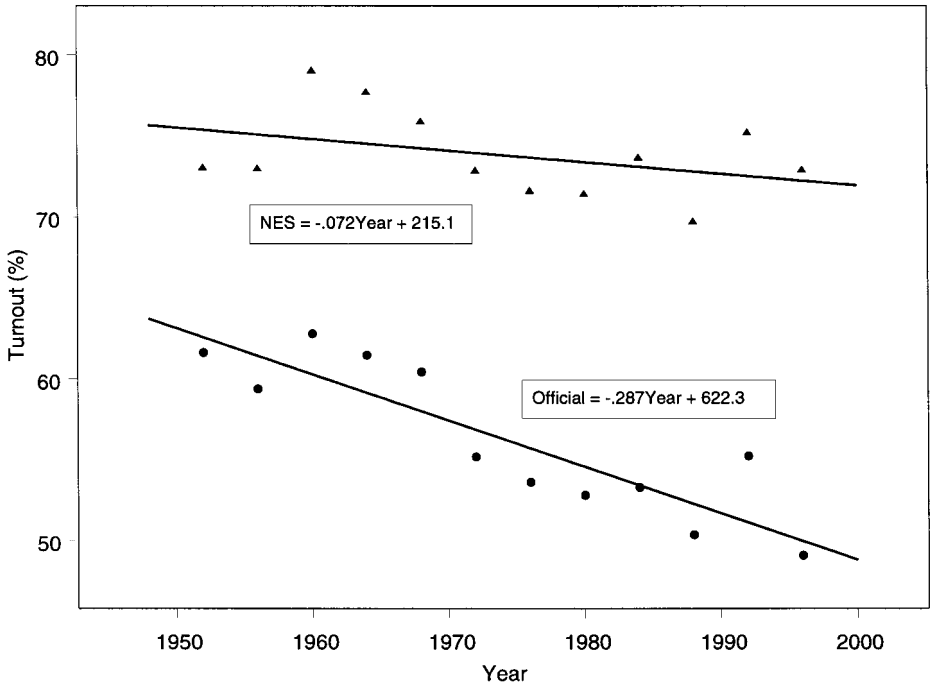

Fig. 1 Plots of the presidential turnout percentages by time from 1952 to 1996.

NES data "closely parallel the real world trends" found in official turnout statistics (see also Texeira 1987).

Despite 30 years' passage, the turnout literature indicates that that NES bias has not budged, yet the NES bias has clearly *not* remained constant. Indeed, the gap between official turnout and NES turnout estimates has *more than doubled* since the survey began regular administration in the early 1950s. This blunt fact has not yet been recognized in the literature. While it is true that official and NES turnout estimates rarely move in *opposite* directions, it is misleading to imply constant bias just because the two time series correlate highly. In Fig. 1 I present simple plots of the presidential turnout percentages by time from 1952 to 1996.[4] The ordinary least squares (OLS) regression lines are superimposed on the two plots to make trends more apparent.

This graph has two important features. First, NES turnout is always higher than official turnout. Second, the gap appears to grow in later elections. The reason is that NES turnout is nearly flat while official turnout exhibits the familiar decline from 1960 onward, with the largest dip occurring in the early 1970s as the 26th Amendment went into effect.[5] Official turnout declines with time, but there is no statistically significant trend in NES turnout.

Figure 2 considers the difference between the NES and the official time series explicitly by graphing the gap between them. The vertical axis is now NES turnout minus official

[4]These are postelection data from the cumulative NES file using weight VCF0009a. NES administrators provided this weight to compensate for panel attrition effects peculiar to the 1996 survey. Analyses of unweighted data produce results that are indistinguishable from those found here (Burden 1999).

[5]Percentages in this range have standard errors of 1.5 to 2.0 percentage points. The downward movement is not random; the percentage of elections with lower turnout than the last is significantly different from 50% ($p < .0001$).

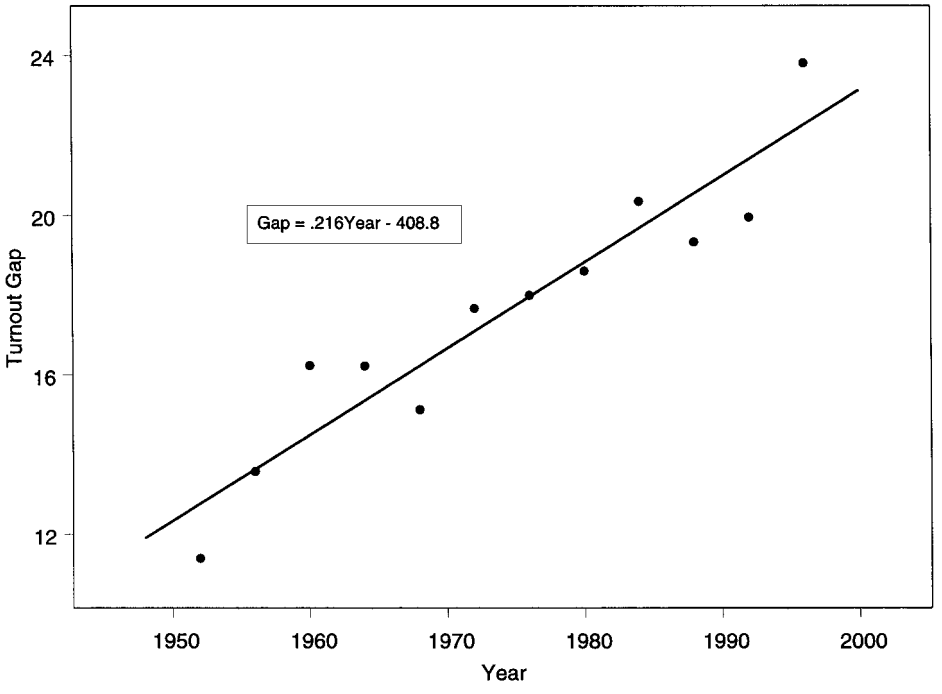

Fig. 2 Graph of the gap between the NES and the official time series.

turnout. It is clear that there is a steady increase in bias as one moves from left to right in the figure. The gap begins around 12 points in the 1950s as *The American Voter* authors found, but nearly doubles to 20 points in the late 1980s and finally reaches the mid-20s in 1996. Contrary to conventional wisdom in the turnout literature, the NES bias has not dropped below 15 points since the 1950s. The OLS line indicates that the change in the gap is nearly linear in years and grows by nearly a full point between successive presidential elections.[6]

Thus far I have focused on turnout in presidential elections. Are there similar patterns in House turnout data? Comparisons with congressional turnout might help one determine the causes of survey biases. For example, if the difference between NES and official turnout estimates grows in congressional elections as it did in presidential elections, it suggests that the difference is due mostly to the nature of the NES sampling procedure or that respondents in both kinds of surveys are equally likely to misreport. On the other hand, if the House bias is fixed over time, there must be something unique about presidential elections that caused the gap to grow only there. This is particularly true if the House gap is also constant in presidential election years since the sampling effects will be controlled.

In all but 3 election years the NES House estimates are higher than official turnout, though the gap between the two is generally small. Both exhibit the well-known oscillation that occurs between presidential and midterm years. Official and NES House turnout exhibit similar declines over time, with OLS coefficients of $-.35$ and $-.24$ (which are statistically

[6]The gap was growing before NES sampling practices changed in the late 1980s to rely more on panel than fresh cross-section respondents and more on telephone than face-to-face interviews. The same pattern emerges if one uses a measure of *relative* bias equal to (NES turnout − official turnout)/official turnout.

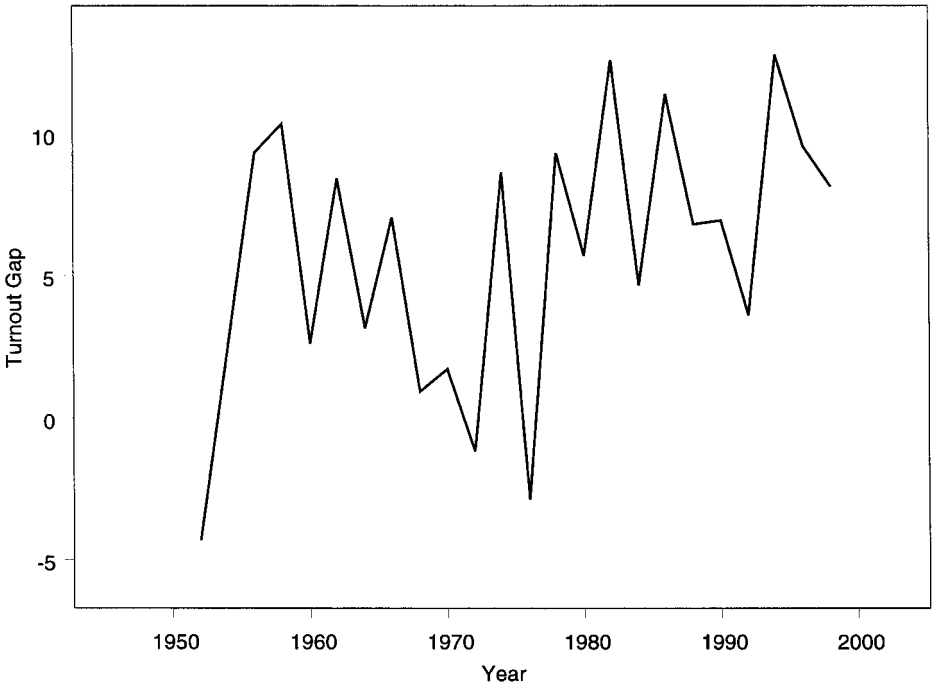

Fig. 3 Graph of the gap between NES and official House turnout estimates by year.

indistinguishable). Linear regression indicates that turnout drops between 0.5 and 0.66 point from one congressional election to the next. Unlike presidential turnout, the gap between these two measures holds steady throughout the survey era, averaging just under 6 points.

Figure 3 graphs the gap between NES and official House turnout estimates by year. In contrast to the dramatic growth in the presidential turnout bias, the House gap increases moderately with time ($b = .13$, $p = .08$). Though the relationship is positive, it is due entirely to a few outliers at the beginning and ends of the period. The relationship immediately weakens to the point of nonsignificance when the 1952 election outlier is removed from the analysis ($p = .26$). In fact, from the mid 1950s until the early 1970s, the gap actually declined.

In summary, though one might suspect that House turnout in the NES increasingly overestimated official turnout as it did for presidential elections, there is little evidence to support that conclusion. The differences between congressional and presidential races and among presidential elections themselves are much greater than among House elections. In the Conclusion I speculate about why the selection mechanism is more correlated with presidential turnout than congressional turnout. For now I merely conclude that the growing NES turnout bias is much more severe in presidential than congressional elections. Increasing presidential election bias is due to the survey's failure to mirror the drop in aggregate turnout percentages; official turnout in presidential elections declined from 1960 onward but NES turnout was flat. I attempt to determine why this is so in the following section.

3 Assessing Overreport Through Survey Participation

The common assumption is that high turnout reports are due to respondents telling postelection interviewers that they voted when they did not. This certainly happens, though

it is difficult to estimate what portion of the NES overestimate is due to such things as lying, faulty recall, and misunderstandings. Campbell et al. (1960) used conservative calculations to estimate that misreporting accounted for about 3 percentage points of the bias in the 1952 and 1956 surveys. Just a few years later, Clausen (1969) used the first and crudest vote validation study and Census data to estimate more accurately that response error in the NES was as high as 8.3 points. Sigelman (1982) later found that reported and validated vote estimates differed by 11.5 percentage points in the 1978 study. The same magnitude appeared in the 1976 study 2 years earlier (Katosh and Traugott 1981) and the 1980 NES 2 years later (Anderson and Silver 1986). Internal NES technical reports also verify that as much as 12 to 14 percentage points of the gap between official and survey turnout estimates is due to misreporting alone (e.g., Traugott 1989).

My purpose here is not to evaluate the utility of the vote validation techniques, but to show that it is difficult to determine what proportion of the NES turnout overestimate is due to misreporting. Estimates hover around 10 to 12 points in the 1970s and 1980s. Meanwhile, the four presidential elections from 1972 to 1984 saw NES estimates ranging from 17.7 to 20.3 points higher than official turnout statistics. This leaves an 8- to 9-point gap between misreported vote totals and official turnout that could be due to survey artifacts such as sample bias, preelection interview stimulation, and other survey artifacts.

Some of the bias results from *instrumentation*. The preelection interview stimulates respondents to vote at higher rates than they would have otherwise (Clausen 1969; Yalch 1976), a kind of "Hawthorne Effect." Though real, this is a minor concern in my analysis for three reasons. First, the last NES study to interview some respondents only in the postelection survey indicated small turnout differences between those people and others were also interviewed in the preelection survey.[7] Second, this effect is probably fairly constant from year to year (though it is now impossible to test this). In contrast, bias resulting from unit nonresponse is not constant.

Third, and most important, overreport is actually less severe for those surveys without a preelection component. Presidential year surveys include pre- and postelection surveys, but midterm year surveys include only the latter. If preelection stimulation effects are real, we ought to see larger discrepancies in presidential than midterm House elections. If anything, the opposite occurs. Looking back at Fig. 3, it appears that the bias is actually greater in midterm House elections than in presidential year House elections.

Even if the cases *drawn* for sampling perfectly represented the population of interest, those who actually participate are an unrepresentative subset of that pool. The first selection, done by survey administrators, is nearly random, but the second type of selection, done by potential respondents, is surely not. In other words, survey participation is generated by an unknown but nonrandom selection process (Brehm 1993; Verba 1996). Some respondents are never reached; many others refuse to participate. Even multiple attempts to reach and convince reluctant respondents fails to bring the response rate above 70 or 75% in many academic surveys such as the NES and GSS. This hard fact of survey research generates a plausible hypothesis: NES turnout estimates should be most biased when response rates are lowest. This negative relationship between response rates and reported turnout is caused by a sample that overrepresents voters. Each difficult respondent added to the sample through multiple callbacks or follow-up letters is less likely to report having voted and thus lowers

[7] In the 1960 NES, 82.3% of pre- and postelection respondents and 76.5% of the strictly postelection respondents reported voting in the presidential election. This difference, due in part to a small number of cases in the latter category, is insignificant ($t = -.626$, $p = .53$).

Table 1 Respondent pursuit and voter turnout[a]

Effort to interview respondents	*Report voting* (%)	*n*
Full sample	77.0	1534
Refusal conversion		
Not required	77.1	1345
Required	75.9	186
Persuasion letter		
Not required	77.3	1454
Required	70.8	67
Telephone calls		
1 or 2	79.9	534
3 or 4	77.7	290
5 to 8	75.8	261
9 or more	68.9	155
Face-to-face calls		
1 or 2	78.8	986
3 or 4	76.3	283
5 to 8	68.7	190
9 or more	60.8	24
Payment		
None	78.2	712
$5 to $10	75.7	746
$15 to $25	69.7	76

[a]Source: 1996 National Election Study (using full sample weight v960003).

the overall turnout percentage. People who are least likely to participate in a political survey such as the NES are also least likely to vote and to report doing so.

As evidence for these assertions, Table 1 reports the mean turnout for different groups of respondents in the 1996 NES. The results extend what Brehm (1993) found using simulations based on an earlier dataset. The efforts made to interview respondents include a verbal attempt to convert an initial refusal, a persuasive letter mailed to the respondent, financial reimbursement for the interview, and multiple calls to the person. Across the board, respondents who are more difficult to interview are less likely to report having voted. Those requiring letters are less likely to vote than those not requiring persuasive letters. The vote percentages also fall as the monetary rewards and the number of calls—by telephone or face-to-face—increase. This does not mean that pursuing respondents with these tactics "causes" them not to vote or report voting less. Rather, the marginal respondents who are difficult to interview successfully are also some of the least frequent voters. The worse our response rate, the fewer of these infrequent voters will be included in the sample used to compute the aggregate turnout percentage.

The nature of the election being studied is an important factor in determining survey biases. Presidential elections have higher turnout than congressional elections because they are more visible and intense. In traditional political behavior jargon, we make distinctions between "high-stimulus" and "low-stimulus" elections (Campbell 1960). Many presidential voters fail to vote in congressional elections since only the most interested and participatory citizens cast ballots. They are apparently willing to admit their disinterest to interviewers. Perhaps social desirability pressures are stronger in presidential elections than in congressional elections. I find, for instance, that when official turnout rises, the NES bias shrinks. So the type of election, particularly in terms of mobilization, influences the turnout bias.

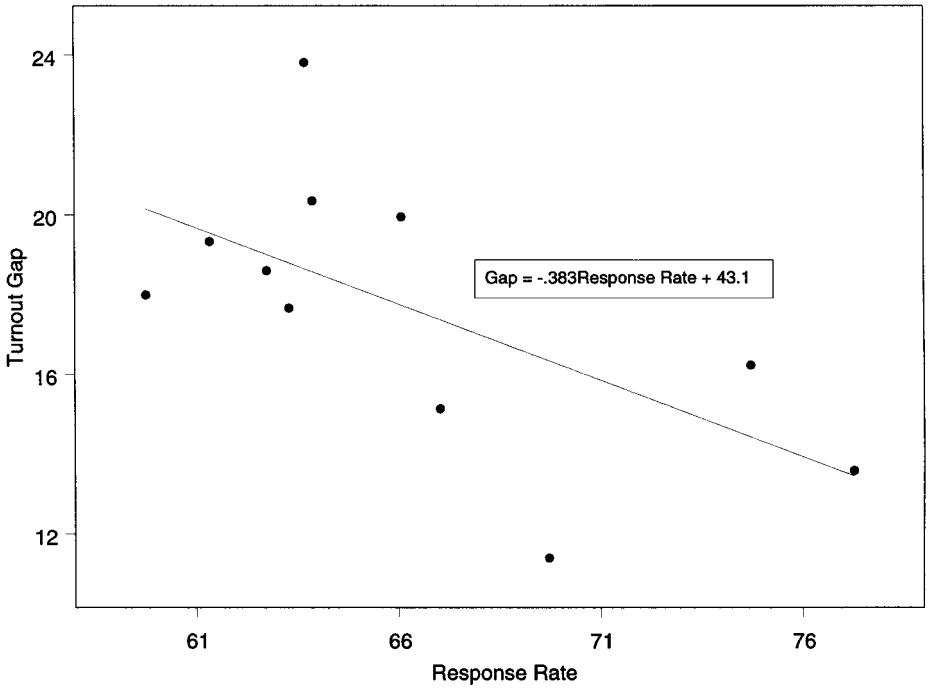

Fig. 4 Scatterplot of the relationship between the NES response rate and the difference between survey and official turnout estimates.

Confining ourselves to presidential elections, there is a more important effect that is due to the response rate. When the NES response rate *increases*, the difference between survey and official turnout estimates *shrinks*. Figure 4 shows this simple relationship in scatterplot form with the OLS line superimposed. There is a clear negative relationship between the two variables. Though response rates are not the sole cause of NES inflation, it is clear that getting more of the people in the sample to interview increases the accuracy of the survey.

It is intriguing that the relationship between bias and response rates does not hold for House turnout. The correlation is insignificant (and positive to boot), and the scatterplot between the two (not presented) also indicates a noisy relationship with no apparent pattern. In contrast, regressing the presidential bias on response rates produces a coefficient of $-.383$ ($p = .04$). Boosting response rates by 10 points to approximate early NES response percentages would probably decrease the voter turnout percentage by about 4 points. Further, a response rate of 100% would still leave a gap of 4.8 points. This rudimentary analysis leads us to the following conclusion: much of the *growing* bias in NES turnout percentages can be attributed to sample bias resulting from worsening response rates. In retrospect, it is ironic that decreasing survey participation has led to higher estimates of electoral participation.

The NES is not alone in dealing with declining response rates, nor have NES response rates been falling monotonically over time. They have generally been declining in all kinds of surveys over the long haul despite clever attempts by researchers to induce greater participation (Bradburn 1992; Groves 1989). It is important to note that falling response rates are due mostly to a growing number of *refusals* rather than a simple lack of availability. So it is not merely that Americans are more mobile and more difficult to catch at home today; an unrepresentative subset of them simply does not want to participate in surveys.

This suggests that survey researchers ought to invest as much effort into improving response rates—by converting refusals and reaching elusive respondents—as they have in improving question wording and other questionnaire design features in recent years (Fink 1995; Schuman and Presser 1981). There have been some novel attempts to reduce vote overreporting *among those interviewed* (e.g., Belli et al. 1999), but these are solutions to different problems. There are techniques such as imputation and weighting that attempt to "correct" surveys that lack full participation. These statistical remedies are not as useful for surveys such as the NES for a theoretical reason—they treat the symptoms, not the underlying illness—and a methodological reason—information about nonrespondents does not exist, which makes comparisons between respondents and nonrespondents impossible.

4 Conclusion

The NES is a truly amazing social science resource of some 40,000 interviews collected over 50 years. The fact that a half-century of opinion data have been gathered allows for the construction of a real-time series of questions in the NES "core" and should facilitate a reexamination of the quality of these data. Turnout is a natural target because of the great normative and empirical interest in it, but it is also convenient because there is an external referent. Though official turnout estimates are not accurate, nor can they ever be exact, they are close to the truth and can be improved through modest adjustments.

The increasing overestimation of voter turnout in the NES had been heretofore undetected and certainly not yet addressed. There is a reservoir of constant bias that stems from such things as misreporting, instrumentation, and sampling frame problems, but change is more intriguing. The growing problem seems due in part to decreasing response rates that are plaguing survey operations. Nonrandom selection of those who are interviewed has kept NES turnouts high and steady while real-world turnout has fallen over the past 40 years. Perhaps the bias in presidential elections is larger than in congressional elections because the greater perceived social desirability to vote in presidential contests leads to more frequent misreporting. These forces interact with selection bias to exaggerate presidential election turnout more in later surveys. Though a minimal level of misreporting is inevitable in survey research, the worsening problems must be due to things beyond this that have changed. Question wording, for instance, is not the culprit since the vote question has been hardly touched in 50 years.

Note that the turnout problem is caused primarily by unit nonresponse—where individuals are selected from the population but not successfully interviewed—rather than item nonresponse, which results from missing data on different questions for each respondent. Fortunately, there are now some new techniques for handling selection biases that result from unit nonresponse in survey data (Brehm 1993, 2000). It would be preferable to arrest the problem at its source rather than develop procedures for repairing the situation later. By investing more heavily in procedures to increase survey response rates, we can reach both cross-sectional and time-series goals while ensuring that substantive conclusions drawn from the data are as valid as possible. As attempts at voter validation and other techniques testify, researchers are motivated to make voting data as accurate as possible.

References

Abramson, Paul R., and John H. Aldrich. 1982. "The Decline of Electoral Participation in America." *American Political Science Review* 76:502–521.

Anderson, Barbara A., and Brian D. Silver. 1986. "Measurement and Mismeasurement of the Validity of Self-Reported Vote." *American Journal of Political Science* 80:771–785.

Belli, Robert F., Michael W. Traugott, Margaret Young, and Katherine A. McGonagle. 1999. "Reducing Vote Overreporting in Surveys: Social Desirability, Memory Failure, and Source Monitoring." *Public Opinion Quarterly* 63:90–108.

Bianco, William T., and Sarah Poggione. 1997. "The Pinocchio Effect: Self-Reports and Verification in the Study of Voter Turnout." Paper presented at the annual meeting of the American Political Science Association, Washington, DC.

Bradburn, Norman M. 1992. "Presidential Address: A Response to the Nonresponse Problem." *Public Opinion Quarterly* 56:391–397.

Brehm, John. 1993. *The Phantom Respondents: Opinion Surveys and Political Representation*. Ann Arbor: University of Michigan Press.

Brehm, John. 2000. "Alternative Corrections for Sample Truncation: Applications to the 1988, 1990, and 1992 Senate Election Studies." *Political Analysis* 8:183–200.

Burden, Barry C. 1999. "Measuring Voter Turnout in the National Election Studies." Paper presented at the annual meeting of the American Political Science Association, Atlanta, GA.

Burnham, Walter Dean. 1986. "Those High Nineteenth-Century American Voting Turnouts: Fact or Fiction?" *Journal of Interdisciplinary History* 16:613–644.

Campbell, Angus. 1960. "Surge and Decline: A Study of Electoral Change." *Public Opinion Quarterly* 24:397–418.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: John Wiley.

Clausen, Aage R. 1969. "Response Validity: Vote Report." *Public Opinion Quarterly* 32:588–606.

Fink, Arlene. 1995. *How to Ask Survey Questions*. Thousand Oaks, CA: Sage.

Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: Wiley.

Katosh, John P., and Michael W. Traugott. 1981. "The Consequences of Validated and Self-Reported Voting Measures." *Public Opinion Quarterly* 45:519–535.

McDonald, Michael P., and Samuel L. Popkin. 1999. "Measuring Turnout Rates Correctly: Turnout Rates Are Not Declining in Recent United States Elections." Unpublished manuscript. Nashville, TN: Vanderbilt University.

Presser, Stanley. 1990. "Can Context Changes Reduce Vote Overreporting?" *Public Opinion Quarterly* 54:586–593.

Rusk, Jerrold G., and John J. Stucker. 1984. "Measuring Patterns of Electoral Participation in the United States." *Micropolitics* 3:465–498.

Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys*. New York: Academic Press.

Shaffer, Stephen D. 1981. "A Multivariate Explanation of Decreasing Turnout in Presidential Elections, 1960–1976." *American Journal of Political Science* 26:47–56.

Sigelman, Lee. 1982. "The Nonvoting Voter in Voting Research." *American Journal of Political Science* 26:47–56.

Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson. 1986. "Who Overreports Voting?" *American Political Science Review* 80:613–624.

Teixeira, Ruy A. 1987. *Why Americans Don't Vote*. New York: Greenwood Press.

Teixeira, Ruy A. 1992. *The Disappearing American Voter*. Washington, DC: Brookings Institution.

Traugott, Santa. 1989. "Validating Self-Reported Vote: 1964–1988." National Election Study Technical Report 34. Ann Arbor: University of Michigan.

Traugott, Michael W., and John P. Katosh. 1979. "Response Validity in Surveys of Voting Behavior." *Public Opinion Quarterly* 43:359–377.

Verba, Sidney. 1996. "The Citizen Respondent: Sample Surveys and American Democracy." *American Political Science Review* 90:1–7.

Yalch, Richard F. 1976. "Pre-election Interview Effects on Voter Turnout." *Public Opinion Quarterly* 40:331–336.