

Vote "Over" Reporting in Surveys:
The Records or the Respondents?

Stanley Presser
The University of Maryland

Michael W. Traugott
The University of Michigan

Santa Traugott
The University of Michigan

Prepared for the International Conference on Measurement Errors, Tucson, Arizona, November 11-14, 1990. This research was supported in part by National Science Foundation Grant Number 8808361. The data utilized in this analysis were made available by the Inter-university Consortium for Political and Social Research. The data for the American National Election Study, 1988, were originally collected by Warren E. Miller and the National Election Studies. Neither the collectors of the data nor the Consortium bears any responsibility for the analyses and interpretations presented here.

One of the most common and consistent concerns of survey researchers has been the accuracy of respondents' reports. Considerable effort has been devoted to developing models for assessing reporting error; and the range of explanations for the error runs the gamut from individual personal and attitudinal characteristics of respondents (Abramson and Claggett, 1984 and 1986) to characteristics of the questionnaire or the individual item (Bradburn et al., 1981) to the nature of the interviewing process itself (Billiet and Loosveldt, 1988; Hanson and Marks, 1958; Silver, Abramson, and Anderson, 1986).

The accuracy of reports has most frequently been evaluated for items that lend themselves to verification or validation in administrative records. These include behaviors and clearly defined personal characteristics, but they exclude attitudes, beliefs, and values. Record checks are not the only way to conduct validation studies; in some instances, even biochemical tests have been used to validate reports of such things as smoking or drug use (Bliss and O'Connell, 1984; Strecher et al., 1989).

There has been a surprisingly wide range of variables for which self reports have been studied through record checks. They include contributions to charity and possession of licenses or permits to drive or borrow library books (Parry and Crossley, 1950); tax evasion (Hessing, Elffers, and Weigel, 1988); age, income and seniority at work (Weaver and Swanson, 1974); participation in public welfare programs (Moore and Marquis, n.d.; Weiss, 1968); hospitalizations (Marquis, 1978); and criminal victimization (LEAA, 1972; Miller and Groves, 1985). One of the most frequent focuses of record checks has been registration and voting. The centrality of electoral participation in models of social and political behavior, its importance for contemporary news coverage, and the fact that official records of such behavior are widely and publicly available all contribute to the frequency with which such

studies have been conducted. Historically, across a wide range of samples and surveys conducted by various organizations, the level of misreporting in these studies - the difference between individual-level rates of reported behavior and those validated in administrative records - has consistently been in the range of 13 to 15 percentage points. Validation studies of these measures have been conducted with municipal, state, and national samples of voters (Clausen, 1968; Miller, 1952; Parry and Crossley, 1950; Traugott, 1989); and they have been conducted by survey researchers from academic, governmental, and commercial survey organizations (Clausen, 1968; Frankovic, 1981; Jennings, 1985; Kohut, 1981; Traugott, 1989; Traugott and Katosh, 1979; Woltman and Isaki, 1978).

This finding usually has been assumed to be the result of respondents' misreporting a socially desirable form of behavior (DeMaio, 1984). However, various attempts to manipulate the survey situation in order to reduce the respondent's motivation to present the socially desirable response, i.e., that he/she voted, have not been effective. Split-ballot experiments in both the 1987 and 1989 American National Election Studies (NES) pilot surveys manipulated vote question wording to minimize social desirability. In both cases, there was no difference between the two groups in the extent to which the claim to have voted was substantiated by the vote validation efforts in 1986 and 1988. There have been at least two attempts to manipulate the context of the vote question. Most recently, Presser (1990) found that preceding the vote question by an information item that should discriminate voters from nonvoters did not affect subsequent misreporting levels, nor did a manipulation that gave respondents an opportunity to report earlier instances of voting. In an experiment on the December supplement to the 1972 Current Population Survey, the

voting and registration questions were reversed, and no difference in misreporting was found (Jennings, n.d.).

In 1984, the National Election Studies altered the placement of the registration and voting questions in the 65-minute post-election survey, from near the end of the questionnaire to near the beginning. It was thought that non-voting respondents who had talked for 55 or 60 minutes about politics would be more likely to misreport that they had voted; but the proportion of misreporters did not decline at all. As part of the same survey, half of the post-election respondents were interviewed by telephone and half face-to-face. Again, there was no difference by mode of interview in the proportion of those who misreported. Jennings (1989) has assembled data on over-reporting by year by state for Current Population Survey Voter Supplement Data, which shows that there is variation in over-reporting by state that is consistent over time. In one state, Alaska, there is consistent under-reporting of vote.

Furthermore, using the catchall explanation of social desirability begs a very important question: why have levels of misreporting remained roughly the same over the last 30 years, when actual levels of voting have declined almost precipitously? If the behavioral norm of voting as a civic duty has declined, suggesting its social desirability might have as well, why has not the tendency to misreport declined at the same time?

While social desirability is undoubtedly operating to some extent in inducing misreporting, the ubiquity of the phenomenon and its resistance to experimental manipulation suggest that other possibilities must be considered which could be making a significant and/or systematic contribution to the error. Reliance on social desirability, and

in turn on individual characteristics of respondents, makes the implicit assumption that the records themselves and the record look up procedures are correct or, at least if wrong, incorrect in minor and non-systematic ways. With the exception of recent work by Abramson and Claggett (1990), this assumption has not been looked at. One source of the mismatch between self-reports and administrative records may be the quality of the administrative records themselves, in terms of the way they are organized, maintained, or made accessible to those who are attempting to validate survey reports.

Because there are neither national registries of voter registration nor balloting, it is impossible to know the "true" registration and turnout rates in the United States. In calculations of turnout, it is common to skirt this issue by producing turnout estimates calculated on the basis of two reasonable approximations of those participating and those who were eligible to participate. The numerator consists of an estimate of the total votes cast and reported to the Clerk of the House of Representatives for the purpose of declaring the winners of presidential and congressional races.¹ The denominator consists of an estimate of citizens 18 years of age and older prepared by the Bureau of the Census,² although Passel (1989) has noted that the Bureau produces several different denominator estimates and some are more error prone than others.

Data are presented in Table 1 that permit the comparison of this aggregate estimate of turnout and the two most common survey-based estimates of turnout in recent presidential elections. One is derived from relatively small national samples of the electorate conducted by NES, usually consisting of between 1,400 and 2,400 respondents interviewed before and after the election who are asked to report on their own registration

status and voting behavior. The other is derived from much larger samples, typically ranging from 30,000 to 40,000 interviews conducted shortly after the November election, in which respondents are asked to report on their own registration status and voting behavior, as well as that of other adults residing in their household. Both of these surveys produce higher estimates of voting than the one generated from the aggregate data, with the NES differences from the aggregate being from two to three times as great (15.8 to 20.5 percentage points) as the Census difference from the aggregate (5.7 to 7.8 percentage points). The basis for these differences, which lie primarily in study design, sample characteristics, and interviewing procedures, have been reported extensively (Clausen, 1968; Jennings, 1985; Traugott and Katosh, 1979).

Table 1 about here

The NES has conducted validation studies in order to understand the sources of error in its estimates and to eliminate nonsampling errors that come from misreporting. In a presidential year, the NES survey consists of both pre-election and post-election interviews that measure a wide variety of politically relevant attitudes and behavior, as well as relevant personal characteristics of each respondent. Since 1978, these studies have also included an attempt to validate respondents' reports of their registration status and voting behavior, usually conducted in the spring of the following year. When the interviewers visit the office of the local official responsible for the electoral district in which it is believed each respondent resides, an interview is conducted with someone in the office as a means of obtaining information about the procedures for storing and maintaining records. This provides both a means of establishing rapport with the office manager as well as providing

the interviewer with a perspective on what kinds of information in what forms will be available in the office for validation. A complete specification of the variables from the 1988 NES survey used in this study is given in Appendix A.

The complexity of the matching operation is seldom fully appreciated. Except for a very few localities, respondents must be registered in order to vote, and they must vote in the jurisdiction in which they are registered. Usually two distinct dependent variables have to be operationalized. The first is whether or not a self-report of registration can be validated, and the second is whether or not a self-report of voting could be validated for those for whom a registration record could be found. In order to understand whether the sources of error for vote reports are different than those for mismatches on registration, it is necessary to hold constant the registration matching status of respondents. It is only through such analyses that an assessment can be made of whether different factors explain relative success in matching these two different kinds of reports, especially as a function of the characteristics of the office and its record keeping requirements, procedures, and formats. Failure to find a registration record does not have the same analytical status as finding a record that is marked in an inconsistent manner relative to the respondent's report. It is impossible to disconfirm the plausible counter-hypothesis that a record cannot be located because of problems of name or address spelling, incorrect addresses, or a misunderstanding about the locality in which the respondent is actually registered.

Traugott (1989) has shown that the discrepancy between self-reported and validated voting, when there is enough information about a respondent to be able to check the records, consists of two parts. It is equally divided between a failure to find a registration

record for those who claim to have voted and a failure to confirm a report of voting for those who have been confirmed as registered. Since 1980, the overall proportion of self-reported voters whose behavior could not be validated has ranged from 11.9 to 16.9%. The proportion of those for whom no registration record could be found has ranged from 5.8 to 7.8%, while the proportion for whom no voting confirmation could be found ranged from 6.1 to 8.7%.³ These data suggest the need to evaluate the results of validation efforts in two parts: the ability to locate a registration record for those who indicate they are registered and the ability to locate a voting record for those for whom registration has been determined.

However, most past voter validation research has ignored this distinction and simply combines the two types of match failures as "misreporting." If the same variables are associated with both kinds of measurement error - whether a record is found and whether it is marked to correspond with a respondent's report - then the distinction may not be problematical. On the other hand, if the same variables are not associated with the failure to find a registration record as the failure to confirm reported voting after a registration record has been found, then there is reason to be concerned about the practice of lumping together both kinds of error under the general heading of "misreport."

In addition, past work has usually interpreted the lack of a match as due solely to respondent error, despite the fact that there may be at least two other factors operating. There may be errors in the records. And there may be errors made in the process of matching the records to the survey data. Unless these factors are carefully considered, and ruled out, there is a danger of putting the theoretical cart before the horse.

The focus of the analysis presented below is an alternative conceptualization of this type of survey measurement error. The main variable to be explained is "matching success," or the rate at which relevant records matched a survey response. In this model, there are two primary explanatory factors for a failure to match: personal factors associated with the characteristics of the respondent and factors related to the matching process itself. In the latter category, there are two important subcategories of explanatory factors. One is the matching procedure employed, which may involve a number of different procedures and algorithms, including the use of computer programs (Fellegi and Sunter, 1969; Jabine and Scheuren, 1986).⁴ A second is the characteristics of the office in which the records are being searched. This includes the rules and regulations that govern registration and voting in the jurisdiction, such as record keeping requirements, the access to the records provided to the person doing the validation, and the search process employed by the validator.

The analysis that follows is based upon data obtained as part of the 1988 NES study and represents a preliminary attempt to evaluate the reconceptualization of the validation process into personal and matching components. Most of the analysis consists of bivariate relationships, with some investigation of interaction effects between relevant personal characteristics and features of the office in which the validation was conducted.⁵ Multivariate analyses must eventually be conducted. The conclusion contains a discussion of the significance of these findings for our understanding of measurement error in surveys, for the conduct of additional measurement error studies, and for improvements in the types of survey items and office characteristic variables which might be collected in order to improve the validation process itself.

Research Results

Data are presented in Table 2 that indicate the "match rates" for such self-reported registration and voting in the entire 1988 NES sample and for certain selected personal characteristics of the respondents. Overall, a registration record could be located for 89.6% of the respondents who indicated they were registered; and a voting record could be located for 91.6% of those for whom a registration record was found. The match rates are not significantly different for men and women, but they do tend to be higher for older respondents, for whites compared to Blacks, and for those who live outside of the very largest cities. Match rates also tend to be lower in the South and relatively higher in the Midwest and West.

Table 2 about here

From the perspective of ascertaining levels of misreporting or evaluating the validation effort, demography should not be the most important factor that explains relative success in finding an administrative record. Rather it should be personal factors of life style and cycle that are, in turn, related to administrative requirements for registration and voting. Since registration is the initial hurdle that must be overcome, these personal factors should be more strongly related to the match rate for registration records than for voting records for those who have been validated as registered. The data presented in Table 3 show that this is indeed the case. It is more difficult to find registration records for renters than home owners, for example, or for those who have resided in their community or at their current address for relatively short periods of time. However, these relationships are much weaker or nonexistent for matching voting records.

Table 3 about here

Some of the strongest differences in registration match rates occur for women who have had a recent change in their legal name or in their marital status. Although the sample sizes for some of these categories are small, the differences are nevertheless striking. Women who have had a legal name change in the past three years are much less likely to have their registration status validated than those who have not (70.8% compared to 89.0%), and women who are separated are also less likely to be matched in the official records. If a registration record was found, there is no difference in the match rate for self-reported voting for women with and without name changes; but the differences for those who are separated are just as large.

There are many ways that characteristics of the office in which the validation effort is being conducted might affect the match rate for the survey reports. One has to do with the size and complexity of the office operations. Data are presented in Table 4 which show that one measure of the office load - the number of precincts in the jurisdiction - is not related to match rates. As the number of administrative units for which the office is responsible increases, there is no effect on match rates. But another measure - the average number of registered voters per precinct - is related. As the population size of precincts increases (the number of individual records to be managed), the match rates decline. This is not simply a function of the way in which the records are managed, as the use of computers is not related to matching either self-reported registration or voting.

Table 4 about here

A second way in which office characteristics might affect the success of the validation effort has to do with the procedures for managing records and providing access to them. Data are presented in Table 5 that show that as the complexity or number of administrative arrangements increases, the success rate of matching declines. And the greater the difficulty which the interviewer might have in getting direct access to the records, the lower the match rates tended to be. But again, there are differences between registration and voting. Whether or not the registration and voting records are kept in the same or separate files affects the match rate for voting but not for registration. This is also true for whether or not all of the voting records are accessible to the interviewer. The match rates are higher in both cases if the record keeping for the 1988 election had been completed by the time of the interviewer's visit to the office in the spring of 1989. While most of the local election administration offices are using a form of computerized records, about 15% of the survey respondents were from places in which the records were not computerized. In these instances, the person doing the validation might or might not have been able to handle the original records himself or herself. Direct access to records did not improve the match rate for registration, but it did result in a higher match rate for voting.

Table 5 about here

Two other variables were used as indicators of possible problems with the records and difficulty in accessing them. In many locales, there was more than one place at which a person could register, indicating that either multiple sets of records had to be consolidated by the clerk or the possibility of some records still being at another location. If the validator was at the only place in which a person could register, then the match rate for registration

was higher, as would be expected. This variable had no effect on the match rate for voting. The other measure was an indication of whether name alone could be used to locate a person or whether exact address was needed as well, as a linkage to precinct. The need for name and address was not related to a difference in match rates for registration but it was related to the match rate for voting. The more information needed, the less likely a voting record was to be found.

These six variables were combined in a simple additive index of record quality and access that measured how difficult it was to locate a voting record.⁶ Given the problems of matching registration records, only the match rate of voting information for those respondents for whom registration records had been found is used in this analysis. The number of "complications" were summed, so that "high" quality indicated a lack of complications. The data presented in Table 6 show there was a strong relationship between the match rate for voting and the index of record quality. In the total sample, about half the respondents lived in areas in which the index had a high value, and the match rate for validating voting was greatest there (95.2%). Another four in ten of the respondents resided in places where the index indicated a medium level of quality and access, and here the match rate was 90.3%. One in twelve respondents resided in areas with low record quality, and their match rate was only 75.0%.

Table 6 about here

The index also provides a useful perspective for understanding the relationship between individual demographic characteristics that have been shown to be related to other measures of misreporting and the index. In 14 out of 15 categories of five variables, the

match rate for validated voting increases with better record quality and access. In the South, there was no difference between the match rate in places with medium record quality and access (92.9%) and high quality (92.0%).

The data do make evident strong interactions between some personal characteristics and the quality of voting records. For example, Black respondents in the NES sample are twice as likely as whites to come from a jurisdiction with a low score on the index (16% compared to 8%). Residents of large central cities are also twice as likely to come from places with low quality records as those who live in small, rural locations (18% compared to 9%). And the match rates for respondents from the Northeast and the South were much lower when the index had a low value than in the Midwest. There were no respondents from such places in the West.

Conclusions

On the basis of vote validation studies conducted over more than 40 years, it is widely assumed that on the order of one in six individuals falsely claim to have voted in the most recent election. The results presented here suggest that the magnitude of this problem may be overstated. Instead of assuming that discrepancies between the survey reports and the administrative records are due solely to respondent error, as is done in almost all prior research, this study investigated the possibility that the information from the record check was in error as well.

In the past, data extracted from administrative records have been treated as the standard against which the respondent reports are assessed. In this revised approach, both sources of information are treated as imperfect indicators of the actual behavior. Instead

of examining whether respondent characteristics are associated with "inaccuracy," this analysis examined whether characteristics of either the respondent or the record checks are associated with "mismatches."

The results indicate that the record check variables are indeed related to matching success. Vote mismatches are particularly likely to occur in offices that keep registration and voting records separately, where an exact address is needed to access records, where not all the records are directly accessible, and where the records are incomplete. The next step is to analyze the extent to which these variables are correlated with respondent-level characteristics, in order to be able to estimate the relative proportion of mismatches that are due to the records as opposed to the respondents.

The results also suggest that the relationships between respondent characteristics and "inaccuracy" reported in prior studies may be overstated. Thus the well-known finding that Blacks misreport voting more than whites appears to be due to the fact that Blacks are more apt to live in areas where voting records are less accessible and complete. When the variation in the quality of and access to records is controlled for, the mismatch difference by race is cut in half.

The record check variables analyzed here are of two types: characteristics of the records themselves and characteristics of the checking procedure. There are several variables of the second type that have not been analyzed. One is the identity of the validator, who in these data are Survey Research Center interviewers and supervisors. Just as there are interviewer effects in data obtained from respondents, it is likely that there are validator effects in data obtained from complex record systems. It should be possible to

take advantage of the fact that different validators were sent to the same place after different election studies to begin to estimate the magnitude of these validator effects.

Another variable is the matching procedure employed. The work described above involves manual procedures by which humans compare a list of names and addresses of respondents to similar information on registration and voting lists. The use of certain computer algorithms for comparing two machine-readable lists would permit matching to be done on a stochastic basis; rather than deciding that a match was "made" or not, a probabilistic weight could be attached to a given match. The use of such procedures would introduce another factor in the explanation of the match rate.

The findings of this study have implications for the conduct of voting validation studies. For example, it may be useful to collect information by mail or phone on the office-level variables that were found to be related to mismatches before the validators are sent to the offices to check records. Then the instructions and forms used by the interviewers could be tailored to the situations which the validators will confront.

The findings may also have implications for validation studies in other subject areas, most of which have also treated the results of the record check as error free. In particular, the demonstration of the usefulness of distinguishing between mismatches due to not finding a record versus those due to finding a record with a value different from the one reported by the respondent may prove useful in studies of other self-reported behaviors.

REFERENCES

- Abramson, P.R. and Claggett, W., "Race-Related Differences in Self-Reported and Validated Turnout," **Journal of Politics**, Vol. 46, 1984, pp. 719-738.
- Abramson, P.R. and Claggett, W., "Race-Related Differences in Self-Reported and Validated Turnout in 1984," **Journal of Politics**, Vol. 48, 1986, pp. 412-422.
- Abramson, P.R. and Claggett, W., "The Quality of Record Keeping and Racial Differences in Validated Turnout," unpublished manuscript, 1990.
- Billiet, J. and Loosveldt, G., "Improvement of the Quality of Responses to Factual Survey Questions by Interviewer Training," **Public Opinion Quarterly**, Vol. 52, No. 2, 1988, pp. 190-211.
- Bliss, R. and O'Connell, K., "Problems with Thiocyanate as an Index of Smoking Status: A Critical Review with Suggestions for Improving the Usefulness of Biochemical Measures in Smoking Cessation Research," **Health Psychology**, Vol. 3, No. 6, 1984, pp. 563-581.
- Bradburn, N.M., Sudman, S., and Associates, **Improving Interview Method and Questionnaire Design**, 1981, Josey-Bass, San Francisco.
- Clausen, A., "Response Validity: Vote Report," **Public Opinion Quarterly**, Vol. 41, No. 1, 1968, pp. 56-64.
- DeMaio, T., "Social Desirability and Survey Measurement: A Review," in Surveying Subjective Phenomena, Vol. 2, Turner, Charles F., and Martin, Elizabeth, eds., NY: Russell Sage Foundation, 1984.
- Fellegi, I. and Sunter, A., "A Theory for Record Linkage," **Journal of the American Statistical Association**, Vol. 64, 1969, pp. 1183-1210.
- Frankovic, K. Unpublished memorandum, CBS News Election and Survey Unit, 1981, New York, NY.
- Hanson, R.H. and Marks, E.S., "Influence of the Interviewer on the Accuracy of Survey Results," **Journal of the American Statistical Association**, Vol. 53, No. 283, September, 1958, pp. 635-655.
- Hessing, D. J., Elffers, H., and Weigel, R.H., "Exploring the Limits of Self-Reports and Reasoned Action: An Investigation of the Psychology of Tax Evasion Behavior," **Journal of Personality and Social Psychology**, Vol. 54, 1988, pp. 405-413.

Jabine, T.B., and Scheuren, F.J., "Record Linkages for Statistical Purposes: Methodological Issues," **Journal of Official Statistics**, Vole. 2, No. 3, 1986, pp.255-277.

Jennings, J.T., "How Right Is the Vote Count?," Unpublished manuscript, September 27, 1985.

Jennings, J.T., "Misreporting Voting in the Current Population Survey," Unpublished manuscript, U.S. Bureau of the Census, n.d., Washington, D.C.

Kohut, A., "A Review of the Gallup Pre-Election Methodology in 1980," Proceedings of the American Statistical Association Annual Meeting, 1981, pp. 41-46.

Law Enforcement Assistance Administration, **San Jose Methods Test of Known Crime Victims**, Statistics Technical Report No. 1, Washington, D.C., June, 1972.

Marquis, K. **Record Check Validity of Survey Responses: A Reassessment of Bias in Reports of Hospitalizations**, Rand Corporation, 1978.

Miller, M., "The Waukegan Study of Voter Turnout Prediction" **Public Opinion Quarterly**, Vol. 16, No. 3, 1952, pp. 381-398.

Miller, P.V., and Groves, R., "Measuring Reporting Accuracy of Criminal Victimization Through Record Checks," **Public Opinion Quarterly**, Vol. 49, Fall, 1985, pp. 366-380.

Moore, J.C. and Marquis, K.H., "Response Errors in SIPP: Some Beginning Estimates Based on Administrative Records," Unpublished memorandum, U.S. Bureau of the Census, n.d., Washington, D.C., pp. 1-6.

Parry, H.J. and Crossley, H.M., "Validity of Responses to Survey Questions," **Public Opinion Quarterly**, Vol. 14, No. 1, 1950, pp. 61-80.

Passel, J.S., "Effects of Population Estimates on Voter Participation Rates," paper presented at the annual meeting of the American Statistical Association, 1989, Washington, D.C.

Presser, S., "Can Changes in Context Reduce Vote Overreporting in Surveys?" **Public Opinion Quarterly**, Vol. 54, No. 4, 1990, pp. 586-593.

Silver, B.D., Abramson, P.R., and Anderson, B.A., "The Presence of Others and Overreporting of Voting in American Elections," **Public Opinion Quarterly**, Vol. 50, No. 2, 1986, pp. 228-239.

Strecher, V., Becker, M., Clark, N., and Prasada-Rao, P., "Using Patients' Descriptions of Alcohol Consumption, Diet, Medication Compliance, and Cigarette Smoking: The Validity of Self-Reports in Research and Practice," **Journal of General Internal Medicine** Vol. 4, 1989, pp. 160-166.

Traugott, S., "Validating Self-Reported Vote: 1964-1988," paper presented at the annual meeting of the American Statistical Association, Washington, D.C., 1989.

Traugott, M.W. and Katosh, J.P., "Response Validity in Surveys of Voting Behavior," **Public Opinion Quarterly**, Vol. 43, No. 3, 1979, pp. 359-377.

Weaver, C.N. and Swanson, C.L., "Validity of Reported Date of Birth, Salary, and Seniority," **Public Opinion Quarterly** Vol. 38, No., 1, 1974, pp. 69-80.

Weiss, C.H., "Validity of Welfare Mothers' Interview Responses," **Public Opinion Quarterly**, Vol. 32, No. 4, 1968, pp. 622-626.

Woltman, H. and Isaki, C., "Measurement Error and Sample Design Research for the 1976 Registration and Voting Survey," paper presented at the annual meeting of the American Association for Public Opinion Research, Roanoke, VA, 1978.

APPENDIX A

Variables from Interviews with Respondents

V8 REGION This variable identifies the four major geographic regions on which the sample was stratified for the purpose of sample selection.

Northeast (Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont)

Midwest (Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin)

South (Alabama, Arkansas, Delaware, D.C., Florida, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Utah, Virginia, West Virginia)

West (Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming)

R22 SIZE AND TYPE OF PLACE Based on V22 Belt code. This variable was coded according to the 1980 Census with additions from Census Population Reports.

V412 RACE Based on interviewer observation.

V413 SEX Respondent's sex from household listing (given by informant).

V31 LEGAL NAME CHANGE *Has your legal name changed in the past three years due to marriage, divorce, or for any other reason?*

R417 AGE Recode based on V415: *What is the month, date and year of your birth?*

V418 MARITAL STATUS *Are you married now and living with your (husband/wife) or are you widowed, divorced, separated, or have you never been married?*

V548 YEARS IN COMMUNITY *How long have you lived in your present (CITY/TOWN/TOWNSHIP/COUNTY)?*

V551 YEAR IN DWELLING UNIT *How long have you lived in this (house/apartment)?*

V552 OWN/RENT HOME *(Do you/Does your family) own your home, pay rent, or what?*

Variables from Interviews with Election Official or Those Determined by Interviewer

V1174 COMPUTERIZED FILE *Is this file on a computer?*

R301 AVERAGE REGISTERED VOTERS/PRECINCT Based on V1215 *As of today, approximately how many registered voters are there in this jurisdiction?* divided by V1217 # of Precincts (see below).

V1217 # OF PRECINCTS *How many (election districts/precincts) do you have in this jurisdiction?*

V1133 SAME FILE Are registration records (or master file) updated with vote information?

V1145 ACCESS OF VOTING RECORDS Determined by interviewer.

V1176 NAME ENOUGH *Do you need to know someone's exact address or precinct to locate them on this file or is having their name enough?*

V1186 ENTRY COMPLETE *Is this entry process complete for the 1988 General Election?*

V1222 ONLY OFFICE TO REGISTER *Is this the only place at which people in this jurisdiction can register?*

V1299 HANDLE RECORDS *Were you allowed to touch or handle the records yourself? (Asked of the interviewer)*

V1144 REGISTRATION RECORD This variable is based on V1118 *Does the respondent have a registration record in this office?* and V1120 *Where did you find the record?*

V1147 VALIDATED VOTE This variable is based on V756 *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. How about you--did you vote in the elections this November?* and V1146 which was a record of the respondents' vote.

Table 1. The Relationship between Aggregate and Survey-Based Estimates of Turnout in Presidential Elections, 1964-1988

	Presidential Election Years					
Alternative Turnout Estimates	<u>1964</u>	<u>1972</u>	<u>1976</u>	<u>1980</u>	<u>1984</u>	<u>1988</u>
Aggregate Vote ^a	61.9%	55.2%	53.5%	52.6%	53.1%	50.2%
Bureau of the Census ^b	69.3%	63.0%	59.2%	59.2%	59.9%	57.4%
National Election Survey ^c	77.7%	72.8%	71.6%	71.4%	73.6%	70.0%
Differences in Estimates ^d						
NES - Aggregate	15.8	17.6	18.1	18.8	20.5	19.8
NES - Census	8.4	9.8	12.4	12.2	13.7	12.6
Census - Aggregate	7.4	7.8	5.7	6.6	6.8	7.2

^aThis rate consists of the Clerk of the House's estimate of votes cast divided by the voting age population (Statistical Abstract of the United States, 1990, Table 443).

^bThis rate is derived from the November Voter Supplement to the Current Population Survey, using self-reported voting behavior divided by a survey-based estimate of the population aged 18 and older.

^cThis rate is derived from the Center for Political Studies' National Election Survey, using self-reported voting behavior divided by the number of post-election respondents.

^dThese differences are expressed in percentage points.

Table 2. The Relationship between Personal Demographic Characteristics and Rates of Validation of Self-Reported Registration and Voting, 1988

	Match Rate for Self-Reported	
	<u>Registration^e</u>	<u>Voting^f</u>
Total Sample	89.6% (1378)	91.6% (1225)
Gender		
Male	90.0% (602)	92.6% (539)
Female	88.1% (776)	90.8% (686)
Age		
18-23	82.6% (886)	87.3% (71)
24-39	87.8% (474)	89.2% (411)
40-65	88.4% (553)	93.5% (491)
66 and Older	94.7% (265)	92.6% (203)
Race		
White	90.4% (1172)	93.2% (1060)
Black	80.9% (162)	79.2% (130)
Size and Type of Place		
Large central cities	88.6% (316)	85.7% (279)
Suburbs of large cities	91.9% (608)	93.1% (550)
Other places	85.2% (454)	93.7% (396)
Region		
Northeast	87.8% (246)	90.3% (216)
Midwest	92.7% (413)	94.6% (390)
South	84.8% (440)	89.1% (366)
West	91.0% (279)	91.7% (253)

^eRate of success in finding a registration record for respondents who reported they were registered.

^fRate of success in validating a respondent's report of voting behavior among those respondents for whom a registration record was found.

Table 3. The Relationship between Personal and Residential Stability and Rates of Validation of Self-Reported Registration and Voting, 1988

	Match Rate for Self-Reported	
	<u>Registration^g</u>	<u>Voting^h</u>
Total Sample	89.0% (1378)	91.6% (1225)
Home ownership		
Yes, own a home	91.3% (956)	92.9% (873)
No, rent	82.8% (383)	88.1% (318)
Years in the Community		
Less than 2 years	84.7% (177)	94.0% (149)
2-4 years	81.3% (75)	90.2% (61)
5-8 years	87.6% (129)	91.2% (113)
9 or more years	90.5% (997)	91.4% (902)
Years in the Dwelling Unit		
Less than 2 years	85.3% (382)	90.7% (322)
2 - 4 years	86.6% (149)	88.5% (130)
5 - 8 years	86.6% (202)	92.5% (174)
9 or more years	92.4% (644)	92.6% (598)
For Female Respondents only:		
Legal name change in last 3 years		
Yes	70.8% (48)	88.6% (35)
No	89.0% (616)	90.7% (547)
Marital Status		
Married	90.3% (401)	92.9% (366)
Never Married	89.2% (102)	86.5% (89)
Separated	71.4% (35)	68.0% (25)
Divorced	86.5% (104)	90.0% (90)

^gRate of success in finding a registration record for respondents who reported they were registered.

^hRate of success in validating a respondent's report of voting behavior among those respondents for whom a registration record was found.

Table 4. The Relationship between Office Characteristics and Rates of Validation of Self-Reported Registration and Voting, 1988

	Match for Self-Reported	
	<u>Registrationⁱ</u>	<u>Voting^j</u>
Total Sample	89.0% (1378)	91.6% (1225)
# of Precincts in Jurisdiction		
1 st Quintile	90.5% (273)	91.5% (246)
2 nd	87.9% (256)	93.1% (218)
3 rd	89.5% (277)	95.7% (257)
4 th	90.8% (260)	87.2% (234)
5 th Quintile	88.1% (269)	89.5% (237)
Average Registered Voters/Precinct		
1 st Quintile	92.1% (266)	95.5% (245)
2 nd	89.7% (281)	94.0% (251)
3 rd	91.3% (219)	90.5% (201)
4 th	84.5% (245)	90.2% (215)
5 th Quintile	89.3% (270)	89.2% (232)
Are registration records computerized?		
Yes	89.1% (1193)	91.2% (1061)
No	91.4% (58)	90.6% (53)

ⁱRate of success in finding a registration record for respondents who reported they were registered.

^jRate of success in validating a respondent's report of voting behavior among those respondents for whom a registration record was found.

Table 5. The Relationship between Record Management and Access Procedures and Self-reported Registration and Vote, 1988

	Match Rate for Self-Reported	
	<u>Registration^k</u>	<u>Voting^l</u>
Total Sample	89.0% (1378)	91.6% (1225)
Are registration and voting records kept in the same file?		
Yes	89.4% (1047)	93.3% (935)
No	87.5% (311)	86.5% (275)
Are all voting records accessible?		
No	87.1% (170)	83.8% (142)
Yes	89.2% (1208)	92.6% (1083)
Is exact address needed to check or is name enough?		
Need precinct or address	88.6% (149)	83.3% (132)
Name enough	89.0% (1183)	92.5% (1051)
Is record keeping for the 1988 election completed?		
Yes	89.1% (671)	92.9% (607)
No	83.6% (61)	80.0% (50)
If the records are not computerized, could the interviewer handle them?		
Yes, could touch records	87.6% (137)	97.3% (149)
No, could not touch records	96.1% (51)	81.6% (49)
Is this the only place at which a person can register?		
Yes	93.1% (354)	92.6% (311)
No	88.3% (1029)	91.2% (899)

^kRate of success in finding a registration record for respondents who reported they were registered.

^lRate of success in validating a respondent's report of voting behavior among those respondents for whom a registration record was found.

Table 6. The Relationship between Index of Record Quality and Access and the Rate of Validation of Self-Reported Vote, among Validated Registrants, 1988

Index of Record Quality and Access				
	<u>Low</u>	<u>Medium</u>	<u>High</u>	<u>Overall</u>
Total Sample	75.0% (108)	90.3% (423)	95.2% (691)	91.6% (1225)
Gender				
Male	71.7% (53)	92.1% (189)	97.0% (296)	92.6% (539)
Female	78.2% (55)	88.9% (234)	93.9% (395)	90.8% (686)
Age				
18-23	62.5% (8)	85.7% (14)	93.9% (33)	87.3% (55)
24-39	58.1% (31)	90.9% (154)	92.5% (240)	89.2% (426)
40-65	89.4% (47)	90.3% (175)	96.6% (267)	93.5% (491)
66 and Older	68.4% (19)	90.3% (62)	97.5% (122)	92.6% (203)
Race of Respondent				
White	80.5% (87)	91.8% (364)	95.9% (609)	93.2% (1060)
Black	52.4% (21)	81.3% (48)	89.7% (58)	79.2% (130)
Size and Type of Place				
Large central cities	70.0% (50)	88.1% (109)	90.0% (120)	85.7% (279)
Suburbs of large cities	76.9% (26)	91.6% (226)	95.6% (298)	93.1% (550)
Other places	77.1% (35)	89.8% (88)	97.1% (273)	93.7% (396)
Region				
Northeast	61.5% (13)	89.1% (119)	96.4% (84)	90.3% (216)
Midwest	82.6% (46)	91.9% (62)	97.2% (282)	94.6% (390)
South	69.2% (52)	92.9% (140)	92.0% (174)	89.1% (366)
West	---- (0)	87.3% (102)	94.7% (151)	91.7% (253)

NOTES

1. The error in this estimate comes from the fact that not everyone who goes to the polls casts a vote for every office on the ballot. Therefore, the total vote cast for a given office, especially one at the "top" of the ticket is a reasonable but inexact count of all those who voted. This estimate most closely approximates the total number of voters in presidential contests.

2. The minimum age for voting eligibility in presidential elections is 18; but in order to participate in an election, a citizen must be registered. This means satisfying state and local residency requirements and filling out registration forms. Many citizens fail to complete these procedures, and therefore the number of citizens 18 years of age and older is much greater than the number that age who are registered to vote.

3. In 1984, the NES stopped trying to validate respondents' reports that they were not registered because it is unclear what the inability to locate a record means in that circumstance.

4. In order to employ computerized record matching, there must be two sets of records available in machine-readable form. One consists of information about the survey respondents, while the other consists of all the relevant data in the administrative records on those who might resemble the survey respondents, including those with similar names, addresses, birthdays and gender, for example. Because the latter information is not available for a national sample of registrants and voters, computerized record matching algorithms have not been employed in past studies. This does not mean that this would not be a fruitful avenue for research in studies employing local samples in jurisdictions in which administrative records are computerized and could be made available for this purpose.

5. Significance tests are not used in this analysis because the National Election Studies samples, as is typical of face-to-face interview studies, are heavily clustered by geographical area. This is a more serious problem than usual because the analyses of individual voters depend so heavily on contextual characteristics of where they live and vote; i.e., the election office characteristics. Significance tests will eventually be performed on the basis of the effective sample sizes, after appropriate calculations have been performed.

6. The index is a summation or count of four kinds of complications which make the matching task more difficult. We used variable 1133, whether the registration record was NOT updated with vote information; variable 1145, all voting records were NOT available; variable 1176, exact address IS needed to locate a respondent on the voting records; variable 1179, the process of record updating

is NOT complete; and variables 1174 and 1299, if the office records are not computerized, the SRC field staff person was NOT allowed to handle them herself. The measure is simple count of the number of these complications in a particular office, and could run from 0-4 although the highest observed score was 3.