



CALTECH/MIT VOTING TECHNOLOGY PROJECT

A multi-disciplinary, collaborative project of
the California Institute of Technology – Pasadena, California 91125 and
the Massachusetts Institute of Technology – Cambridge, Massachusetts 02139

**TITLE Correcting for Survey Misreports using Auxiliary
Information with an Application to Estimating Turnout**

**Name Jonathan N. Katz
University California Institute of Technology**

**Name Gabriel Katz
University California Institute of Technology**

Key words:

**VTP WORKING PAPER #74
May 2009**

Correcting for Survey Misreports using Auxiliary Information with an Application to Estimating Turnout

Jonathan N. Katz

Gabriel Katz

Abstract

Misreporting is a problem that plagues researchers that use survey data. In this paper, we give conditions under which misreporting will lead to incorrect inferences. We then develop a model that corrects for misreporting using some auxiliary information, usually from an earlier or pilot validation study. This correction is implemented via Markov Chain Monte Carlo (MCMC) methods, which allows us to correct for other problems in surveys, such as non-response. This correction will allow researchers to continue to use the non-validated data to make inferences. The model, while fully general, is developed in the context of estimating models of turnout from the American National Elections Studies (ANES) data.

A question fundamental to the study of politics is what determines a citizen's decision to participate in the political process. In particular, many studies have sought to understand what drives an individual to turnout to vote on election day.¹ The empirical studies of turnout are all interested in how the probability of an individual voting varies according to relevant observable factors, such as citizen's level of political information, registration laws, or demographic characteristics. That is, these studies are interested in estimating the conditional distribution of the turnout decision given characteristics of interest. This almost always leads to estimation of the common logit or probit models, since the turnout decision is dichotomous, although there are alternatives such as scobit (Nagler 1994) or non-parametric models (Härdle 1990) for discrete choice models.

A problem arises because we do not (easily) observe the decision to vote because of the use of secret ballot in the U.S. Even if we could observe turnout from the official ballots we would not, in general, be able to also observe all the characteristics — e.g., the voter's policy preferences or information about the candidates — that presumably affect

¹The literature is far too vast to even begin to fully cite here, however see Aldrich (1993) for a review of the theoretical literature and Wolfinger and Rosenstone (1980) for an influential empirical study.

the decision. Instead we rely on the use of survey instruments, such as the American National Election Study (ANES) or the Current Population Survey (CPS), that includes both measures of relevant characteristic and the respondent's *recalled* voting behavior. The problem is that recall is hardly perfect: Respondents sometimes report recalling to vote when in fact they did not.

The evidence that misreporting is a problem in the recalled turnout data can be found in a series of validation studies that the ANES conducted in 1964, 1976, 1978, 1980, 1984, 1988 and 1990. The validation studies were possible, but expensive, because voting is matter of public record, although for whom a voter voted is not. After administering a post-election survey to a respondent an official from ANES was sent to the respondent's local registrar of elections to see if in fact they were recorded as having voted in the election. This is not an easy task, since the respondent's often do not know where they voted, election officials differed in their ability to produce the records in a usable form, as well as difference is spelling of names. This means that the validated data may also be mis-measured, but for this paper we will assume it is correct. That said, the ANES for these years included both the respondent's *recalled* vote and the *validated* vote. The differences between the two measures are fairly shocking; depending on the election year between 8 and 13 percent of the respondents claiming to have voted did in fact not according to the public records. This finding lead to a cottage industry analyzing the causes of misreporting (Abramson and Claggett 1984, 1986, 1991; Hill and Hurley 1984; Katosh and Traugott 1981; Sigelman 1982; Silver, Anderson and Abramson 1986; Weir 1975) and even to a debate about how to best measure misreporting (Anderson and Silver 1986). All of these studies find that misreporting varies systematically with characteristics of interest, but none offers a complete characterization of when this misreporting will be problem for inference or provide an estimation solution to correct for possible misreporting.

Respondents misreporting of voting behavior may be an honest mistake or because there are social pressures to report voting. There is some evidence for the latter explanation because almost no respondent misreported not voting when in fact she had. For the purpose of estimating the conditional distribution of turnout, the purpose of these studies, the cause of misreporting matters little.

The open question then is what to do about the problem of respondents misreporting. One option would be to use only validated data. At some level this is an appealing option. If we are sure that the validated data is the correct, then estimation and inference straight forward. Collecting the validated data, unfortunately, is difficult and costly. The ANES has stopped collecting the validated turnout data for this reason. Even if it were free, some states, such as Indiana, make it impossible to validate vote. If we are going to limit ourselves to only fully validated data, then our sample will be much smaller. We are also throwing away the useful information in the already collected but non-validated data.

In this paper we develop an estimation model that allows researchers to continue to use the self-reported data in a manner that will produce correct estimates and inferences

even in the presence of misreporting. The method relies on auxiliary information to estimate the conditional probability of misreporting. The model is developed in the context of estimating the conditional probability of turning out to vote, but the method is general and will be applicable whenever there is systematic misreporting of a discrete outcome in a survey. Another likely example of misreporting in surveys of participation in social programs. Since there is often a social stigma to receiving government aid, we might expect to see respondents misreporting their participation in a program.

The key to the model is the incorporation of the probability of misreporting into the specification. This information need not come from the sample of interest used to estimate the turnout probabilities. In the example presented here we will use the earlier validation studies to estimate misreporting probabilities. However, even small pilot validation studies and possibly even aggregate data might be used to gain this information. The estimation is done via Markov Chain Monte Carlo simulation, that will all us to also simultaneously address another problem with survey data, namely missing data.

The paper proceeds as follows. The next section formally lays out the estimation problem in the presence of misreporting and presents some findings about how bad the problem is in practice. In Section 2, we develop the proposed estimation solution. In Section 3, we provide two applications of our methodology. First, we fit a model of voter turnout accounting for misreporting with data from all the National Elections Studies between 1978 and 1990 for which vote validations were undertaken, and compare the results with those obtained using both the recalled and the validated vote for those years. Second, using information on misreport probabilities from the 1988 and 1990 ANES, we apply our methodology to a model voter turnout for the 1992 Presidential and 1994 Midterm elections, for which no validated data is available. In both applications, we compare the results from a complete-case analysis with those obtained using Bayesian imputation methods for dealing with missing data. Finally, Section 4 concludes.

1. DEFINING AND MEASURING THE PROBLEM

Let y_i be an indicator (dummy) variable that is 1 if individual i voted and 0 otherwise and let \mathbf{x}_i be a vector of associated characteristics that influences her choice.² We want to estimate the conditional distribution of y_i given \mathbf{x}_i , $\Pr[y_i|\mathbf{x}_i]$. Generally y_i is unobserved. Instead we observe the recalled vote, \tilde{y}_i . \tilde{y}_i is coded as 1 if respondent i recalled voting in the election and 0 otherwise.

Most studies use this recalled turnout measure, typically running either a probit or logit model to estimate $\Pr[\tilde{y}_i = 1|\mathbf{x}_i]$. We need to know the relationship between $\Pr[\tilde{y}_i = 1|\mathbf{x}_i]$ and $\Pr[y_i = 1|\mathbf{x}_i]$ in order to know when this substitution leads to correct

²As will be clear as the argument proceeds, it extends easily to multichotomous choice although not to the case of continuous responses.

inferences. We can always write

$$\Pr[\tilde{y}_i = 1|\mathbf{x}_i] = \Pr[\tilde{y}_i = 1|\mathbf{x}_i, y_i = 1] \cdot \Pr[y_i = 1|\mathbf{x}_i] + \Pr[\tilde{y}_i = 1|\mathbf{x}_i, y_i = 0] \cdot \Pr[y_i = 0|\mathbf{x}_i]$$

by using standard accounting identities of probabilities. All that we have done is rewrite the probability that the respondent recalled voting into two components: when they have and have not actually voted. Also noting that $\Pr[\tilde{y}_i = 0|\mathbf{x}_i, y_i = 1] = 1 - \Pr[\tilde{y}_i = 1|\mathbf{x}_i, y_i = 1]$ we can re-write the relationship as

$$\Pr[\tilde{y}_i = 1|\mathbf{x}_i] = (1 - \pi_i^{01} - \pi_i^{10}) \Pr[y_i = 1|\mathbf{x}_i] + \pi_i^{10} \tag{1}$$

where $\pi_i^{01} = \Pr[\tilde{y}_i = 0|\mathbf{x}_i, y_i = 1]$ is the probability that the respondent falsely responds not voting when she did and $\pi_i^{10} = \Pr[\tilde{y}_i = 1|\mathbf{x}_i, y_i = 0]$ is the probability the voter falsely claims to have voted when in fact she did not. It is important to note that the probability of both types of misreporting is conditional on \mathbf{x}_i .

The conditional distribution of y_i given \mathbf{x}_i is usually considered known up to some set of parameters θ , for example the logit or probit specifications, that are estimated via maximum likelihood or as the case here we will estimate complete posterior distribution using Markov Chain Monte Carlo (MCMC) methods. The question, then, is under what conditions will maximizing the likelihood function defined by $\Pr[\tilde{y}_i = 1|\mathbf{x}_i; \theta]$ to estimate θ yield equivalent estimates as the true model $\Pr[y_i = 1|\mathbf{x}_i; \theta]$? By examining Equation 1, it obvious that if $\pi_i^{01} = \pi_i^{10} = 0 \quad \forall i$ then the two likelihood functions are identical and must therefore lead to same estimates. In other words, when there is no response error our estimation works as expected.

The estimates, however, will also be equivalent under a more general assumption. If it is the case that $\pi_i^{10} = \pi^1$ and $\pi_i^{01} = \pi^2 \quad \forall i$, where π^1 and π^2 are constants — clearly in $[0,1]$ since they are probabilities — the ML estimates will also coincide between the true and recalled vote. This is the case where there is *random* response error that is independent of the characteristics, \mathbf{x}_i , we are interested in. In this case the the likelihood function for the recalled vote is just a monotonic transformation of the likelihood function for the true vote. From standard results about maximization, monotonic transformation do not alter the optimization problem and lead to identical results. These are asymptotic results, that is the ML estimator $\hat{\theta}$ from the recalled data is consistent in the presence of random response error. However, if there is large amounts of random response error, we might need very large samples to be sure that we were making correct inferences. In principle, this problem could be by using a Bayesian approach based on Gibbs sampling, which allows obtaining arbitrarily precise approximations to the posterior densities without relying on large-sample theory (Albert and Chib 1993).

1.1. An illustrative example

Thus, we know the conditions under which misreporting will be a problem theoretically: when the probability of misreporting varies systematically with characteristics we

are interested in conditioning on. The question, however, is this a problem in practice. Will our inference using recalled as opposed to validated vote actually differ? Some evidence that this is a real problem can be found by fitting a simple model of turnout using data from all the ANES studies for which vote validations were undertaken since 1978. This dataset comprises three Midterm (1978, 1986, 1990) and three Presidential elections (1980, 1984, 1988).³ As can be seen in Figure 1, there are considerable differences between turnout computed using recalled and validated data; validated turnout is systematically lower in all the ANES studies under analysis, and while both rates tend to follow similar trends, differences vary considerably across years, ranging from 6 percentage points in 1980 to almost 12 percentage points in 1978.

However, the real test to see if the conditional results vary. Using these validated data, we fit a hierarchical probit model allowing for election year and regional effects with both self-reported and validated turnout as the response variable:

$$\Pr[\tilde{y}_i = y_i^{Reported}] \sim \text{Bernoulli}(p_i);$$

$$p_i = \Phi(\alpha_t + \gamma_r + \beta' \mathbf{x}_i)$$

and

$$\Pr[y_i = y_i^{Validated}] \sim \text{Bernoulli}(p_i);$$

$$p_i = \Phi(\alpha_t + \gamma_r + \beta' \mathbf{x}_i);$$

where the $k=1, \dots, K$ elements of β are assigned diffuse prior distributions:

$$\beta_k \sim N(\mu_\beta, \sigma_\beta^2)$$

and α_t and γ_k are election- and region- random effects distributed

$$\alpha_t \sim N(\mu_\alpha, \sigma_\alpha^2), \quad t = 1978, 1980, 1984, 1986, 1988, 1990;$$

$$\gamma_r \sim N(\mu_\gamma, \sigma_\gamma^2), \quad r = \text{Northeast, North Central, South, West}$$

The regressors included in \mathbf{x} are indicators for the demographic and socio-economic conditions; a complete list of the variables used in the analysis may be found in Appendix A. We should note that while this specification is similar to most found in the literature, it does not examine other we might plausibly believe alter turnout behavior — for example, any role for political information (Alvarez 1997) or differences in state-level ballot laws Wolfinger and Rosenstone (1980). The samples used in the analysis consists of 6452 observations for the 6 elections under study and were constructed so that they are identical for both models: only the respondents who answered the turnout question in each election study and for whom the ANES staff could validate their vote are included in the results. The remaining observations were dropped using list-wise deletion,

³We use data from the 1978-1990 validated ANES in order to preserve the comparability regarding the questions about the conditions of the interview; specifically, items concerning the interviewers' perceptions about respondents' cooperation and sincerity levels. We will use this information in the application of our model of misreporting in Section 3.

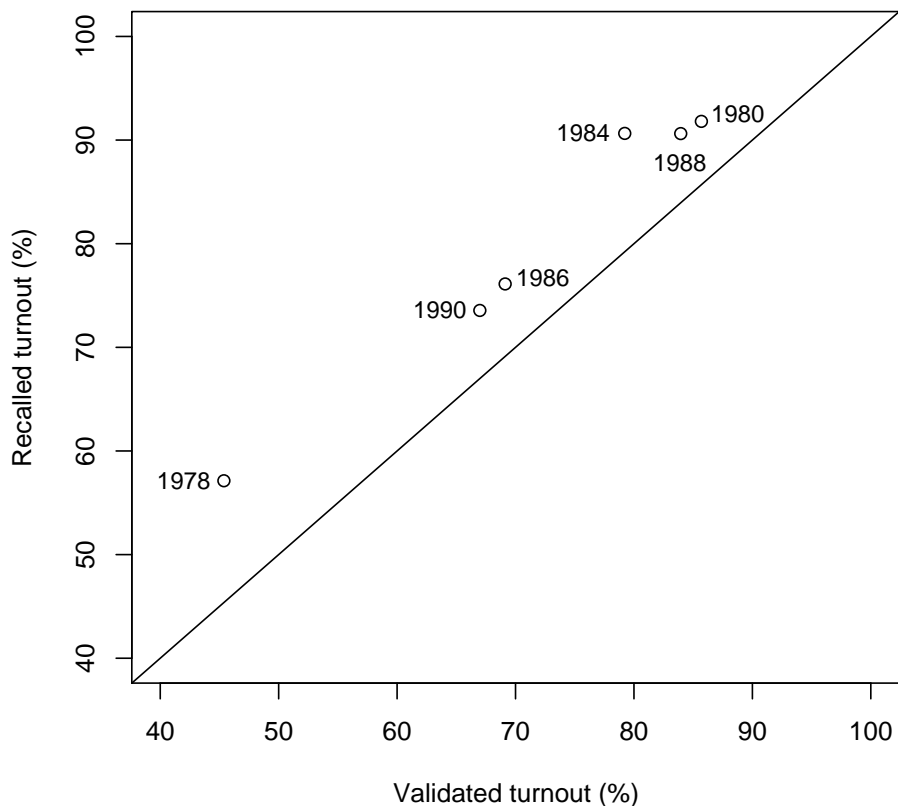


Figure 1: *Estimated Turnout from Recalled vs. Validated Responses 1978–1990.* The graph shows the reported and validated turnout from the 1978–1990 ANES only in years for which there were vote validation studies. Recalled turnout rates are systematically larger than the validates ones.

generally not a good idea (Little and Rubin 1987), but easy to implement and often done in practice; we will deal with the problem of missing data in Section 3.3.

Figure 2 presents the main results of the models. The left panel summarizes the posterior distributions of the model’s coefficients using recalled vote as the dependent variable, and the right panel re-does the analysis with the ANES validated vote. The posterior distribution of the coefficients of most predictors are quite similar in both models, and inferences drawn regarding the role of these predictors on the probability of voting agree with common expectations. For example, for both sets of estimates, wealthier and more educated respondents are more likely to turn out to vote: a respondent whose household belongs to the fifth quintile of income is approximately 7.4 percentage points more likely to vote than one whose household belongs to the first quintile, while a respondent with university education is 0.28 more likely to vote than one with 8 years of schooling or less. Likewise, respondents are much more likely to turn out to vote in

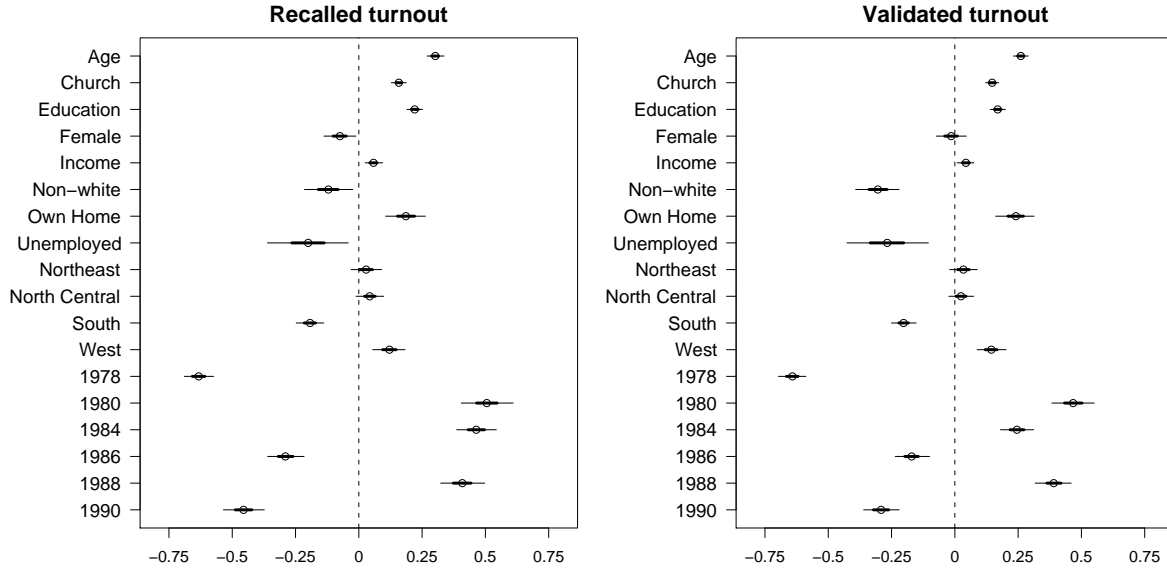


Figure 2: *Coefficients of the probit model from Recalled vs. Validate Turnout. The graph summarizes the posterior distribution of the coefficients of the turnout model, using recalled and validated turnout as the response variable. The center dots correspond to the point estimates, the thicker lines to the 50% confidence interval, and the thinner lines to the 90% confidence interval.*

Presidential than in Midterm elections, and are less likely to vote if they live in the South: the average marginal effect of a Presidential election is to raise the probability of voting by 16 percentage points, while living in the South lowers the probability of voting by 6.2 percentage points. These results are essentially similar using either recalled or validated vote as the dependent variable, and there are also coincidental findings in the two models for the role of age, church attendance, education and employment status: the magnitudes of the estimated marginal effects of these predictors on voter turnout are roughly similar for both sets of data.

However, there are some interesting differences between the two sets of results regarding the role of some socio-demographic variables such as gender and race. For instance, while the coefficient for *Female* is negative and significant at the 0.05 level in the model for recalled vote, it is not statistically significant even at the 0.5 level in the model for validated vote. Also, the mean posterior of the coefficient for the race indicator is more than 2.5 times larger (in absolute value) using validated vote than using recalled vote as the response variable. These differences lead to substantial divergences in the inferences drawn from both models regarding the impact of these variables on voter turnout. In order to illustrate this fact, Figure 3 plots the marginal effect of race on the probability of voting using recalled and validated vote for each election under analysis. As seen in the figure, the impact of race on voting is considerably higher when validated vote is used as the response variable: the estimated marginal effect is 7 percentage points higher than if we look only at the recalled vote data, and this difference reaches almost 11 percentage

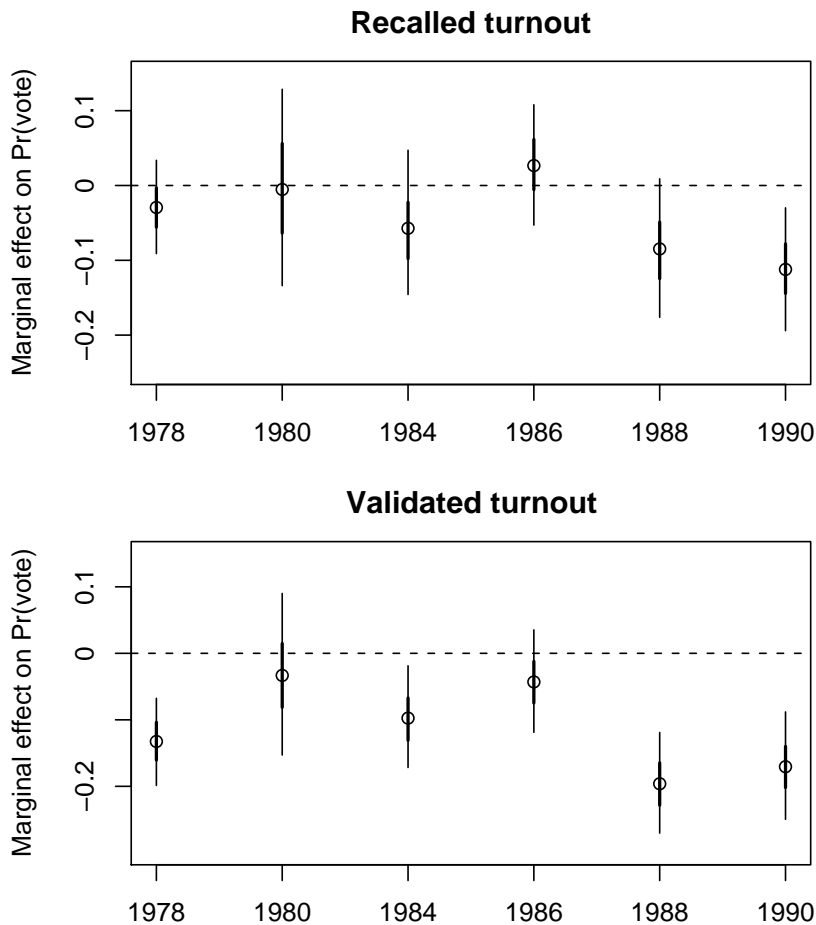


Figure 3: *Marginal effect of race on turnout.* The graph shows the marginal effect of the race indicator on the likelihood of voting for each year under study, using both recalled and validated vote. The center dots correspond to the point estimates, the thicker lines to the 50% confidence interval, and the thinner lines to the 90% confidence interval.

points in the 1978 and 1988 elections. Moreover, the substantive conclusions drawn from the analysis would also change considerably: while a researcher would conclude that race had no significant effect on the probability of voting in the 1978, 1984 and 1988 elections at the usual confidence levels, the results obtained using validated data indicate that it did in fact matter. Hence, fitting a model of turnout using recalled vote as the response variable will generally tend to overpredict the probability of voting among non-white respondents.

Thus, there is direct evidence that the probability of misreporting does vary systematically with characteristics we might be interested in, at least using the ANES data between 1978 and 1990. Any study presenting findings using the recalled vote measure should be considered suspect. Unfortunately, the ANES has stopped collecting the validated data because of the cost and difficulty in collecting the data as well as the fact that few researches used the validated data. We need, therefore, a model that can use the

available data to correct the estimates using the recalled vote data. This is the subject of the next section.

2. A POSSIBLE SOLUTION

If we must use the recall vote that is subject to misreporting, we need to develop a model that accounts for this error in order to correctly estimate $\Pr[y_i|\mathbf{x}_i]$. Since recalled data is dichotomous, we can start by assuming the observations conditional on observable characteristics are independently and identically distributed according to a Bernoulli distribution — just as in the standard logit or probit. We can therefore write the probability of the sample as

$$\prod_{i=1}^N \Pr[\tilde{y}_i|\mathbf{x}_i, \theta]^{\tilde{y}_i} (1 - \Pr[\tilde{y}_i|\mathbf{x}_i, \theta])^{1-\tilde{y}_i}$$

We will further assume that $\Pr[y_i = 1|\mathbf{x}_i] = F(\beta' \mathbf{x}_i)$, where $F(\cdot)$ is some cumulative density function. For ease of exposition, we will assume that $F(\cdot)$ is the standard normal distribution denoted by $\Phi(\cdot)$. This will lead to a probit model with a correction for misreport. The use of the logistic cumulative density function would lead to a logit model with a correction for misreporting. We can substitute for $\Pr[\tilde{y}_i|\mathbf{x}_i, \theta]$ in Equation 1 arriving at

$$\begin{aligned} \mathcal{L}(\theta|\mathbf{Y}, \mathbf{X}) = \prod_{i=1}^N & \left[(1 - \pi_i^{0|1} - \pi_i^{1|0})\Phi(\beta' \mathbf{x}_i) + \pi_i^{1|0} \right]^{\tilde{y}_i} \\ & \times \left[(1 - \pi_i^{1|0} - \pi_i^{0|1})(1 - \Phi(\beta' \mathbf{x}_i)) + \pi_i^{0|1} \right]^{1-\tilde{y}_i} \end{aligned}$$

which represents the probability of observing the sample under misreporting. Taking the logs in the above expression, we arrive at the log likelihood function for the sample, given by:

$$\begin{aligned} \ln \mathcal{L}(\theta|\mathbf{Y}, \mathbf{X}) = \sum_{i=1}^N & \left[\tilde{y}_i \ln[(1 - \pi_i^{0|1} - \pi_i^{1|0})\Phi(\beta' \mathbf{x}_i) + \pi_i^{1|0}] + \right. \\ & \left. (1 - \tilde{y}_i) \ln[(1 - \pi_i^{1|0} - \pi_i^{0|1})(1 - \Phi(\beta' \mathbf{x}_i)) + \pi_i^{0|1}] \right], \end{aligned}$$

This log likelihood function is strikingly similar to the standard probit specification. The observations are essentially weighted according to how likely individual i misreported her vote. The parameters that we need to estimate from the data are $\theta = \{\pi_i^{1|0}, \pi_i^{0|1}, \beta'\}$. Unfortunately, $\pi_i^{1|0}$ and $\pi_i^{0|1}$ are not identified if the only data we have is recalled vote, since they represent the probabilities that a respondent misreports her vote. Hence,

when estimating the model using a Bayesian approach, we would typically assign non-informative priors to $\pi_i^{1|0}$ and $\pi_i^{0|1}$. However, we can use auxiliary, or out of sample, data, to fit the model.

Suppose, for example, we were interested in estimating this turnout model with the 1994 ANES data for which there is no validation study. We could fit a model of the misreport probabilities using data from (some or all of) the ANES validated studies, say, via a probit model. Let d_j^1 be a dummy variable that is true when the respondent j falsely recalls voting when she did not, d_j^2 a dummy variable that is true when j recalls not voting when she did, and \mathbf{z}_j^1 and \mathbf{z}_j^2 denote sets of regressors that are useful in predicting the two types of misreports — where the notation allows for the fact we may use different regressors to predict the two types of misreporting. We could then estimate the misreport probabilities fitting the probit models:

$$\begin{aligned}\Pr[d_j^1] &\sim \text{Bernoulli}(\pi_j^{1|0}); \\ \pi_j^{1|0} &= \Phi(\gamma_1' \mathbf{z}_j^1)\end{aligned}$$

and

$$\begin{aligned}\Pr[d_j^2] &\sim \text{Bernoulli}(\pi_j^{0|1}); \\ \pi_j^{0|1} &= \Phi(\gamma_2' \mathbf{z}_j^2)\end{aligned}$$

with non-informative prior distributions

$$\begin{aligned}\gamma_{l,1} &\sim N(\mu_{\gamma_1}, \sigma_{\gamma_1}^2), \quad l = 1, \dots, L; \\ \gamma_{m,2} &\sim N(\mu_{\gamma_2}, \sigma_{\gamma_2}^2), \quad m = 1, \dots, M;\end{aligned}$$

for the elements of γ_1 and γ_2 .

We can incorporate this auxiliary information on misreport probabilities in our turnout model for 1994. In order to do so, we use a Bayesian approach, fitting our turnout model via MCMC simulations. This approach has two basic advantages in this setting. First, we can include the results from previous statistical studies by repeatedly using Bayes' Rule. Second, MCMC simulations directly account for the extra uncertainty in the variance of β caused by substituting estimates of the misreport probabilities instead of their true values. In the context of maximum likelihood estimation, accounting for the added uncertainty introduced by substituting $\pi_i^{1|0}$ and $\pi_i^{0|1}$ with their estimates $\hat{\pi}_i^{1|0}$ and $\hat{\pi}_i^{0|1}$ would require additional “post-estimation” steps, such as bootstrapping or applying the results of [Murphy and Topel \(1985\)](#) for two-step estimators.

Let i index respondents in the the 1994 ANES study, j index respondents in a previous validated study, and M, N denote the sample sizes in the validated and non-validated studies, respectively. Then, assuming conditional independence throughout, the joint posterior density of the unknown parameters in our turnout model for 1994 is given by:

$$\begin{aligned}p(\beta, \gamma_1, \gamma_2 | \mathbf{Y}, \mathbf{X}, \mathbf{D}^1, \mathbf{D}^2, \mathbf{Z}^1, \mathbf{Z}^2) &\propto p(\mathbf{Y} | \beta, \gamma_1, \gamma_2, \mathbf{X}, \mathbf{Z}^1, \mathbf{Z}^2) \times p(\mathbf{D}^1 | \gamma_1, \mathbf{Z}^1) \\ &\times p(\mathbf{D}^2 | \gamma_2, \mathbf{Z}^2) \times p(\beta) \times p(\gamma_1) \times p(\gamma_2); \end{aligned} \quad (2)$$

and, from the arguments above, it can be written as

$$\begin{aligned}
p(\beta, \gamma_1, \gamma_2 | \mathbf{Y}, \mathbf{D}^1, \mathbf{D}^2, \mathbf{X}, \mathbf{Z}^1, \mathbf{Z}^2) &\propto \prod_{i=1}^N \left[\left[(1 - \Phi(\gamma_1 \mathbf{z}_i^1) - \Phi(\gamma_2 \mathbf{z}_i^2)) \Phi(\beta \mathbf{x}_i) + \Phi(\gamma_1 \mathbf{z}_i^1) \right]^{\tilde{y}_i} \right. \\
&\quad \times \left. \left[(1 - \Phi(\gamma_1 \mathbf{z}_i^1) - \Phi(\gamma_2 \mathbf{z}_i^2)) (1 - \Phi(\beta \mathbf{x}_i) + \Phi(\gamma_2 \mathbf{z}_i^2)) \right]^{1 - \tilde{y}_i} \right] \\
&\quad \times \prod_{j=1}^M \left[\left[\Phi(\gamma_1 \mathbf{z}_j^1) \right]^{d_j^1} \times \left[1 - \Phi(\gamma_1 \mathbf{z}_j^1) \right]^{1 - d_j^1} \right] \\
&\quad \times \prod_{j=1}^M \left[\left[\Phi(\gamma_2 \mathbf{z}_j^2) \right]^{d_j^2} \times \left[1 - \Phi(\gamma_2 \mathbf{z}_j^2) \right]^{1 - d_j^2} \right] \\
&\quad \times \prod_{k=1}^K N(\beta_k | \mu_\beta, \sigma_\beta^2) \times \prod_{l=1}^L N(\gamma_{l,1} | \gamma_1, \sigma_{\gamma_1}^2) \\
&\quad \times \prod_{m=1}^M N(\gamma_{m,2} | \gamma_2, \sigma_{\gamma_2}^2).
\end{aligned} \tag{3}$$

Although Equation 3 is intractable analytically, inference on β , γ_1 and γ_2 can be performed using Gibbs sampling to repeatedly draw samples from each unknown parameter’s full conditional posterior distribution in order to form the corresponding marginal distributions (Gelfand and Smith 1990; Casella and George 1992). While the corresponding conditional posterior densities have no closed forms (see Appendix B), draws of β , γ_1 and γ_2 can be obtained using Adaptive Rejection Sampling (ARS) (Gilks and Wild 1992). Under mild regularity conditions (Gilks, Richardson and Spiegelhalter 1996), for a sufficiently large number of iterations, samples from these complete conditionals approach samples from the marginals used for Bayesian inference. The means and standard deviation of the convergent Gibbs samples can be used to summarize the posterior distributions of the model’s coefficients and to compute the marginal effects of the regressors on the probability of voting through “average predictive comparisons” (Gelman and Hill 2007).

Thus, we need only to have done a validation study at some point in order to account for misreporting in our model of voter turnout for 1994, although we must maintain the assumption that the process generating misreporting has not changed since the time the validation study was conducted. Even if we did not have access to any validation study, there still may be aggregate information available on misreporting that could be used to improve estimates of turnout. For example, we can observe turnout in small geographic areas, such as counties or congressional districts. We would set the misreport probabilities for all individuals from a given area, so that the estimated sample turnout for the area matched the known actual turnout.

3. ESTIMATING TURNOUT ACCOUNTING FOR MISREPORTING

In this section we provide two applications of the methodology developed in Section 2. First, we re-estimate the model of turnout presented in Section 1 for the 1978 – 1990 ANES using self-reported vote data but with the correction for misreporting developed above. Applying our correction for misreporting to data from the 1978–1990 ANES studies, however, is a somewhat roundabout way to estimate a turnout model: in this sample we have the validated vote making, direct estimation possible without having to concern ourselves with misreporting problems. In typical applications we would not have such a validated vote. Instead, we would use the estimates of the misreporting probabilities from an earlier validated or pilot study to correct the self-reported vote. We do this in Section 3.2, applying our correction for misreporting to the 1992 Presidential election and the 1994 Midterm election, for which there is no validated data available, and comparing the results obtained using self-reported vote with those of the corrected model.

Both applications are based on a complete-case analysis — i.e., using list-wise deletion. We deal with the problem of incomplete data in Section 3.3, where we fit the models using a Bayesian procedure to impute the values of the missing variables (Ibrahim et al. 2005). In all cases, we focus on the marginal effect of race on voter turnout since, as evidenced in Section 1.1, inferences drawn about the impact of race on the probability of voting differ substantially when using recalled and validated vote.

Although our model accounts for the possibility of two types of misreporting with probabilities $\pi_i^{1|0}$ and $\pi_i^{0|1}$, almost no one ever reports not voting when they had in the ANES studies: in the 1978 – 1990 surveys, only 46 respondents — 0.8% of the respondents in the sample — claimed to not have voted when the official record suggested they did. This fact lends some credence to social pressures argument for misreporting and should help alleviate some of our concerns about inaccurate records; why should errors only be of false positives, reporting voting when the official record contradicts this claim? Some of the codebooks — e.g., in the 1988 study — even suggest that these cases are probably errors. Therefore, in the applications of our model we will assume that $\pi_i^{0|1} = 0$, and we therefore only need to account for $\pi_i^{1|0}$. The models were fit using WinBUGS 1.4, as called from R 2.4.1.⁴ All the hyperparameters were assigned diffuse priors in order to let the data dominate the form of the posterior densities: the scalars μ_β and μ_{γ_1} were set to 0, while σ_β and σ_{γ_1} were set to 10^4 . In order to ensure that inferences are data dependent, several alternative values for the hyperparameters were tried, yielding essentially similar results. Three parallel chains with dispersed initial values reached approximate convergence after 5,000 iterations, with a burn-in of 2,000 iterations.⁵

⁴The code is available from the authors upon request.

⁵ Approximate convergence is achieved for values of Gelman and Rubin’s estimated Potential Scale Reduction factor below 1.1 (Gelman and Rubin 1992).

3.1. *Re-estimating Turnout for the 1978–1990 ANES*

Estimating the corrected model for the 1978 – 1990 data has the obvious benefit of allowing us to directly compare the corrected estimates to a known benchmark, the same model estimated directly on the validated data. Presumably these estimates from the validated vote are the “correct” estimates, although there is some potential debate on this point.

The concern is that the validation studies are far from perfect. As stated at the outset, vote validation is expensive and difficult. The ANES is conducted in two parts, a pre- and post- election survey. In the studies from 1978, 1980, 1984, 1986, 1988 and 1990 there were in total 11,632 completed post election surveys. Unfortunately of these completed surveys, the ANES was unable to validate 3,097 respondents, about 26.6 percent of the usable sample.⁶ The majority of these cases were caused either because no registration records were found or because the local election office refused to cooperate with the ANES. If we are willing to maintain the assumption that these errors are essentially random (in the sense of being independent of the characteristic of interest), then there is no real harm done. The measurement error will merely result in less efficient estimates of the misreporting model and a corresponding reduction in efficiency of the corrected turnout model. However, if there is systematic error, then we are just substituting one form of measurement error for another.

If you recall from Figure 3 in Section 1, there was a substantial difference between from the recalled and validated vote regarding the effect of race on turnout decisions. This result is due largely to the fact that many more nonwhite respondents are classified as misreporting voting compared to their white counterparts: only 8.9 percent of the whites in the sample are classified as misreporting where as 16.4 of nonwhites, and this difference holds across all election years. If it is the case that nonwhites, whom are more likely to be from poorer areas, are more likely to be incorrectly validated — or excluded from the validation study because no records can be found — because these elections offices are more likely to be poorly staffed and maintained, then this result could very well be an artifact. It is difficult to rule this claim out. [Abramson and Claggett \(1984\)](#) argue that the difference in misreporting are least partially caused by real differences and are not just mistakes in validation. Their argument is that black respondents in surveys in other contexts are more likely to give socially acceptable answers, for example having an opinion on a relatively obscure congressional bill. Further, given that blacks had to struggle to be granted suffrage, they fell more obligated to appear to have voted. In the end, it will be difficult to directly answer this concern. We will operate under the assumption that validation data is correct, or at least not subject to systematic bias. There is some information contained in these studies about the local elections office which might be used to indirectly test this claim, but we will leave that to others to debate and analyze.

⁶The rate of non-validation varies considerably across Election Studies, from around 3 percent of sample in 1978 to almost 39% in 1990.

Instead we will proceed in estimating the corrected model of turnout. Since we will need a model of the misreporting to correct the estimates, we will first estimate such a model from the validated study. We will then use these to correct the estimates of the turnout model using the recalled vote.

Misreport probabilities Estimates from a model of the probability misreporting are presented in Figure 4. As with the turnout model presented in the first section, the model is fairly simple. It consists almost exclusively of easily observable demographic characteristics. In order to potentially improve fit, we also include three covariates that tried to at least tap into the social pressure argument. The first is an indicator of whether the interview was conducted while the respondent was alone. According to the social pressure argument, a respondent should be more likely to lie about voting if others will learn of the statement. Second, we included two indicators for the interviewer’s evaluation of the respondent after the post-election survey. All interviewers were asked to rate the cooperativeness and sincerity of the respondent after the completion of the survey.

The results of the estimation are for most part what one would expect. Race has the highest effect on the probability of misreporting: non-whites are on average 0.12 more likely to misreport turning out to vote. This finding should not come as a big surprise: from the comparison of the estimates using recalled versus validated vote, race was the only variable for which there were substantial differences regarding its marginal effect on the probability of voting. The only other variable that has a significant effect on the probability of misreporting at the 0.05 level is *Own Home*: home owners are on average 4.4 percentage points less likely to misreport their vote. None of the other variables has a statistically significant effect on misreporting at the 0.05 level; in particular, the interviewers seem unable to pick up a “feeling” that is not otherwise captured by the characteristics observable from the survey. This is probably caused by the fact that very few of the interviewers were willing to rank a respondent as uncooperative and/or insincere.

Perhaps the biggest concern is that the model does not predict misreporting very well: the mean error rate of the model is 41%, while a null model that simply predicts that no respondent misreports has an error rate of 10%. This is not an unexpected finding: given the very few respondents misreport, it is hard to do better than by just predicting the mode that a respondent is not lying. However, what we really need is a model that well separates misreporters from non-misreporters. The results here also seem a bit mixed: the model correctly classifies 61% of the cases, and the mean predicted probability of misreporting averaged across simulations is 0.51; ideally this would be near zero or one for the entire sample. But the proof is really in how well the correction helps the estimates which is where we will turn now.

The effect of race on voter turnout. We estimated the turnout model using the recalled data corrected for misreporting and computed the marginal effect of the race indicator on the probability of voting. Figure 5 compares the average marginal effect using validated

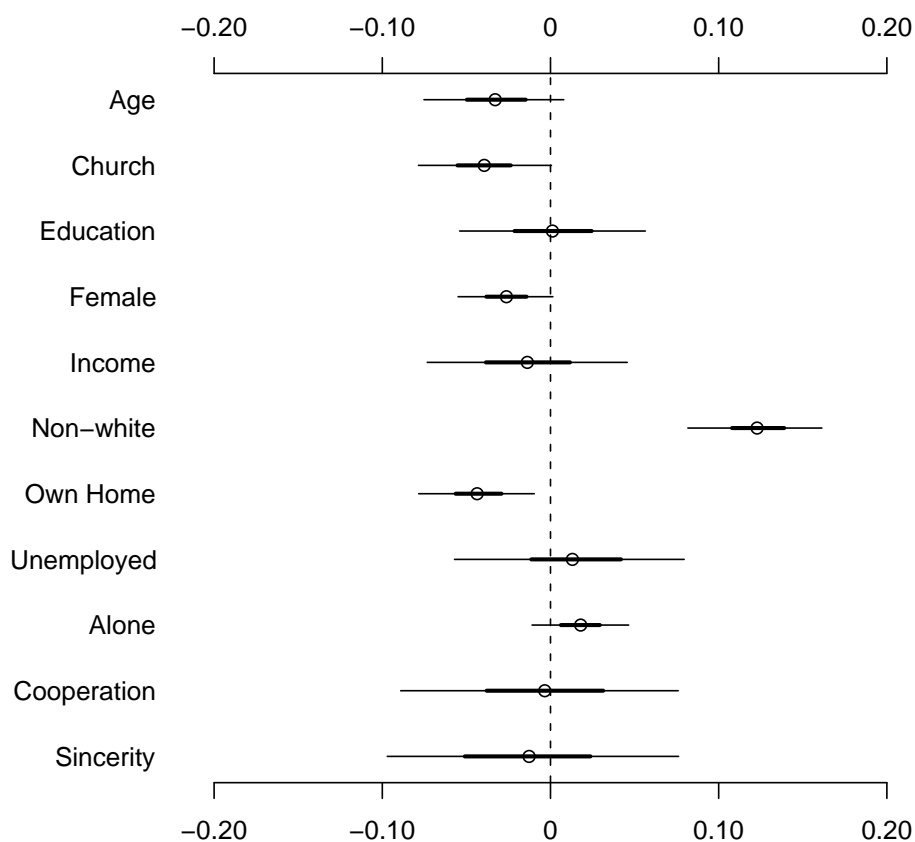


Figure 4: ***Determinants of misreporting.*** The graph shows the marginal effect of each of the regressors on the probability of misreporting. The center dots correspond to the point estimates, the thicker lines to the 50% confidence interval, and the thinner lines to the 90% confidence interval.

and corrected turnout for each year in the sample, and allows us to assess the performance of the correction for misreporting by comparing the results with those obtained using recalled vote. The results show that the correction method does seem to work considerably well, in the sense that, for each of the election-years considered, the marginal effects of race on voting estimated from the corrected self-reports are closer to those obtained using validated data than when estimated from recalled vote. The difference between the marginal effects computed using corrected self-reports and the known benchmark – i.e., validated vote – is 3.9 percentage points for all election-years in the sample, versus 6.8 percentage points when computed from the self-reported vote.

Moreover, comparing the results in Figure 6 with those presented in Figure 3 above, we see that, after correcting for misreporting, the impact of race in the 1978 and 1988 elections is now statistically and substantively significant at the 0.1 level, unlike the esti-

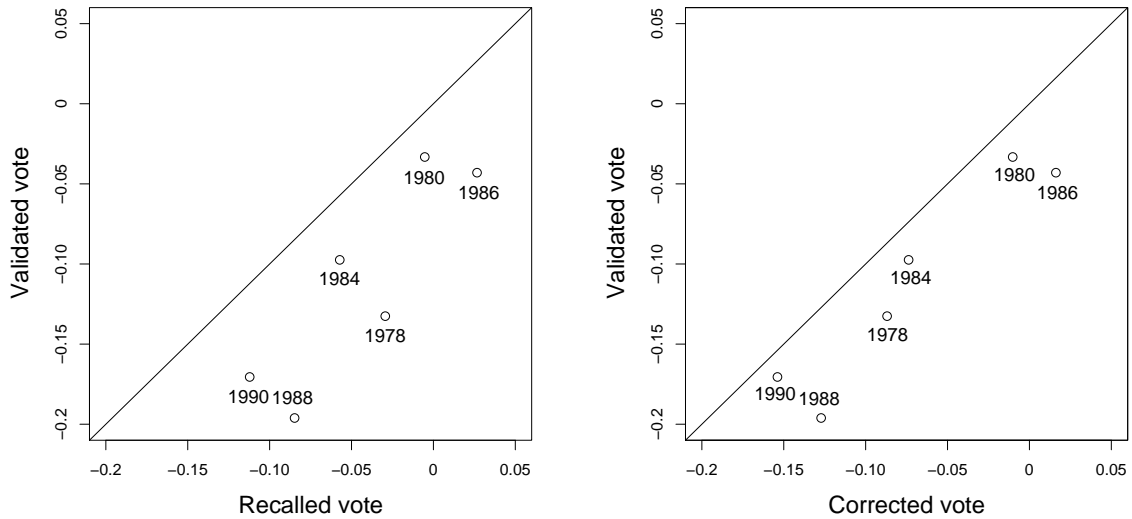


Figure 5: *Marginal effect of race on turnout in the three models.* The left panel of the graph compares the marginal effect of being non-white on the probability of voting using recalled vs. validated vote, and the right panel compares the effect using validated vs. corrected vote.

mates obtained from self-reported vote.⁷ Ideally, the estimated marginal effects obtained using our methodology would be almost identical to those obtained using validated vote; while this is the case for some elections (e.g., 1980, 1984, 1990), differences can be relatively large for some others (e.g., 6.9 percentage points for 1988). Nonetheless, these results would presumably improve further with a better model of misreporting. Even with the rather simple model estimated here, however, the improvements in terms of the magnitude of the marginal effects are important.

3.2. Applying the model using auxiliary information

We also apply our correction for misreporting in order to estimate the probability of voting in the 1992 and 1994 elections, for which there is no validated data. As in the model presented in Section 2, we incorporate auxiliary information from previous Election Studies for which validated data is available; specifically, we use the misreport probabilities from the 1988 and 1990 ANES, the last Presidential and Midterm elections for which validated vote is available, in order to fit the turnout model for 1992 and 1994, respectively.⁸

⁷In the case of the 1988 election, the marginal effect of *Non-white* from the corrected model is also significant at the 0.05 level.

⁸As shown above, the probability of voting is considerably higher in Presidential than in Midterm elections. Therefore, in order to avoid additional complexities derived from comparing different types of races, we used the misreport probabilities from the last comparable election for which validated data is

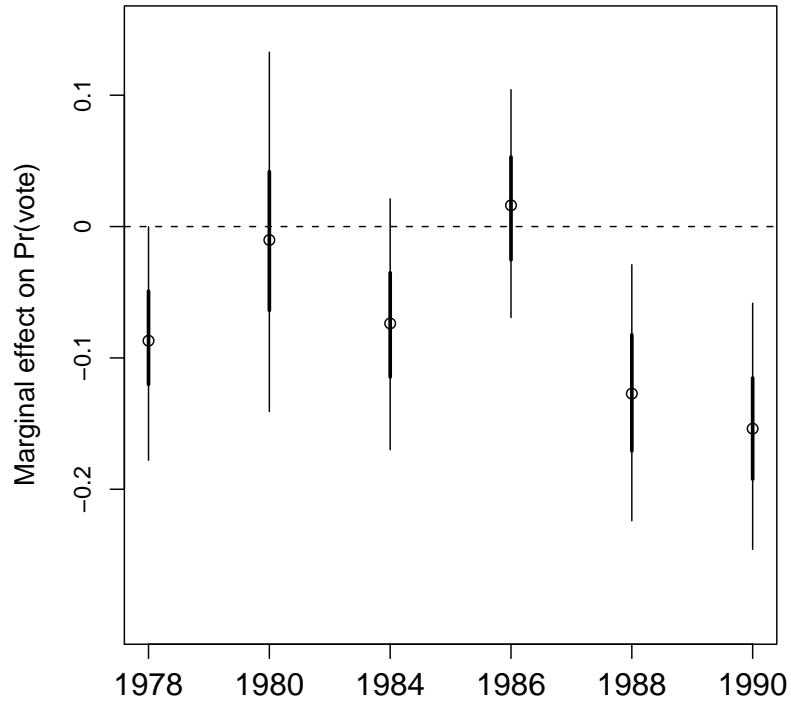


Figure 6: *Marginal effect of race on turnout for each election-year.* The graph shows the marginal effect of race on the likelihood of voting estimated from the corrected model for each election under analysis. The center dots correspond to the point estimates, the thicker lines to the 50% confidence interval, and the thinner lines to the 90% confidence interval.

Figure 7 illustrates the results, comparing the posterior densities of the coefficients of the selected regressors using recalled and corrected vote for the 1992 election. For some of the regressors – e.g., *Church*, *Education*, *Own Home*, – the posterior distribution of the coefficients remain essentially unchanged when applying the correction for misreporting. However, using auxiliary information does substantially affect the posterior distribution of the coefficients of other socio-demographic variables. In particular, and in line with the evidence presented above, accounting for misreporting does substantially affect the distribution of coefficient for *Non-white*: the mean posterior of $\beta_{Non-white}$ is 1.6 times larger (in absolute value) when using the corrected self-reports. More importantly, accounting for misreporting affects once again the substantive conclusions drawn from the turnout model: while inference based on recalled vote suggests that race had no significant effect on the probability of voting in the 1992 election, using the corrected vote would lead to conclude that being non-white in fact reduced the likelihood of voting by 10 percentage

available.

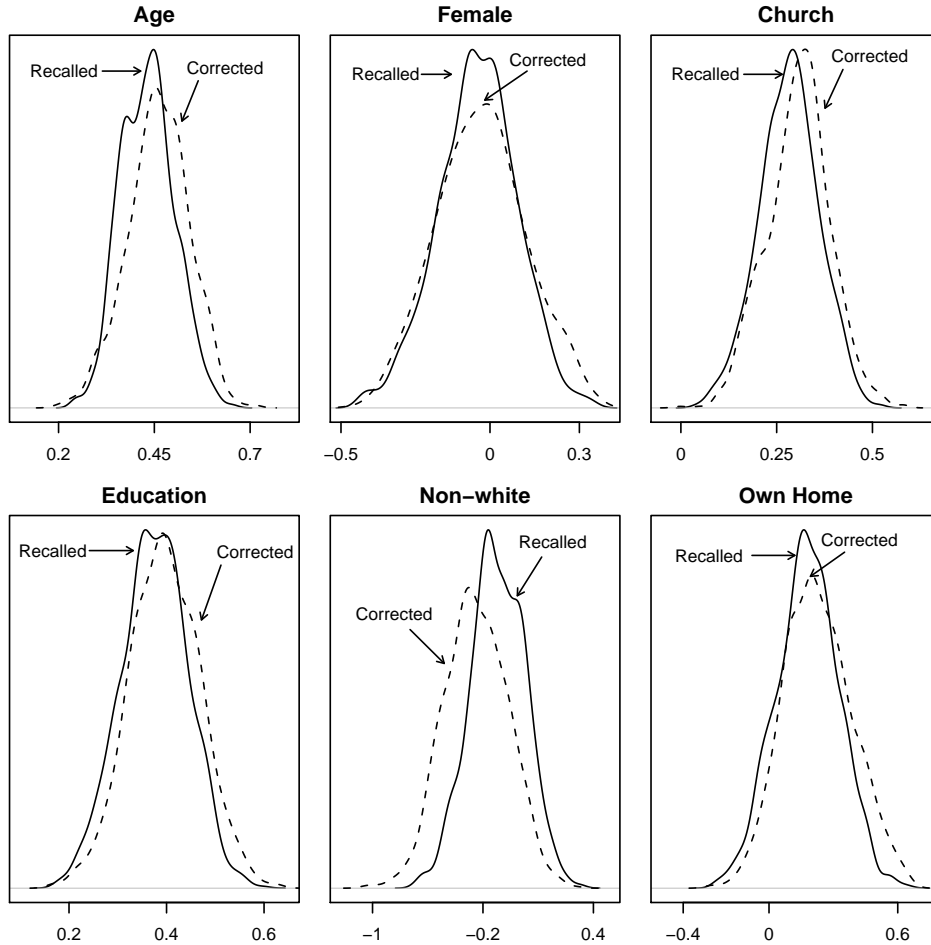


Figure 7: *Posterior densities using self-reported and corrected vote.* The figure compares the posterior densities of selected coefficients using self-reported and corrected vote in the 1992 Presidential election. The solid line plots the posterior distributions obtained using recalled vote, and the dashed lines to the ones obtained from the model accounting for misreporting.

points, and this effect is significant at least at the 0.1 level. Similar results are obtained for the 1994 election.

3.3. Accounting for missing data

Both applications of our methodology in Sections 3.1 and 3.2 have been based on a complete-case analysis, including in the sample only those respondents who are completely observed. This approach has serious drawbacks which have been extensively documented (Little and Rubin 1987; Lipsitz and Ibrahim 1996; Ibrahim et al. 2005). First, simply omitting missing data from the analysis leads to valid inferences if the data are missing completely at random. If, on the other hand, respondents with complete data

are systematically different from those with missing data, a complete-case analysis can produce biased results.⁹ In our sample from the 1978–1990 ANES studies, the percentage of missing observations for non-whites is 18.9% higher than for whites. In addition, complete-case analyses are unnecessarily wasteful and can be quite inefficient; in our case, list-wise deletion due to missing values in the response variable and the predictors leads to discard approximately 45% of the respondents in the 1978–1990 ANES and more than two-thirds of the respondents in the 1994 ANES. Appendix C reports the rates of item nonresponse for all the variables included in the turnout models from Sections 3.1 and 3.2.

Ad-hoc approaches to imputation aimed at keeping the full sample, such as mean imputation, are easy to implement, but exhibit several potential biases (Gelman and Hill 2007). In this section, we use a model-based, Bayesian approach to address these concerns. This will require specifying distributions for the variables with missing data as well as priors on all the parameters; missing variables are then sampled from their conditional distribution through Gibbs sampling. Hence, this approach implies incorporating just an “extra-layer” in the Gibbs steps compared to the complete case analysis, and can thus accommodate missing data without resorting to new techniques for inference. For this reason, Bayesian methods are a powerful and general tool for dealing with missing data (Ibrahim et al. 2005).¹⁰

We fit a separate imputation model for each of the ANES studies, assuming that the data are missing at random (MAR) and that the parameters of the missing-data process are distinct from the parameters of the data model, so that the missing-data mechanism is *ignorable* (Rubin 1976).¹¹ If only the response variable had missing data, we would just specify the model presented in Section 2 and draw a value for each missing value of \tilde{y}_i based on its predictive density (Congdon 2001).¹² In our case, however, we also have missing values in most of the covariates, so an important issue is the specification of a parametric model for the missing covariates (Ibrahim, Lipsitz and Chen 1999). Let $\mathbf{w}_i = (w_{i,1}, \dots, w_{i,P})$ denote the P -dimensional vector of missing covariates in our model for misreporting. Following Lipsitz and Ibrahim (1996) and Ibrahim, Lipsitz and Chen (1999), we model the joint distribution of the missing covariates as a product of one-dimensional parametric conditional distributions.¹³ That is, we write the joint

⁹It is worth mentioning, however, that there are situations in which inference based on a complete-case analysis might yield unbiased estimates and outperform imputation methods even when the data are not missing completely at random (Little and Wang 1996; Ibrahim et al. 2005).

¹⁰A detailed review of the different commonly used model-based imputation methods is beyond the scope of this paper. See Schafer and Graham (2002); Ibrahim et al. (2005); Horton and Kleinman (2007), among others, for a detailed discussion.

¹¹See Gelman, Gary and Liu (1998) for an approach to multiple imputation for multiple surveys using hierarchical modeling

¹²WinBUGS handles missing data in the dependent variable automatically, so imputation of missing data in the response variable \tilde{y}_i as well as in the misreport indicator d_j^1 is straightforward.

¹³Obviously, a model needs to be specified only for those covariates that have missing values. If some of the covariates are completely observed for all respondents in a survey, they can be conditioned on when constructing the distribution of the missing covariates.

distribution of \mathbf{w}_i as:

$$p(w_{i,1}, \dots, w_{i,P} | \alpha) = p(w_{i,P} | w_{i,1}, \dots, w_{i,P-1}, \alpha_P) \times p(w_{i,P-1} | w_{i,1}, \dots, w_{i,P-2}, \alpha_{P-1}) \times \dots \times p(w_{i,1} | \alpha_1) \quad (4)$$

where α_j is a vector of parameters for the j th conditional distribution, the α_j 's are distinct, and $\alpha = (\alpha_1, \dots, \alpha_P)$. Among other advantages, specification 4 eases the computational burden in the Gibbs sampling scheme required for sampling from the posterior distribution, which can be implemented via Gilks and Wild (1992)'s ARS algorithm (Ibrahim, Chen and Lipsitz 2002; Ibrahim et al. 2005). In addition, modeling the joint covariate distribution in this fashion allows us to specify Bernoulli distributions for the dichotomous variables, rather than having to specify a multivariate normal distribution for all covariates, as is generally the case with other imputation procedures (Schafer 1997). In our application, probit regression models were specified for all the dichotomous covariates in the model – *Non-white*, *Own Home*, *Unemployed*, *Alone* –, while the remaining categorical covariates were assigned conditional normal distributions and discrete values were afterward imputed for the missing responses (Lipsitz and Ibrahim 1996; Gelman, Gary and Liu 1998).¹⁴ As in the case of the parameters of the data model, we used non-informative independent normal priors for the components of α_j , $j=1, \dots, P$, and 20,000 Gibbs iterations were used (with a burn-in of 3,000 iterations).

Figure 8 illustrates the results obtained using Bayesian imputation for the 1978 and 1994 ANES, the Election Studies with the lowest and largest rates of item non-response, respectively. The left panel shows the marginal effect of race on the probability of voting for the 1978 ANES using reported, validated and corrected vote. The right panel presents the estimated marginal effects for the 1994 ANES, for which we use auxiliary information from the 1990 ANES in order to correct for misreporting, as mentioned in Section 3.2. In both cases, estimates obtained using Bayesian imputation are compared to those in the complete-case analyses.

Two remarkable facts emerge from the figure. First, for both election-studies, the marginal effect of race on the probability of voting estimated from the Bayesian imputation model is not statistically different from the effect obtained using list-wise deletion, at least at the 0.1 level. However, the standard errors tend to be lower when missing values are imputed, as can be found by comparing the results in the left panel of Figure 8 with those presented in Figures 3 and 6: for instance, while using recalled vote as the dependent variable in the complete-case analysis would lead to conclude that race had no significant effect on voting in the 1978 election, the inference from the Bayesian imputation model indicates otherwise. This result holds in fact for most of the election-years under analysis, suggesting that by omitting the cases with missing values, much information is lost on the variables that are completely or almost completely observed, thus leading to less efficient estimates for the models' coefficients and the marginal effect

¹⁴The substantive results are essentially unchanged if, instead of the normal distributions, one-dimensional conditional gamma distributions are specified for these covariates, all of which are strictly positive.

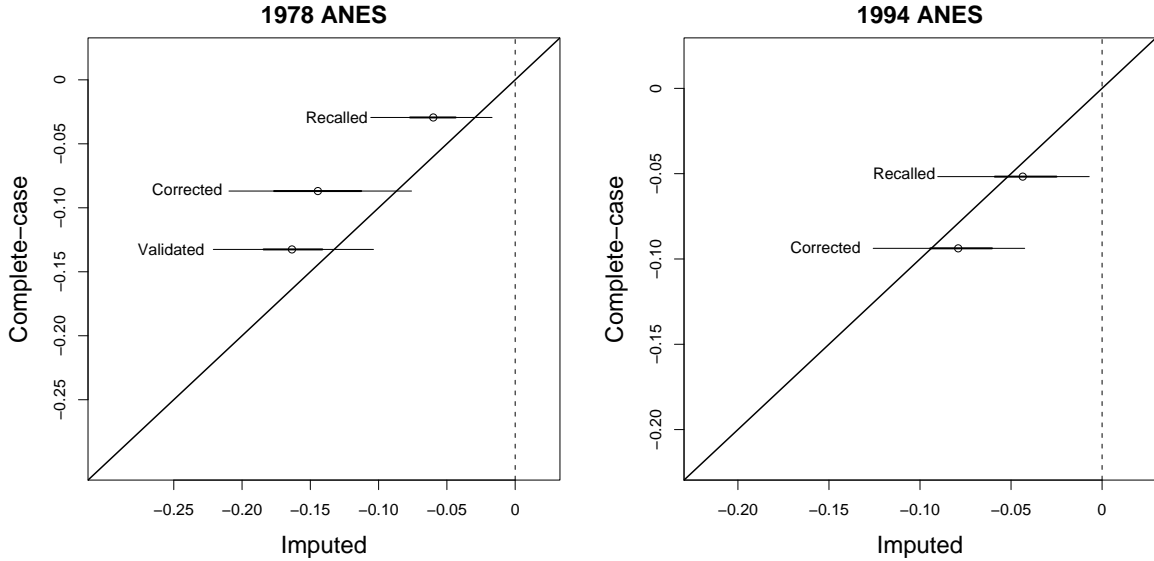


Figure 8: *Marginal effect of race with list-wise deletion versus Bayesian imputation.* The graph compares the marginal effect of race on the probability of voting in the 1978 and 1994 ANES, using list-wise deletion and Fully Bayesian imputation. The center dots correspond to the point estimates, and the horizontal bars indicate the 90% and 50% confidence intervals for the models with imputed missing values.

of race on turnout probabilities. This is likely to be an important concern in the Election Studies examined here, given that there is substantial variation in the rates of item nonresponse, with most of the variables exhibiting relatively low percentage of missing values while a few others show very high rates of nonresponse (see Appendix III). Second, imputing missing values does not change the substantive findings reported above regarding the performance of our methodology. The results for the 1978 ANES show that the estimated effects from our model correcting for misreporting are again closer to the benchmark case – using validated vote – than the effects estimated using recalled vote, and this result holds for all the ANES with validated vote. For the 1994 election, the marginal effect of race obtained from the corrected turnout model is also higher than in the uncorrected model, as was in the complete-case analysis. Once again, then, the main conclusions drawn from the model correcting for misreporting differ from those obtained using recalled vote.

4. CONCLUSIONS

The contribution of this paper are two fold. First, we have defined the conditions under which misreporting in samples will be problematic for inference. This is crucial and has largely been ignored by the literature on the misreporting that grew out of the validation studies conducted by the ANES. These studies have been more more concerned with the causes and measurement of misreporting, rather than its impact on models of

turnout.

Second, we have developed a model that corrects for this possible misreporting. The model is fully general and modular. It follows from the accounting identities of probabilities, so few new assumptions are necessary to estimate the model. It should, therefore, be of use in contexts of misreporting other than turnout estimated from ANES data. Also the modularity of the model will make extensions to the method fairly easy to implement. For example, the model directly extends to multichotomous choice. Another potentially fruitful line of extension would be to semi-parametric methods to estimate both the misreporting and turnout models. We leave these to possible extensions for further research.

The last virtue of this estimation model is that it does not require full validation studies, which are expensive, to be conducted every time a researcher is concerned about potential misreporting. As long as enough data exists to reasonably estimate the misreporting probabilities, the non-validated samples can be safely used for inference. This is clearly important in the case of the ANES. Without a method that corrects for misreporting, huge amounts of data on turnout would be lost since validation studies were only conducted a handful of years. This unvalidated data does provide important and useful information.

In addition, while the primary focus of the paper has been on estimation techniques as opposed to substantive findings, the implication for researchers interested in race seems clear. Race does have a clear negative impact on turnout and that the null previous finding have been due to problems of misreporting as had been found by [Abramson and Claggett \(1984, 1986, 1991\)](#). With the correction for misreporting developed in this paper, however, researchers could now better estimate the effect of race over the length of the ANES datasets and not just the few years with validation. In addition, researchers might wish to revisit [Wolfinger and Rosenstone \(1980\)](#) findings of the affect of registration laws to see if properly correct misreporting re-enforces or diminishes their findings.

A. VARIABLES USED IN THE TURNOUT MODEL

1. Socioeconomic indicators

Age: 1 if $Age < 30$; 2 if $30 \leq Age < 45$; 3 if $45 \leq Age < 60$; 4 if $Age \geq 60$.

Church: Frequency of church attendance. Coding: 1 if never; 2 if a few times a year; 3 if once or twice a month; 4 if every week or almost every week.

Education: Highest grade of school or year of college completed. Coding: 1 if 8 grades or less; 2 if 9–12 grades with no diploma or equivalency; 3 if 12 grades, diploma or equivalency; 4 if some college; 5 if college degree.

Female: 1 if the respondent is female, 0 if male.

Income: Household income. Coding: 1 if 0–16th percentile; 2 if 17th–33rd percentile; 3 if 34th–67th percentile; 4 if 68th–95th percentile; 5 if 96th–100th percentile.

Non-white: 0 if white, 1 otherwise.

Own Home: 1 if the respondent owns his house, 0 otherwise.

Unemployed: 1 if unemployed, 0 otherwise.

2. Additional covariates to account for misreporting

Alone: 1 if the respondent was interviewed alone, 0 otherwise.

Uncooperative: Respondent's level of cooperation in the interview, as evaluated by the interviewer. Coding: 1 if very good; 2 if good; 3 if fair; 4 if poor; 5 if very poor.

Sincerity: How sincere did the respondent seem to be in his/her answers, as evaluated by the interviewer. Coding: 1 if often seemed insincere; 2 if usually sincere; 3 if completely sincere.

In order to speed convergence, all variables were centered at their mean values, and redundant parameters (Gelman and Hill 2007) were used when fitting the hierarchical models.

B. CONDITIONAL POSTERIOR DENSITIES IN THE MODEL CORRECTING FOR MISREPORTING

From Eq 3, the conditional posterior densities required for Gibbs sampling are:

$$\begin{aligned}
 p(\beta|\gamma_1, \gamma_2, \mathbf{Y}, \mathbf{D}^1, \mathbf{D}^2, \mathbf{X}, \mathbf{Z}^1, \mathbf{Z}^2) &\propto \prod_{i=1}^N \left[\left[(1 - \Phi(\gamma_1' \mathbf{z}_i^1) - \Phi(\gamma_2' \mathbf{z}_i^2)) \Phi(\beta' \mathbf{x}_i) + \Phi(\gamma_1' \mathbf{z}_i^1) \right]^{\tilde{y}_i} \right. \\
 &\quad \times \left. \left[(1 - \Phi(\gamma_1' \mathbf{z}_i^1) - \Phi(\gamma_2' \mathbf{z}_i^2))(1 - \Phi(\beta' \mathbf{x}_i) + \Phi(\gamma_2' \mathbf{z}_i^2)) \right]^{1-\tilde{y}_i} \right] \\
 &\quad \times \exp \left[(\sigma_\beta^{-2} \times (\beta - \mu_\beta I_K))' (\beta - \mu_\beta I_K) \right]; \\
 p(\gamma_1|\beta, \gamma_2, \mathbf{Y}, \mathbf{D}^1, \mathbf{D}^2, \mathbf{X}, \mathbf{Z}^1, \mathbf{Z}^2) &\propto \prod_{i=1}^N \left[\left[(1 - \Phi(\gamma_1' \mathbf{z}_i^1) - \Phi(\gamma_2' \mathbf{z}_i^2)) \Phi(\beta' \mathbf{x}_i) + \Phi(\gamma_1' \mathbf{z}_i^1) \right]^{\tilde{y}_i} \right. \\
 &\quad \times \left. \left[(1 - \Phi(\gamma_1' \mathbf{z}_i^1) - \Phi(\gamma_2' \mathbf{z}_i^2))(1 - \Phi(\beta' \mathbf{x}_i) + \Phi(\gamma_2' \mathbf{z}_i^2)) \right]^{1-\tilde{y}_i} \right] \\
 &\quad \times \prod_{j=1}^M \left[\left[\Phi(\gamma_1' \mathbf{z}_j^1) \right]^{d_j^1} \times \left[1 - \Phi(\gamma_1' \mathbf{z}_j^1) \right]^{1-d_j^1} \right] \\
 &\quad \times \exp \left[(\sigma_{\gamma_1}^{-2} \times (\gamma_1 - \mu_{\gamma_1} I_L))' (\gamma_1 - \mu_{\gamma_1} I_L) \right]; \\
 p(\gamma_2|\beta, \gamma_1, \mathbf{Y}, \mathbf{D}^1, \mathbf{D}^2, \mathbf{X}, \mathbf{Z}^1, \mathbf{Z}^2) &\propto \prod_{i=1}^N \left[\left[(1 - \Phi(\gamma_1' \mathbf{z}_i^1) - \Phi(\gamma_2' \mathbf{z}_i^2)) \Phi(\beta' \mathbf{x}_i) + \Phi(\gamma_1' \mathbf{z}_i^1) \right]^{\tilde{y}_i} \right. \\
 &\quad \times \left. \left[(1 - \Phi(\gamma_1' \mathbf{z}_i^1) - \Phi(\gamma_2' \mathbf{z}_i^2))(1 - \Phi(\beta' \mathbf{x}_i) + \Phi(\gamma_2' \mathbf{z}_i^2)) \right]^{1-\tilde{y}_i} \right] \\
 &\quad \times \prod_{j=1}^M \left[\left[\Phi(\gamma_2' \mathbf{z}_j^2) \right]^{d_j^2} \times \left[1 - \Phi(\gamma_2' \mathbf{z}_j^2) \right]^{1-d_j^2} \right] \\
 &\quad \times \exp \left[(\sigma_{\gamma_2}^{-2} \times (\gamma_2 - \mu_{\gamma_2} I_M))' (\gamma_2 - \mu_{\gamma_2} I_M) \right].
 \end{aligned}$$

Although, as mentioned before, these conditional posterior densities have no closed forms, draws of β , γ_1 and γ_2 can be obtained using Adaptative Rejection Sampling (ARS) ([Gilks and Wild 1992](#)).

C. RATES OF NONRESPONSE FOR THE VARIABLES INCLUDED IN THE
VOTER TURNOUT MODELS

Variable	1978-1990 validated ANES	1992 ANES	1994 ANES
Age	0.00	0.00	0.00
Church	0.13	0.33	0.61
Education	0.01	0.03	0.02
Female	0.00	0.00	0.00
Income	0.13	0.10	0.10
Non-white	0.01	0.01	0.02
Own Home	0.01	0.06	0.00
Unemployed	0.00	0.00	0.00
Alone	0.01	0.01	0.02
Cooperation	0.01	0.00	0.01
Sincerity	0.00	0.00	0.01
Recalled turnout	0.01	0.00	0.00
Validated turnout	0.27	-	-
Misreport (d_j^1)	0.27	-	-
Total sample	11,632	2,255	1,795
Complete-case sample	6,452	1,210	595

REFERENCES

- Abramson, Paul R. and William Claggett. 1984. "Race-Related Differences in Self-Reported and Validated Turnout." *Journal of Politics*. 46:719–738.
- Abramson, Paul R. and William Claggett. 1986. "Race-Related Differences in Self-Reported and Validated Turnout in 1984." *Journal of Politics*. 48:412–422.
- Abramson, Paul R. and William Claggett. 1991. "Race-Related Differences in Self-Reported and Validated Turnout in the 1988 Presidential Election." *Journal of Politics*. 53:186–197.
- Albert, James H. and Siddharta Chib. 1993. "Bayesian analysis of binary and polychotomous data." *Journal of the American Statistical Association*. 88:669–679.
- Aldrich, John H. 1993. "Rational Choice and Turnout." *American Journal of Political Science*. 37(1):246–278.
- Alvarez, R. Michael. 1997. *Information and Elections*. Michigan: The University of Michigan Press.
- Anderson, Brady A. and Brian D. Silver. 1986. "Measurement and Mismeasurement of the Validity of the Self-Reported Vote." *American Journal of Political Science*. 30:771–785.
- Casella, George and Edward L. George. 1992. "Explaining the Gibbs Sampler." *The American Statistician*. 46(3):167–174.
- Congdon, Peter. 2001. *Bayesian Statistical Modelling*. London: Wiley.
- Gelfand, Alan E. and Adrian F. Smith. 1990. "Sampling-Based Approaches to Calculating Marginal Densities." *Journal of the American Statistical Association*. 85(410):398–409.
- Gelman, A., K. Gary and C. Liu. 1998. "Not Asked and Not Answered: Multiple Imputation for Multiple Surveys." *Journal of the American Statistical Association*. 93(443):846–857.
- Gelman, Andrew and Donald B. Rubin. 1992. "Inference for iterative simulation using multiple sequences." *Statistical Science*. 7:457–472.
- Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multi-level/Hierarchical Models*. New York: Cambridge University Press.
- Gilks, Walter R. and P. Wild. 1992. "Adaptive rejection sampling for Gibbs sampling." *Applied Statistics*. 41:337–348.
- Gilks, W.R., S. Richardson and D.J. Spiegelhalter. 1996. *Markov chain Monte Carlo in practice*. London: Chapman and Hall.

- Härdle, Wolfgang. 1990. *Applied nonparametric regression*. New York: Cambridge University Press.
- Hill, Kim Q. and Patricia A. Hurley. 1984. "Nonvoters in Voters' Clothing: The Impact of Voting Behavior Misreporting on Voting Behavior Research." *Social Science Quarterly*. 65:199–206.
- Horton, Nicholas J. and Ken P. Kleinman. 2007. "Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models." *The American Statistician*. 61(1):79–90.
- Ibrahim, J.G., M. Chen, S.R. Lipsitz and A.H. Herring. 2005. "Missing-Data Methods for Generalized Linear Models: A Comparative Review." *Journal of the American Statistical Association*. 100(469):332–346.
- Ibrahim, Joseph G., Min-Hui Chen and Stuart R. Lipsitz. 2002. "Bayesian Methods for Generalized Linear Models With Missing Covariates." *The Canadian Journal of Statistics*. 30:55–78.
- Ibrahim, Joseph G., Stuart R. Lipsitz and Min-Hui Chen. 1999. "Missing Covariates in Generalized Linear Models When the Missing Data Mechanism Is Nonignorable." *Journal of the Royal Statistical Society B*. 61:173–190.
- Katosh, John P. and Michael W. Traugott. 1981. "The Consequences of Validated and Self-Reported Voting Measures." *Public Opinion Quarterly*. 45:519–535.
- Lipsitz, S.R. and J.G. Ibrahim. 1996. "A Conditional Model for Incomplete Covariates in Parametric Regression Models." *Biometrika*. 83:916–92.
- Little, R.J. and Y. Wang. 1996. "Pattern-Mixture Models for Multivariate Incomplete Data With Covariates." *Biometrics*. 52:98–111.
- Little, Roderick J. A. and Donald B. Rubin. 1987. *Statistical analysis with missing data*. New York: Wiley.
- Murphy, Kevin M. and Richard H. Topel. 1985. "Estimation and Inference in Two Step Econometric Models." *Journal of Business and Economic Statistics*. 3:370–379.
- Nagler, Jonathan. 1994. "Scobit: An Alternative Estimator to Logit and Probit." *American Journal of Political Science*. 38(1):230–255.
- Rubin, Donald B. 1976. "Inference and missing data." *Biometrika*. 63:581–592.
- Schafer, Joseph L. 1997. *Analysis of Incomplete Multivariate Data*. London: Chapman and Hall.
- Schafer, Joseph L. and John W. Graham. 2002. "Missing Data: Our View of the State of the Art." *Psychological Methods*. 7(2):147–177.

- 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.
- Weir, Blair T. 1975. "The Distortion of Voter Recall." *American Journal of Political Science*. 19:53–62.
- Wolfinger, Raymond E. and Steven J. Rosenstone. 1980. *Who Votes?* New Haven: Yale University Press.