Political Analysis, 8:2

Alternative Corrections for Sample Truncation: Applications to the 1988, 1990, and 1992 Senate Election Studies

John Brehm
Duke University

High levels of nonresponse plague all three waves of the National Election Studies' Senate Studies. Each of the studies failed to elicit interviews from close to one of every two selected sample persons, a rate far worse than the NES regular Pre- and Post-Election Studies. This paper addresses three interdependent problems: Given limited data about the nonrespondents, how can we model the causes of nonresponse? Using these models, how can we adjust our analysis of the Senate Study data to compensate for nonresponse? What difference does nonresponse make for our understanding of the dynamics of Senate elections?

1 Introduction

ACADEMIC PUBLIC OPINION surveys will never interview every single person selected for our samples. Practitioners even adjust their budgets and alter the allocation of cases per sample segment because they expect a nonnegligible nonresponse rate. While a nonresponse rate of 30% (plus or minus a few percent) might be acceptable for some academic studies (e.g., the General Social Survey or the National Election Studies), there are times when response rates fall well below expectations. The 1988, 1990, and 1992 Senate Election Studies are perhaps a case in point: the three waves failed to interview about one out of every two people selected to the sample. (The response rate was 43% for the 1988 study, 46% for the 1990 study, and 56% for the 1992 study.)

In principle, we have a well-developed set of techniques to correct for selection bias in general (Heckman 1976, 1979; Achen 1986) and survey nonresponse in particular (Little 1982; Rubin 1987; Brehm 1993). In order to apply these corrections, we require information about the nonrespondents. Of course, for any study of nonrespondents, the information that would best identify differences between respondents and nonrespondents is the very questionnaire we could not administer. In many circumstances, all we will know about the nonrespondents is contextual information and the response rate of the sample. This is exactly the situation faced by analysts of the 1988–1992 Senate Election Study. The only

Author's note: Thanks go to John Aldrich, Bill Bianco, Charles Franklin, Paul Gronke, Gary King, and Doug Rivers for comments on early drafts of the manuscript.

Copyright 1999 by the Society for Political Methodology

information we have, in the published releases of the Senate Study, is the nonresponse rate for each state in the sample and state-level contextual demographic information.[1] Because the only information we have about the nonrespondents is state–level contextual data, all our analyses of the Senate data are based upon a "truncated" sample. While there is a clear econometric basis for handling "censored" samples (where we have some information about the nonrespondents), the technology for "truncated" samples is less settled.

This paper suggests ways in which we might address the high nonresponse rates by "converting" the truncated sample to a censored sample, as well as ways in which we might draw upon an assumption of normality in the truncated sample. The first method draws upon parallels between the statewide response rate and the inverse Mills' ratio estimated in the usual Heckman/Achen two–stage selection model. The second method creates dummy nonrespondents and employs statewide contextual information as regressors in the first stage of the Heckman/Achen selection model. The third method employs a "stochastic truncation" approach, which builds upon an assumption that respondents are observed only if the systematic component for selection exceeds a random truncation threshhold. Each of the three methods entails strong assumptions, but all of them quite different. The strength of the corrections is obtained not by opting for one of the approaches but in employing all three, triangulating between methods which make different assumptions about the data. This paper will report on the relative strengths of the methods, namely, that the stochastic truncation approach generates more precise estimates of standard errors than the other two but that the other two approaches are more assured of convergence.

This paper wrestles solely with what is referred to in the literature on missing data as "unit nonresponse," the condition where entire cases are missing from the sample. This is opposed to the condition referred to as "item nonresponse," where some of the observations for particular variables are missing. The unit nonresponse condition arises when a selected sample person refuses to answer a questionnaire; the item nonresponse condition arises when a respondent refuses to answer a particular question but is willing to answer others. There is now a very strong method for the item nonresponse approach; see King et al. (1998) for a recent political science application. The purpose of this paper is an exposition of methods for a particular form of the unit nonresponse condition. Obviously, there is also item nonresponse in the applications to be presented below (indeed, in virtually every conceivable model to be run on data such as the Senate Election Studies).

Based on the state response rate and guided by general research on nonresponse, how far can we go in assessing the damage of the high nonresponse rates? This paper proceeds in three steps. First, I examine potential causes of variation in aggregate response rate by state across the two waves of the Senate study. Second, I explicate three corrections for nonresponse in regression analysis. The first two methods extend from econometric corrections for selection bias developed by Heckman (1976, 1979) and Achen (1986). The third method uses maximum-likelihood models of sample truncation similar to those touched upon by King (1989) or Maddala (1983). This section also discusses the comparative strengths and weaknesses of the three methods. Finally, I demonstrate the effect of the corrections on a model of voting for Senator based on issue positions.

[1] This level of information is less than in the regular Pre- and Post-Election Studies. For every major release since 1978, the NES includes limited information about the amenability of the respondents such as the interviewer's assessment of cooperativeness, whether a persuasion letter was sent to the respondent or a refusal conversion attempted, and (recently) whether the respondent refused initially. For the 1986–1990 studies, the NES also gathered limited demographic data about the nonrespondents (from the interviewer's observation or from contacts with other household residents or neighbors).

2 Identifying Causes of Nonresponse

The corrections to follow proceed from the assumption that aggregate data informs us about response behavior of individuals. In this second section, I examine the extent to which aggregate information indentifies causes of nonresponse to the Senate Election Studies. The models developed in this section are used in the corrections developed in Section 3, and applied in Section 4.

Many of us appreciate how many hurdles a sample person must cross before he or she becomes a respondent to our studies. Not only must the individual be selected to the sample, not only must an interviewer be able to contact this individual, not only must the interviewer determine that the contacted person is capable of being interviewed and eligible under the sampling frame, but also, finally, the respondent must comply with the interviewer's request for an interview.[2] Most nonresponse results from refusals, and this is especially so for telephone surveys. [Typically, telephone surveys achieve better rates of *contacting* respondents but worse *refusal* rates than face-to-face surveys (Groves 1988).]

The theoretical problem of why people choose or refuse to participate in surveys is thus quite complex. In a separate work (Brehm 1993), I develop a model of survey participation, drawing upon multiple techniques to gauge the attitudes and characteristics of the nonrespondents. It is beyond the scope and intent of this paper to explicate such a model.

Nonetheless, we might speculate about the possible characteristics of the nonrespondents. One view of nonrespondents stems from parallels between participation in surveys about politics and political participation generally. Plausibly, nonrespondents are less likely to participate in a range of political activities, including voting.[3] But one should remember that the nonrespondents refused to participate in the Senate Election Studies well before they knew what the survey would be about. Would we expect nonrespondents to a medical survey to be nonvoters as well? Perhaps, but the instrumental links between participation in surveys and politics become less clear.

An alternative characterization of the nonrespondents views them as socially marginalized individuals, alienated from politics and social life. These individuals may be less informed about politics, hold views on policy contrary to both major parties, or otherwise be distanced from mainstream politics.[4] As with the first characterization, it is awkward to see how this population extends to nonpolitical surveys. The purpose of this essay is *not* to develop a general model of survey participation, but to examine how nonresponse affects analysis of the Senate Election Studies. With this purpose in mind, we need only to examine participation in the present survey.

We lack virtually all information about the nonrespondents except for state-level demographic contextual information.[5] Can we identify potential causes of nonresponse from the characteristics of the states? Such a contextual approach is somewhat plausible. We have reason to suspect that conditions of the respondent's environment (urbanness, level of crime, density of population) affect response rates (Groves and Couper 1998; House and

[2] The NES calculation for response rate is the ratio of the number of interviews divided by the number of interviews plus refusals plus noncontacted sample persons.

[3] There is strong evidence to support this claim. Goyder (1986) reports from experimental evidence, interviewing populations of known voters and nonvoters, that the nonrespondents vote less frequently than voters. Brehm (1993) demonstrates that models of turnout overestimate turnout rates without correction for nonresponse.

[4] There is support for this view as well. Brehm (1993) demonstrates that nonrespondents are less informed about politics to a substantial degree than respondents.

[5] We could conceivably improve upon the level of aggregation by working with contextual information organized by area code and exchange.

Wolf 1978; Steeh 1981). We also know that income, age, race, working status, income, and gender covary with response rates (Groves and Couper 1998; Goyder 1986; Benus 1971; DeMaio 1980; Smith 1983; Weaver 1973; Hawkins 1975; O'Neil 1979; Cobb, King and Chen 1957; Dohrenwend and Dohrenwend 1968). How far does an aggregate approach to response rates take us? As the reader will see in the subsequent section, we might use the aggregate predictors for response as a means to convert the truncated sample to a censored sample, and apply two-stage corrections for selection bias in such cases.

Any examination of the correlates of the aggregate characteristics of the state and state response rates is inherently inefficient and subject to the usual problems of ecological inference. As an initial demonstration of both the plausibility of examining aggregate attributes of states and the inherent inefficiency of such an approach, Fig. 1 plots the response rates for the three studies against one another by state. The diagonal line marks where response rates would be equal from year to year. The first observation should be that the there is a loose linear relationship between response rates in each of the three years. This finding should reinforce the plausibility of an effect of aggregate characteristics of the states on response. Of course, our second observation should be that this fit is quite loose, and we must temper

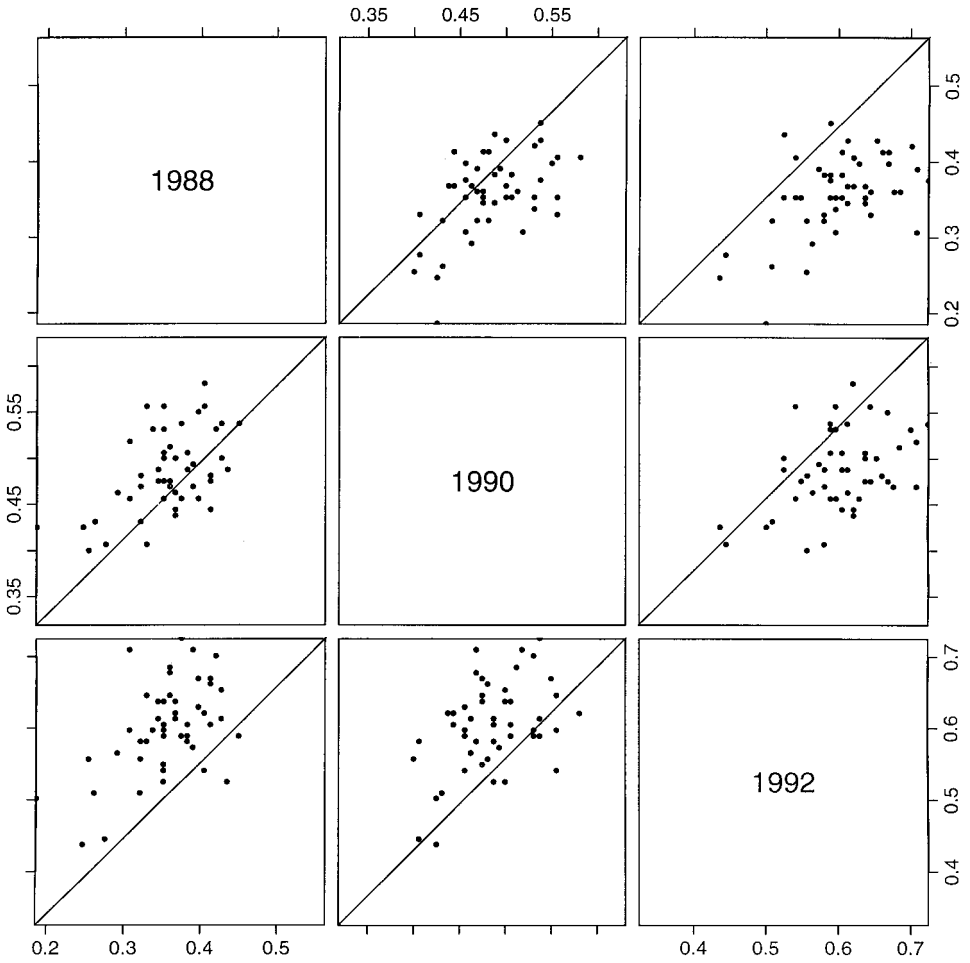

Fig. 1 Scatterplot matrix of response rate by state, 1988, 1990, and 1992 Senate Election Studies.

our expectations about the predictive power of aggregate characteristics for state response rates. (The R^2 for a regression of the 1990 response rate on the 1988 response rate is only 0.28, and the R^2 for a regression of the 1992 response rate on the 1990 response rate is even weaker, at 0.18.) Third, we should note that the response rate for most states rose from 1988 to 1990 and, again, from 1990 to 1992. The dashed line denotes where response rates for 1990 would equal those for 1988, or the boundary between rising and falling response rates. All of the 10 states with the lowest response rates in 1988 improved for 1990, and all but 1 of the 10 states with the lowest response rates in 1990 improved for 1992. Of course, part of this effect is simply a regression to the mean, but we might also attribute greater success in these states to greater efforts by study staff.[6]

We can model the effect of aggregate characteristics on state response rates through what is sometimes referred to as a "dose–response" probit [or what King (1989) refers to as "grouped uncorrelated binary variables"]. Dose–Response probit stems from pharmacological studies where a fixed number of subjects receive the same dose of a trial medication, and where we observe a binary response among some subset of the subjects. Here, we observe a number of responses (interviews) out of a variable number of subjects (interviews plus nonrespondents) in each state. The coefficients of the dose–response probit are exactly equivalent to coefficients from the more familiar probit estimated for individuals.[7]

Table 1 displays the estimates for the dose–probit analysis of response by state for selected demographics for the 1988, 1990, and 1992 Senate Election Studies. (All of the demographics are taken directly from the 1988 Current Population Surveys.) The columns immediately to the right of the dose–response probit estimates display estimates of the marginal effects.[8]

Virtually all of the predicted changes in probability of compliance cohere with the substantial literature on correlates of nonresponse. (Where exceptions from the established literature arise, the dose–response probit coefficients aren't statistically distinguishable from zero due to wide standard errors, even though the first difference may be large.) The more urban the state, the lower the response rate (at least for 1990 and 1992): a state which is 10% metropolitan (e.g., New Jersey) would have a response rate 8% lower than the mean for 1990 and 15% lower than the mean for 1992. A wide range of studies confirms lower response rates in urban areas compared to rural areas, although there is little consensus as to the cause (e.g., House and Wolf 1978; Steeh 1981).

[6] Partial rejection of statistical regression to the mean appears at the opposite end of the graphs: among the states with the highest response rate for 1988 or 1990, about half the states worsened, while another half improved. If the effect were solely regression to the mean, we should expect to see a larger proportion of the states with high response rates in the first year to have lower response rates in the second year.

[7] The likelihood function for the dose–response probit is

$$\mathcal{L} = \prod_{i=1}^{50} \pi_i^{y_i} (1 - \pi_i)^{N_i - y_i}$$

where y_i is the number of respondents in the state, N_i is the number of sample persons (respondents + nonrespondents), and π refer to the probability of response, estimated via a probit link function for the parameters of interest. When the number of subjects (N_i) equals one, this model is exactly the same model as the more common probit.

[8] These marginal effects are computed by taking the difference in the predicted probability of compliance when substituting in the maximum value for an independent variable found among the 50 states holding all other variables at the mean, from the predicted probability of compliance holding all variables at the mean. That is, the first differences for unemployment do not reflect 100% unemployment, but the maximum level of unemployment in each year (10.0%).

Table 1 Dose–response probit estimates of state response rate as a function of demographic contextual information, 1988–1990 Senate Study[a]

Variable	*Mean*	*Maximum*	*1988*	*First difference*	*1990*	*First difference*	*1992*	*First difference*
Constant			1.67		1.81		2.15	
			(0.49)		(0.49)		(0.61)	
% in metropolitan area	6	10	0.36	0.07	−2.40	−0.08	−0.86	−0.15
			(1.11)		(1.19)		(1.36)	
% white-collar workers	51	61	−1.25	−0.22	−1.74	−0.34	−1.56	−0.40
			(0.66)		(0.72)		(0.88)	
% blue-collar workers	31	41	−0.81	−0.07	−1.58	−0.23	−0.97	−0.23
			(0.44)		(0.54)		(0.62)	
% unemployed 1989	6	10	0.17	0.07	3.50	0.15	−0.81	−0.10
			(0.91)		(1.36)		(1.24)	
% rental	32	51	−2.27	−0.32	−0.97	−0.17	−1.81	−0.39
			(0.41)		(0.39)		(0.45)	
% elderly	11	17	−2.18	−0.09	−2.06	−0.12	−1.23	−0.15
			(0.91)		(0.85)		(0.99)	

[a]The dependent variable is the state response rate, where the unit of analysis is a state. The independent variables refer to the percentage of persons in the state meeting the given criteria (e.g., elderly, unemployed, etc.). Standard errors are in parentheses. "First differences" represent the approximate net change in probability of response if the given independent variable were at its maximum, holding all the other variables at their means.

As the fraction of the workforce that is employed increases, the probability of response decreases. In a state with a 61% white-collar workforce, we should expect a drop in the response rate of between 22% and 40%. In states with a 41% blue-collar workforce, we could expect a drop in response rate from 7% (in 1988) and 23% (in 1990 and 1992). Conversely, in states with high unemployment (10% in 1989), we could see an increase in response rates of about 7% for 1988 and 15% for 1990, but a drop in response rates of about 10% in 1992. As I have stated elsewhere (Brehm 1993; see also Groves and Couper 1998), the most prominent reason that wealthier people express for refusing to participate in surveys is "I'm too busy." If "I'm too busy" is more than a polite dismissal, then it is quite plausible that employed people are busier than unemployed people and thus more likely to refuse. These findings run counter to our intuition about socioeconomic bases of high refusal rates but are, I believe, better grounded in an understanding of the process of eliciting response.

The more frequently that the state's population resides in rental housing, the lower our response rates. In 1988, we should expect a response rate nearly 32% below the mean in states with the maximum (51%) of its population living in rental housing. In 1990, that estimate of a first difference drops to about 17%, but it returns to 39% for 1992. This finding is consistent with the lower response rates in metropolitan areas (which tend to have a greater proportion of rental housing than urban areas). It is also consistent with the greater difficulty interviewers have contacting respondents who live in apartment buildings or who may have greater mobility (Brehm 1993, Groves and Couper 1998).

Finally, as the fraction of the state's population that is elderly rises, our response rate falls. In states with 17% of its population over the age of 65, we could expect response rates to be 9–15% lower than the average. We can attribute some of the lesser response among the elderly to a greater chance that the selected respondent is not healthy enough to participate (Brehm 1993). Other writers (Goyder 1986; House and Wolf 1978) attribute lower response

rates among the elderly to a greater fear of crime (which seems frankly implausible as a reason for nonresponse to a telephone survey).

Obviously, all that we gain from the preceding paragraphs is a description of the characteristics of high and low response states. Description provides at least limited insight into why response rates vary so much by differing states, and perhaps into the causes of nonresponse itself. These descriptive approaches will prove useful in Section 4, where I attempt to compensate for the effects of nonresponse in models of politics. As the next section details, if we are to adjust our research to compensate for nonresponse, one of our best approaches is to incorporate the reasons for nonresponse into our substantive models.

3 Methods for Compensating for Nonresponse

Corrections for nonresponse to regression models follow from a now well-developed statistical literature on selection bias (Tobin 1956; Heckman 1976, 1979; Achen 1986; Little 1982; Rubin 1987; Fay 1986). Although the methods of all differ, all require models to predict whether we observe a case (i.e., models predicting whether a sample person responds or not).

A little algebra demonstrates both the problem and the potential cure. Suppose that our equation of interest (the *outcome model*) is a simple linear model:

$$Y_i = \beta' X_i + \epsilon_i \tag{1}$$

(That is, Y is a linear function of the exogenous X's and some error ϵ.) Suppose that Y is observed if and only if a second linear model (the *selection model*) exceeds zero:

$$c_i = \delta' Z_i + u_i > 0 \tag{2}$$

Suppose further that the errors (ϵ and u) are jointly distributed normally:

$$\begin{pmatrix} \epsilon_i \\ u_i \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{pmatrix}\right] \tag{3}$$

That is, ϵ and u have a covariance $\rho\sigma$, where σ is the standard deviation of ϵ. This covariance ($\rho\sigma$) represents the extent to which other causes of the outcome, not explicitly appearing on the right hand side of the outcome equation, are related to the other causes of selection into the sample.

The selection mechanism [Eq. 2] causes inconsistency in estimates of the outcome model (1):

$$\begin{aligned} \mathrm{E}(Y_i \mid Y_i \text{ observed}) &= \mathrm{E}(\beta' X_i + \epsilon \mid \delta' Z_i + u_i > 0) \\ &= \beta' X_i + \rho\sigma \frac{\phi(\delta' Z_i)}{\Phi(\delta' Z_i)} \end{aligned} \tag{4}$$

[where $\phi(\,)$ is the normal probability density function and $\Phi(\,)$ is the cumulative normal probability density function]. That is, the expected value of Y_i is the true value plus a function [the "inverse Mill's ratio," $\phi(\,)/\Phi(\,)$] of the probability that the case is selected. Only if the error terms are unrelated or all cases are selected with certainty does the estimate of Y_i equal the true value under nonresponse.

Many of the default responses of our commonly used statistical packages (e.g., listwise deletion, weighting, substitution of the sample mean for missing cases) are of no help to us in these circumstances. Listwise deletion simply replicates the results of selection bias in the outcome model [see King et al. (1998) for a spirited demonstration of the consequences of listwise deletion]. Weighting for nonresponse does not remove, or even address, the underlying problem of the correlated error structure (Brehm 1993). Substitution of the mean value for missing cases reduces the covariance between all regressors, resulting in an attentuation of any regression estimates (Kalton 1983).

One successful approach proceeds directly from the foregoing specification of the nature of selection bias.[9] If we can specify the selection equation [Eq. 2], we could incorporate the inverse Mill's ratio in our equations of interest and obtain consistent estimates (Heckman 1976, 1979; Achen 1986). Typically, we would estimate the selection equation on the basis of data collected about the respondents and nonrespondents alike. In other works (Brehm 1989, 1993), I suggest using any of three sets of data which may be gathered about the nonrespondents: the nonrespondents' reasons for refusal, demographic information that may be gathered by observation or by contact with neighbors, or study administration measures (such as whether persuasion letters were sent or the number of calls to the household). Obviously, telephone surveys like the Senate Studies pose additional difficulties collecting such information over the traditional face-to-face Pre–Post studies. These difficulties are hardly insurmountable; as with the face-to-face studies, the survey research centers routinely collect information about the administration of telephone studies which might be used in the study administration selection model. Likewise, the administrators might record reasons for refusal and hazard crude estimates of the age, race, and gender of refusals.[10]

No such individual–level information about nonrespondents is available for any of the three waves of the SES. While the Heckman–Achen methods are clear in cases of *censored samples*, where we have some individual–level information about the nonrespondents, the situation of the Senate study is more akin to a *truncated sample*, where we lack all information about omitted cases. We might adopt three approaches: an ad hoc approximation to the inverse Mill's ratio correction; substitute dummy cases for the nonrespondents, limiting our analysis of selection to aggregate contextual data; or maximum-likelihood estimates of the truncated sample. I take each approach in turn.

3.1 *Aggregate Approximates to Individual Response*

While we cannot compute the inverse Mill's ratio (IMR) for individual cases, we can compute an equivalent to the IMR at the level of states. Consider the components of the denominator of the IMR: the $Z\delta$ represent the predicted probability of compliance, expressed as units of the standard normal distribution. Computing the cumulative normal density of the $Z\delta$ yields a predicted probability of compliance, $\Phi(Z\delta)$. Aggregated over the subsample for a state, we have the $\Phi^*(Z\delta)$: the response rate for the state, which may serve as (an admittedly inefficient) proxy for the denominator. By taking the inverse of the cumulative normal for the response rate for the state, we obtain an estimate for the $Z\delta$. An

[9] There are a wide number of similar, model-based approaches; see Little and Rubin (1987) for a several multiple imputation approaches.

[10] Additionally, two of the corrections below (the "pseudo-IMR" and "dummy nonrespondent" approaches) would benefit from response rates for substate aggregations that could be linked to census data, such as the area code or three-digit exchanges.

approximation to the IMR may be expressed as

$$\text{PseudoIMR} = \frac{\phi(\Phi^{-1}(\text{State Response Rate}))}{\text{State Response Rate}} \tag{5}$$

Substituting a transformation of the statewide response rate for individual level responsiveness should make intuitive sense. The reader should note that the pseudo-IMR assumes that we can accurately describe individual–level responsiveness by aggregate responsiveness. As the size of the aggregate goes to one, the expected value of the pseudo-IMR goes to the individual IMR. With alternative data, it is possible to compare this pseudo-IMR and one derived directly from selection models. With the 1988 Pre–Post, where we have a relative wealth of data about the nonrespondents, I estimate three selection models. The correlations between the IMR generated from the selection models and a pseudo-IMR generated from the state response rate range from 0.45 to 0.67. While these correlations hardly approach unity, the signs and magnitude are within reason. As a second preliminary demonstration of the usability of the pseudo-IMR, turn to Section 3, where I examine what differences arise between the three corrections and the uncorrected estimates.

The pseudo-IMR approach could prove counterproductive, actually introducing new inconsistency, in certain circumstances. Because the IMR is a nonlinear function, the expected value of the pseudo-IMR will not be the expected value of the true IMR [since the expected value of a nonlinear function is not the same as the nonlinear function of the expected value; see Kennedy (1990, p. 5)]. However, the IMR in the area for medium to high probability of response is close to linear [see Achen (1986, p. 104) for a graphical representation], and the difference between the pseudo-IMR and the true IMR will be small for such cases. As long as one has reasonable grounds to assert that the respondents to the study had a reasonably high probability of complying with the request for an interview, the pseudo-IMR should come close to the true IMR. If one believes that the respondents had a very low probability of response, then this method could introduce inconsistency.

It should be transparent that the pseudo-IMR approach is not consistent in the sense that it gets no closer to the true value as the sample size increases. If one continues to use the state response rate as the basis for the predicted probability of response for individuals, then this estimator does not improve as our sample size approaches infinity.

We know the basis for the derivation of the pseudo-IMR in the consistent, but unavailable, individual–level response rate. The pseudo-IMR is reasonably close to, although not identical to, the individual IMR under alternative data sets. And as the reader will see shortly, the results from the pseudo-IMR are very close to the results generated by the two wholly different approaches to follow.

The pseudo-IMR, like the other two methods to follow, entails assumptions about the data. First, the pseudo-IMR approach assumes that the aggregate response rate provides an unbiased estimate of the individual–level proclivity to respond. Relatedly, this approach does not take advantage of any information about the aggregate that might induce higher or lower response rates. Instead this approach simply relies upon a conversion of the aggregate response rate to the approximation of the IMR.

Second, the pseudo-IMR approach explicitly depends upon a linear model for the outcome equation, when much of the work in political science which will use survey data is based on limited dependent variables. A linear probability model could be used (indeed, will be in the application below), correcting the standard errors as appropriate, with all the unwanted assumptions about the uniform effect of the variables on the dichotomous outcome throughout the range of probability.

If there is a clear advantage of this correction over the two to follow, it is that the pseudo-IMR is very simple to construct. All one needs to know is the response rate for the state, which is then easily entered as a new variable in the dataset. The transformation itself relies upon the normal density and inverse cumulative normal density functions which are standard in many statistical packages [or may be approximated (see Abramowitz and Stegun 1964)]. The next correction is only marginally more complex but adheres more closely to the demands of the Heckman–Achen two-stage corrections.

3.2 *Dummy Nonrespondents and Aggregate Predictors*

A second approach exploits the aggregate, state–level demographic information discussed in Section 2 and our knowledge about the number of cases. Obviously, the state-level demographic characteristics apply equally to respondents as to nonrespondents (or to any other resident of the state). Further, we know how many nonrespondents there are for each state (the number of sample persons allocated for the state less the number of respondents). It is a relatively simple matter to create "dummy" nonrespondents and record the state-level demographic information to be used as the Z in the selection model [Eq. 2]. For the present circumstances, I use the regressors in the dose–response probit (Table 1) as the Z.

This procedure permits one to directly employ the Heckman/Achen two-stage corrections. Note, however, that the Heckman/Achen two-stage corrections, like any system of equations, must be identified. The only technical requirement for identification is that the first stage (selection equation) include at least one variable not present in the second stage. Since the corrections in this paper use the same battery of regressors found in the dose–response probit (Table 1), and since *none* of the regressors in this table appear in any of the substantive models to follow, the selection model is identified.[11]

A singular advantage of this approach is that a developed literature proves the consistency of the Heckman–Achen two-stage corrections. The disadvantage of such an approach is that we are limited to the use of aggregate predictors of compliance. As we know from Section 2, the state-level demographic information provides very weak clues as to the process of compliance. This disadvantage is significant: if we use poor predictors of compliance in the first stage of the correction (estimation of the selection model), we increase the risk of introducing substantial error to the substantive model.

Like the pseudo-IMR approach, the dummy nonrespondent approach requires an assumption about aggregates. Here, the dummy nonrespondent approach works only if the characteristics of the aggregate (e.g., the state) influence compliance by the individual respondent. There is a long literature in survey research which contends that such characteristics of the aggregate as the urbanness of the community does influence response (e.g., House and Wolf 1978; Steeh 1981), but as mentioned above, there is little consensus as to *why* urban areas discourage response. The difference between the pseudo-IMR and the dummy nonrespondent approaches with regard to the assumption about aggregates is that the former approach assumes that aggregate response rates are an unbiased measure of individual responsiveness, while the dummy nonrespondent approach assumes that characteristics of the aggregate are a good predictor of individual responsiveness.

[11] Further, the models are theoretically identified as well. Examine the list of regressors, and ask whether any of these regressors should properly be included in the substantive models. Some might, but are omitted from the original models. But under what grounds would "percentage of rental housing" be related to one's vote choice?

3.3 *Stochastic Truncation*

A third, wholly different approach to adjusting our models for nonresponse derives from maximum-likelihood models of "stochastic truncation" (Bloom and Killingsworth 1985; King 1989; Maddala 1983). Let y be the dependent variable of interest, observed only when some other variable c (the "truncation variable") is greater than 0. Further, let σ_y^2 be the variance of y, σ_c^2 be the variance of c, and $\rho\sigma$ be the covariance of y and c [as above; Eq. 3]. Then the likelihood function for the stochastic truncation model may be written

$$\mathcal{L} = \prod_{i=1}^{n} \phi\left(y_i \mid \mu_y, \sigma_y^2\right)\left[\left(1-\Phi\left(0 \mid \mu_c + \frac{\sigma_{yc}}{\sigma_y^2}(y_i - \mu_{y_i}), \sigma_c^2 - \frac{\rho\sigma}{\sigma_y^2}\right)\right)\Big/\left(1-\Phi\left(0 \mid \mu_c, \sigma_c^2\right)\right)\right] \tag{6}$$

where the equivalent to the outcome equation [Eq. 1] appears as the parameter μ_y,

$$\mu_y = X\beta$$

and the equivalent to the selection model [Eq. 2] is the parameter μ_c

$$\mu_c = Z\delta$$

(Note that in the covariance structure laid out above [Eq. 3], the variance σ_c^2 is fixed at 1, much like the standard probit model. Here, we have to fix σ_c^2 to 1 in order to identify the model.)

In principle, the stochastic truncation approach is the best of the three. Aside from the usual advantages of maximum-likelihood approaches (i.e., invariance to reparameterization, invariance to sample plans, consistency), the stochastic truncation approach should be more efficient as a full information correction than the limited information two-stage corrections (Nathan 1981). Also, unlike the previous two approaches, the set of regressors used in the selection component of the model need not be limited to regressors available for all sample persons (respondents and nonrespondents alike) and can use regressors from the survey itself. (In the application below, I restrict the regressors to only those available for all three approaches for purposes of comparability).

In practice, the stochastic truncation approach has a serious constraint upon its use. All three of the methods require exclusions in the selection and outcome models for identification, but the stochastic truncation method is the only one of the three which estimates both equations simultaneously. Hence, identification becomes an immediate issue for purposes of obtaining estimates. One could regard this as an advantage in that the stochastic truncation approach flags unidentified selection models by failing to converge. In simulation tests based on the three approaches, the stochastic truncation approach would fail to converge in about 10% of the replications.

We now have three distinct approaches to correct for sample truncation in the 1988–1992 Senate Election Studies. All three approaches have advantages and disadvantages. The ad hoc approximation to the IMR is the most straightforward to compute and requires no additional information beyond subsample response rates (in this case, by state), yet it is also the most jerry-rigged of the lot, inherently inefficient, and of unproven consistency. The dummy nonrespondent approach employs proven statistical procedures, requires only a moderate amount of data processing to generate the dummy nonrespondents, and may use

sensible aggregate predictors of individual response, yet also forces a potentially undesirable ecological inference. The stochastic truncation correction proceeds most directly from the assumptions behind the joint error structure of the selection and outcome models, yet is also most dependent upon those assumptions for the estimation, computable in only a handful of statistical packages and (in practice) most sensitive to model specification. I now proceed to an application of these techniques to three widely respected models based on the 1988 Senate Election Study.

4 Applications of Nonresponse Corrections

In this section, I compare the uncorrected estimates with the three corrections for one well-respected model of Senate elections, a model of vote for Senator as a function of issue positions (Abramowitz and Segal 1990). The dependent variable is a dichotomous one, and hence, analysts would conventionally and appropriately use a probit model. In the demonstrations to follow, I use a linear probability model: for one, the methods developed above require linear outcome models; further, we can directly compare coefficients to assess the effect of corrections across the three approaches (since the assessment of effects in the probit case depends upon the full set of coefficients, a comparison across four sets of estimates is much more awkward).

Abramowitz and Segal (1990) model voting for Senator as a function of issue positions. Their broader purpose is to examine the role that different issues had upon voting for President, Senator, and Representative. Here, I consider the vote for Senator. The 1988 presidential election appeared to many observers as one devoid of issues. But Abramowitz and Segal demonstrate that the respondent's perception of the condition of the national economy, attitudes toward spending, liberal–conservative ideology and (potentially) abortion affected the respondent's vote for President. Some of these issues (namely, perceived condition of the national economy and attitudes toward spending) also affect the vote for Senator.

I replicate Abramowitz and Segal's model of voting for Senator in Tables 2–4 using their original coding of the variables. (The sampling details, question wording, and coding of variables are available on the *Political Analysis* web page.) As Abramowitz and Segal noted, partisanship, presidential vote (in 1988 and 1992) and incumbency are the strongest predictors of the vote for Senator. In 1988, the respondent's attitude toward government spending and assessment of the condition of the national economy were also substantively significant. For 1990, the effect of the respondent's attitude toward government spending and assessment of the national economy became indistinguishable from zero. However, in 1990, personal finances and approval of George Bush become statistically significant predictors, although substantively still slight. In 1992, attitudes toward government spending become enormously important, trumping even the effect of senate incumbency (but still behind the presidential vote). Attitudes toward abortion also becomes statistically and substantively significant.

For the 1988 incarnation of Abramowitz and Segal's model, the most important and consistent changes for the three corrections affect the very variables that Abramowitz and Segal point to as illustrations of the importance of issues. The effect of the respondent's assessment of the national economy and attitudes toward government spending are overestimates before correction for nonresponse. Substantively, does it make sense that issues matter more for respondents than for nonrespondents? Perhaps, as nonrespondents are less likely to attend to politics in the first place.

Table 2 Uncorrected and corrected estimates, model of voting for Senate as a function of issue positions, 1988 Senate Study[a]

		Corrected		
Variable	*Uncorrected (GLS)*	*Pseudo-IMR*	*Dummy NonRs*	*Truncation*
Constant	0.522	0.489	0.489	0.491
	(0.051)	(0.037)	(0.037)	(0.044)
Party ID	0.106	0.090	0.090	0.095
	(0.028)	(0.018)	(0.018)	(0.023)
Presidential approval	0.019	0.019	0.042	0.038
	(0.017)	(0.012)	(0.011)	(0.016)
Personal finances	−0.014	−0.017	−0.017	−0.043
	(0.015)	(0.010)	(0.010)	(0.018)
National economy	0.039	0.032	0.032	0.046
	(0.012)	(0.013)	(0.014)	(0.020)
Abortion	−0.030	−0.010	−0.011	−0.009
	(0.029)	(0.018)	(0.018)	(0.013)
Help Blacks	0.014	0.009	0.009	0.010
	(0.011)	(0.007)	(0.009)	(0.013)
Government spending	−0.018	−0.012	−0.012	−0.012
	(0.005)	(0.003)	(0.003)	(0.005)
Liberal–conservative index	0.027	0.028	0.028	0.004
	(0.015)	(0.010)	(0.011)	(0.016)
Presidential vote	0.135	0.154	0.154	0.345
	(0.058)	(0.038)	(0.036)	(0.048)
Senate incumbent	0.191	0.168	0.167	0.162
	(0.022)	(0.013)	(0.013)	(0.019)
$\rho\sigma$		−0.155	−0.240	−2.205
		(0.157)	(0.251)	(0.058)

[a]The dependent variable in these analyses is reported vote, where voting for the Republican is coded 1, voting for the Democrat is coded 0, and nonvoters are coded missing. Standard errors are in parentheses. Pseudo-IMR reflects a nonlinear least-squares regression (single stage) incorporating the transformed state response rate as a proxy for the inverse Mills' ratio. Dummy NonRs represents a nonlinear least-squares regression (two stage) using dummied cases to represent nonrespondents, coding statewide aggregate demographic information as predictors in the first stage. Truncation represents a maximum-likelihood estimate of a stochastic truncation model, using statewide aggregate demographic information to predict truncation (single stage).

Several substantive changes occur in the 1990 model upon correction for non-response. Note that the effect of presidential approval jumps with all three corrections—doubling for the pseudo-IMR and dummy nonrespondent approach and trebling for the stochastic truncation correction. The effect of attitudes toward abortion becomes substantively plausible upon correction for nonresponse (although not statistically significant except for the stochastic truncation model). The coding for the abortion variable is such that those opposed to abortion would be coded negative: the implication of the uncorrected model is that opposition to abortion is a (very) weak contributor to voting for Democrats; the corrections reverse sign. Further down the table, the effect of liberal–conservative ideology triples or doubles depending on the correction, and the effect of incumbency increases by 1–3 percentage points.

Table 3 Uncorrected and corrected estimates, model of voting for Senate as a function of issue positions, 1990 Senate Study[a]

		Corrected		
Variable	*Uncorrected (GLS)*	*Pseudo-IMR*	*Dummy NonRs*	*Truncation*
Constant	0.416	0.475	0.471	0.472
	(0.048)	(0.024)	(0.027)	(0.014)
Party ID	0.201	0.171	0.171	0.185
	(0.040)	(0.012)	(0.013)	(0.009)
Presidential approval	0.022	0.039	0.039	0.067
	(0.010)	(0.007)	(0.008)	(0.005)
Personal finances	−0.031	0.003	0.003	−0.024
	(0.005)	(0.009)	(0.010)	(0.004)
National economy	−0.002	−0.001	−0.001	−0.004
	(0.009)	(0.011)	(0.013)	(0.004)
Abortion	0.004	−0.026	−0.022	−0.022
	(0.011)	(0.014)	(0.016)	(0.011)
Help Blacks	−0.010	−0.003	−0.003	−0.024
	(0.004)	(0.006)	(0.006)	(0.004)
Government spending	0.004	−0.004	−0.004	0.014
	(0.003)	(0.004)	(0.004)	(0.003)
Liberal–conservative index	0.010	0.032	0.032	0.019
	(0.006)	(0.009)	(0.009)	(0.006)
Senate incumbent	0.165	0.193	0.193	0.174
	(0.036)	(0.011)	(0.011)	(0.007)
$\rho\sigma$		0.021	−0.094	−0.139
		(0.130)	(0.223)	(0.008)

[a]The dependent variable in these analyses is reported vote, where voting for the Republican is coded 1, voting for the Democrat is coded 0, and nonvoters are coded missing. Standard errors are in parentheses. Pseudo-IMR reflects a nonlinear least-squares regression (single stage) incorporating the transformed state response rate as a proxy for the inverse Mills' ratio. Dummy NonRs represents a nonlinear least-squares regression (two stage) using dummied cases to represent nonrespondents, coding statewide aggregate demographic information as predictors in the first stage. Truncation represents a maximum-likelihood estimate of a stochastic truncation model, using statewide aggregate demographic information to predict truncation (single stage).

The effect of the corrections on the 1992 data is very small, mattering only in the second or third decimal place. The 1992 wave also obtained the best response rate, by far, of the three studies (although still well below the response rates for the traditional Pre–Post Election Studies).

It is worth making some provisional interpretations of the estimates for the coefficient on $\rho\sigma$ under the three different corrections. First, one might be struck by the considerable variability in the magnitude of the estimates in the 1988, 1990, and 1992 corrections. Recall that this coefficient represents the correlation between the error terms in the outcome and selection equation. Since the regressors for the selection equation all differ, it shouldn't be surprising that the correlation of the error terms vary too. Second, one might note that the estimates of $\rho\sigma$ don't usually attain statistical significance. This is also not out of the ordinary for this family of corrections; indeed, the equivalent coefficients for the models evaluated by Heckman (1976, 1979) and Achen (1986) don't attain conventional levels of

Table 4 Uncorrected and corrected estimates, model of voting for Senate as a function of issue positions, 1992 Senate Study[a]

		Corrected		
Variable	*Uncorrected (GLS)*	*Pseudo-IMR*	*Dummy NonRs*	*Truncation*
Constant	0.723	0.735	0.722	0.718
	(0.036)	(0.040)	(0.040)	(0.037)
Party ID	0.100	0.096	0.097	0.097
	(0.029)	(0.030)	(0.030)	(0.030)
Presidential approval	0.007	0.008	0.008	0.008
	(0.013)	(0.013)	(0.014)	(0.013)
Personal finances	0.030	0.031	0.032	0.032
	(0.015)	(0.016)	(0.016)	(0.015)
National economy	0.014	0.012	0.013	0.014
	(0.018)	(0.018)	(0.018)	(0.018)
Abortion	−0.045	−0.040	−0.044	−0.044
	(0.025)	(0.025)	(0.026)	(0.026)
Help Blacks	0.016	0.016	0.016	0.017
	(0.009)	(0.010)	(0.010)	(0.009)
Government spending	−0.182	−0.174	−0.179	−0.181
	(0.061)	(0.061)	(0.062)	(0.062)
Liberal–conservative index	0.040	0.045	0.043	0.043
	(0.013)	(0.013)	(0.013)	(0.013)
Presidential vote	0.382	0.378	0.374	0.373
	(0.056)	(0.057)	(0.057)	(0.057)
Senate incumbent	0.135	0.138	0.137	0.137
	(0.017)	(0.017)	(0.017)	(0.017)
$\rho\sigma$		−0.340	−0.057	−0.363
		(0.385)	(0.308)	(0.452)

[a]The dependent variable in these analyses is reported vote, where voting for the Republican is coded 1, voting for the Democrat is coded 0, and nonvoters are coded missing. Standard errors are in parentheses. Pseudo-IMR reflects a nonlinear least-squares regression (single stage) incorporating the transformed state response rate as a proxy for the inverse Mills' ratio. Dummy NonRs represents a nonlinear least-squares regression (two stage) using dummied cases to represent nonrespondents, coding statewide aggregate demographic information as predictors in the first stage. Truncation represents a maximum-likelihood estimate of a stochastic truncation model, using statewide aggregate demographic information to predict truncation (single stage).

statistical significance either. The key to evaluating the corrections lies in the substantive coefficients of the model.

The substantive interpretation of the effect of the changes is that issues may matter less in voting for Senator than Abramowitz and Segal first identified. The two variables that they thought captured issue voting—attitudes toward the national economy and government spending—become less significant (although still sizable) upon correction for nonresponse. In 1990, these two variables are of negligible importance even in the uncorrected model. Furthermore, we appear to be underestimating the importance of presidential approval, incumbency and liberal–conservative ideology in the 1990 election.

Finally, the reader should note that despite the very different assumptions and methods for correction for sample truncation, in most cases, the corrected coefficients are closer to each other than to the uncorrected model. In combination with the substantive significance

of the changes in estimated effects, the unified picture from the three corrections is that survey nonresponse can matter and that corrections for nonresponse are worth application.

5 Conclusion

None of the three approaches to correction *or* the implicit approach of ignoring nonresponse is devoid of problems. If we do not correct for selection bias, we will have inconsistent estimates to our models to the extent that nonresponse is nonignorable, i.e., where we have any level of nonresponse and outcome models where the error terms are correlated with the process of selection. Clearly, both Senate Studies fail to pass the first criterion: we have sizable nonresponse, more nonresponse than in most of our surveys. And we probably fail to pass the second criterion: if participation in a survey about politics is at all related to participation in politics itself, then it is probable that the error in our substantive models is correlated with the error in selection.

Unfortunately, we also know that the corrections themselves are capable of introducing greater inefficiency to our estimates, spurious error in their own right, or greater sensitivity to model specification. By far the best alternative to escape nonresponse bias is to develop a better understanding (and thus better measures) of nonresponse itself. As the National Election Studies have been doing with the traditional Pre–Post, the NES should continue collecting and disseminating information about nonrespondents. With better understanding about why some people choose or refuse to participate in the surveys, not only do we develop better means of improving response rates and addressing the legitimate concerns of our respondents, but also we gain the prospect of more accurate corrections for nonresponse itself.

But we cannot go backward in time to retrieve lost respondents. What recommendations can we draw from these results? First, we have good reason to suspect problems due to nonresponse in some models, and some attention to nonresponse in these cases is warranted. If participation in surveys is related to participation in politics, then models employing political information, interest, turnout, and vote choice are all susceptible. One of these corrections is relatively simple to employ—the pseudo-IMR correction makes use only of statewide response rates and represents, at most, a few lines of additional code. If one suspects that nonresponse might be a problem, then it is of little additional cost to try, at least, the pseudo-IMR correction. This leads to a second recommendation: if one identifies differences due to nonresponse with one correction, it is prudent to try all three. If the pattern of differences between uncorrected and corrected estimates is the same across the three approaches, it seems reasonable to conclude provisionally that nonresponse affects one's results.

References

Abramowitz, Alan I., and Jeffery A. Segal. 1990. "Beyond Willie Horton and the Pledge of Allegiance: National Issues in the 1988 Elections." *Legislative Studies Quarterly* 15:565–580.

Abramowitz, Milton, and Irene Stegun. 1964. *Handbook of Mathematical Functions*. New York: Dover.

Achen, Christoper H. 1986. *The Statistical Analysis of Quasi-Experiments*. Berkeley: University of California Press.

Benus, Jacob. 1971. "The Problem of Nonresponse in Sample Surveys." In *Working Papers on Survey Research in Poverty Areas*, eds. John B. Lansing et al. Ann Arbor: Institute for Social Research.

Brehm, John. 1989. "How Survey Nonresponse Damages Political Analysis." Paper presented at the 1989 American Political Science Association Meetings, Atlanta, GA.

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

Cobb, Sidney, Stanley King, and Edith Chen. 1957. "Differences Between Respondents and Nonrespondents in a Morbidity Survey Involving Clinical Examination." *Journal of Chronic Diseases* 6(2):95–108.

DeMaio, Theresa J. 1980. "Refusals: Who, Where and Why?" *Public Opinion Quarterly* 44:223–233.

Dillman, Donald. 1978. *Mail and Telephone Surveys*. New York: John Wiley & Sons.

Dohrenwend, Barbara S., and Bruce P. Dohrenwend. 1968. "Sources of Refusals in Surveys." *Public Opinion Quarterly* 32:74–83.

Dubin, Jeffrey A., and Douglas Rivers. 1990. "Selection Bias in Linear Regression, Logit and Probit Models." *Sociological Methods and Research* 18:360–390.

Fay, Robert E. 1986. "Causal Models for Patterns of Nonresponse." *Journal of the American Statistical Association* 81:354–365.

Goyder, John. 1986. "Surveys on Surveys: Limitations and Potentialities." *Public Opinion Quarterly* 50:27–41.

Groves, Robert M., and Mick P. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: John Wiley & Sons.

Groves, Robert M. 1990. *Survey Errors and Survey Costs*. New York: Wiley Interscience.

Groves, Robert M., and Lars E. Lyberg. 1988. "An Overview of Nonresponse Issues in Telephone Surveys." In *Telephone Survey Methodology*, eds. Robert M. Groves et al. New York: John Wiley & Sons.

Hausman, Jerry A., and David A. Wise. 1979. "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment." *Econometrica* 47:455–473.

Hawkins, Darnell F. 1975. "Estimation of Nonresponse Bias." *Sociological Methods and Research* 3:461–488.

Heckman, James J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5(4):475–492.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47:153–161.

House, James, and Sharon Wolf. 1978. "Effects of Urban Residence and Interpersonal Trust and Helping Behavior." *Journal of Personality and Social Psychology* 10:222–226.

Kalton, Graham. 1985. *Compensating for Missing Survey Data*. Ann Arbor, MI: Survey Research Center.

King, Gary. 1989. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. Cambridge: Cambridge University Press.

King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 1998. "Listwise Deletion is Evil: What to Do About Missing Data in Political Science." Paper presented at the 1998 Annual Meetings of the Political Methodology Society, San Diego, CA.

Little, Roderick J. A. 1982. "Models for Nonresponse in Sample Surveys." *Journal of the American Statistical Association* 77:237–250.

Maddala, G. S. 1983. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Nathan, Gad. 1981. "Regression Analysis Under Differential Nonresponse." *Proceedings of the Section on Survey Research Methods*. American Statistical Association, pp. 618–622.

O'Neil, Michael. 1979. "Estimating the Nonresponse Bias Due to Refusals in Telephone Surveys." *Public Opinion Quarterly* 43:218–232.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

Smith, Tom W. 1983. "The Hidden 25%: An Analysis of Nonresponse on the 1980 General Social Survey." *Public Opinion Quarterly* 47:386–404.

Steeh, Charlotte. 1981. "Trends in Nonresponse Rates." *Public Opinion Quarterly* 45:40–57.

Stewart, Charles, and Mark Reynolds. 1990. "Television Markets and U.S. Senate Elections." *Legislative Studies Quarterly* 15:495–524.

Tobin, James. 1958. "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26:24–36.

Weaver, Charles N., Sandra L. Holmes, and Norval D. Glenn. 1973. "Some Characteristics of Inaccessible Respondents in a Telephone Survey." *Journal of Applied Psychology* 60(2):260–262.

Wright, Gerald C. 1990. "Misreports of Vote Choice in the 1988 NES Senate Study." *Legislative Studies Quarterly* 15:543–564.

Wright, Gerald C. 1991. "Errors in Measuring Vote Choice in the National Election Studies, 1952–1988." Paper presented at the 1991 American Political Science Association Meetings, Atlanta GA.