Nonresponse Adjustment in a Longitudinal Survey of African Americans

Monica L. Wolford, Senior Research Fellow, Program on International Policy Attitudes, Center on Policy Attitudes and the Center for International & Security Studies, University of Maryland

Myriam Torres, Senior Research Associate, Program for Research on Black Americans, Institute for Social Research, University of Michigan

Abstract

The growing number of panel studies in the field of social research and the need for quality longitudinal data about minority populations makes it increasingly important to understand the factors associated with whether a respondent will participate in consecutive waves. The analyses presented are from the first full national probability sample of African Americans and its three-wave panel. These data were first collected in 1979-1980 by the Program for Research on Black Americans at the University of Michigan. The original sample size of 2,107 respondents and the large number of measures collected allow for wide variety in the range of measures used to predict nonresponse for each panel year. They reveal some interesting patterns in nonresponse behavior. The questions these analyses address are: what items are important to include in the first year of a panel study in order to arrive at stable nonresponse weights for future study years?; what groups of demographic, content-related, and interest-related variables are useful in adjusting for nonresponse once the survey is completed?; and, are patterns of nonresponse consistent form wave to wave in a minority population?

Introduction

In social science research we are deeply concerned with issues of causality. There is no better method for determining the nature and direction of the causal flow than the observation of changes over time. No cross-sectional study can provide us with comparable information. Yet, the process of conducting longitudinal research presents a unique set of challenges. Two of these are the difficulties of maintaining a representative sample and compensating for the systematic loss of individuals over the course of years. While even the random loss of respondents can create problems by eroding the power of analyses, many authors have noted that the most critical problem is the bias that the disproportionate loss of particular sub-groups introduces into the analyses. This presentation focuses on the dynamics that are associated with longitudinal nonresponse in the National Panel Survey of Black Americans. We will

examine the sources of systematic nonresponse in order both to compensate for it statistically through the use of weights and to develop better procedures for identifying respondents most likely to be lost at later waves for intensified tracking efforts.

The data used for these analyses are from the National Panel Survey of Black Americans (NPSBA), a survey of 2,107 black Americans (Jackson, 1991). The first wave of the study was collected in 1979-80. These respondents were followed up in three data collections in 1987-88, 1988-89, and 1992. NPSBA, more well known as NSBA, is part of a major research project undertaken at the Institute for Social Research to collect and analyze national survey data on African American populations. This survey explores neighborhood-community integration, services, crime and community contact, political ideology and behavior, the role of religion and the church, physical and mental health, and self-esteem. It also examines employment, the effects of chronic unemployment, the effects of race on the job, and interaction with family and friends. In addition, the survey includes measures of racial attitudes, race identity, group stereotypes, and race ideology.

The NPSBA has several unique attributes with regard to nonresponse across multiple waves of data collection. First, it is a national probability sample panel survey of African Americans. As such, it is one of very few surveys with the ability to inform us about nonresponse rates for all African Americans rather than only those living in highly concentrated areas. For instance, while the Americans' Changing Lives (ACL) study found that black informants had a higher second wave response rate than other respondents, they attributed this to the manner in which blacks were over sampled, resulting in the skewing of the sample toward rural southern blacks (Kalton, Lepkowski, Montanari and Maligalig, l990). The relatively large number of black Americans in this study provides us with a strong base for talking about nonresponse even with considerable attrition.

Second, unlike many of the studies in which patterns of nonresponse were examined, this is a study of an individual's attitudes. Proxy interviews or interviews from any of several sources such as those used in the Survey of Income and Program Participation (SIPP) could not be used in this study. This will result in some differences in the patterns of nonresponse as well as an attenuated response rate.

Third, the unusual time frames of the panel data present a challenge. The first wave of the study was collected through face-to-face interviews during 1979 and 1980. The National Survey of Black Americans was not designed as a panel study and therefore only very limited recontact information was collected. During 1987 and 1988 a second wave of data was collected using telephone interviews, to follow up respondents eight years later. Naturally, such a long period of time and limited recontact information has led to considerable attrition. Of the original 2,107 respondents,

102 were known to have died and 935 were reinterviewed for an overall response rate of 47 percent (See Table 1). A study is currently underway to match original respondents with the National Death Index to ascertain how many of those not reinterviewed were in fact deceased at the time of the second wave of data collection. This should result in a more accurate response rate.

The third wave of data was collected one year later and 779 (84 percent) of the second wave respondents were reinterviewed. Eleven respondents were known to have died in the interim. The fourth wave of data was collected three years later during which time an additional 34 respondents were known to have died. Of the 779 respondents recontacted, 652 (84 percent) were interviewed (see Table 1).

Table 1.

NATIONAL PANEL SURVEY OF BLACK AMERICANS -- RESPONSE RATES*

	N	% of Wave I	Response Rate
WAVE I (1979-80)	2107	NA	67.0%
WAVE II (1987-88)	935	46.6%	46.6%
WAVE III (1988-89)	779	39.1%	84.3%
WAVE IV (1992)	652	33.4%	83.6%

*Known deceased were eliminated from calculation of the Response Rate

Finally, while this is a study of non-institutionalized African American adults, tracking information did not allow for the exclusion from the sample frame of original respondents who had since entered an institution. This has lead to an underestimated response rate.

Sources of Nonresponse

In order to more effectively examine the sources of nonresponse, we must characterize the underlying processes leading to nonresponse. We began this procedure by assessing the indicators of nonresponse used in cross-sectional studies. These are general demographics such as region, urbanicity, gender, age, and ethnicity. Usually, this is all the information that is known about the respondent. When we are seeking to account for attrition in the later waves of a survey however, there is an abundance of data about those who did not respond to subsequent data collections. The characteristics used successfully in the few longitudinal studies in which nonresponse has been examined were added to those from cross-sectional studies. These are presented here in the categories we think represent the primary sources of nonresponse.

Demographics: It is most reasonable to assume that many of the dynamics that determine nonresponse in a cross-sectional survey would continue to influence response rates in a longitudinal context. Generally, the data available about respondents in most cross sectional surveys is limited to demographic data collected from an informant or extrapolated from census data. Additionally, demographic indicators of the likelihood that a respondent will change residences and their facility with survey formats and understanding of research agendas must be considered.

Work: Another important determinant of response rates is the nature of the respondent's employment situation. This goes beyond the mere presence or absence of employment, to include the nature of the job search for the unemployed, and the stability of employment for those currently employed. Obviously, those respondents currently looking for work who find work in another locality are more difficult to trace. Similarly, people who lose their jobs are more likely to leave their original location. The type of employment and how the individual feels about that work may have an impact on his or her staying in that job as well as their willingness to continue to participate in the study.

Income and Assets: Financial stability and the resources to weather economic downturns should result in less movement among respondents.

Community Involvement: The degree of involvement the individual has with the community through interaction with neighbors, club memberships, important relationships with others, and civic activities would seem to be related to both an individual's willingness to complete a survey and the likelihood of their being locatable at recontact.

Topic of the Study: To some extent, we might expect that the topic of the study should effect a respondent's willingness to continue to participate. Certainly, as source of

nonresponse, this has the potential to be the most damaging to the viability of the data collected in later waves. The systematic loss of respondents with particular attitudes toward the topic under study introduces a severe bias into any longitudinal analysis.

Affective States: Finally, the individual's affective states and attitudes toward the survey should be related to their likelihood of participating in later waves. Kalton et al. (1990) report that the interviewer's rating of the cooperativeness of the respondent was positively related to further participation. Negative emotional states such as self-reports of depression and tenseness had a negative effect on response rates. These effects are likely to be attenuated in this study by the eight-year time span between waves and the volatility of emotional states.

Analyses

Since no clear theory exists at this time for selecting which variables to use for predicting nonresponse, all of the non-open ended items (approximately 1,200 variables) were given the opportunity to enter the analyses. Following the process outlined by Kalton, Lepkowski, and Lin (1985), the numerous social and demographic characteristics of the respondents measured in the study were reduced to those with highly significant bivariate relationships to nonresponse. Since a large number still remained, these were then reduced further to those which showed a significant relationship with nonresponse in multivariate analyses. While most work on nonresponse relies on examining the marginal distributions, O'Muircheartaigh (1989) notes that particular subgroups in the population are subject to the greatest attrition. We chose therefore to base the final analyses on the SEARCH algorithm, or Automatic Interaction Detector developed by Sonquist, Baker, and Morgan (1973). Those variables with significant multivariate linear relationships with nonresponse were then combined with the categorical variables from the questionnaire, and those variables deemed likely to have nonlinear effects and entered into a SEARCH as joint predictors.

Results

A total of 28 variables entered the final SEARCH analyses, which were run separately for each of the three waves. The analyses shown in Table 2 are weighted to compensate for unequal selection probabilities and cross-sectional nonresponse at Wave I. It is clear that many subgroups do end up under-represented in later waves. In particular, in Wave II the subgroup with one of the lowest response rates were individuals who did not live in a family-owned home and whose household unit needed repair, who primarily desired good interpersonal relations and challenging work in a job, were male and had either a very high or very low household income. This group had a response rate of 13.54%. In contrast, one of the groups in Wave II with the highest response rate were home owners with at least some college education, who

TABLE 2: PERCENT OF VARIANCE IN NONRESPONSE EXPLAINED USING SEARCH

		Wave II	Wave III	Wave IV
Demographics	Home Ownership	3.19%	4.33%	3.60%
	Household Unit in Need of Repair	1.03%	0.90%	1.41%
	Education	1.07%	1.03%	0.00%
	Number of Persons in Household	0.50%	0.00%	1.52%
	R's Age	0.18%	1.02%	0.55%
	Living Three Generational Family	0.00%	0.51%	0.52%
	Urbanicity	0.31%	0.38%	0.23%
	Gender	0.77%	0.00%	0.00%
	Marital Status	0.00%	0.43%	0.00%
	Type of Dwelling	0.23%	0.00%	0.00%
Work	Occupation (9 Categories)	0.55%	0.65%	2.30%
	Most Important Factor in a Job	0.77%	0.44%	1.50%
	Frequency of Absence from Work	0.44%	0.99%	0.37%
	Employment Search Status	0.00%	0.00%	1.75%
	Amount of 1978 Employed	0.00%	0.64%	0.00%
Income and Assets	Household Income	1.22%	1.80%	0.81%
	Personal Income	0.96%	0.33%	1.19%
	Interviewer's Estimation of Household Income	0.00%	0.67%	0.46%
	Paid Salary, Hourly, or Other Type of Wage	0.60%	0.00%	0.00%
Community Involvement	Frequency of Other Church Activities	1.14%	1.41%	0.69%
	Voted in Presidential Election	0.00%	0.76%	1.86%
	Main Romantic Involvement	0.43%	0.00%	0.00%
Topic of Study	Would Support a Black Political Party	0.00%	0.51%	0.45%
	Would Support a Black Platform	0.58%	0.00%	0.00%
Affective States	R was Suspicious -- Open	0.24%	0.75%	0.00%
	R was Bored -- Interested	0.25%	0.45%	0.00%
	R was Hostile -- Friendly	0.00%	0.00%	0.56%
	Frequency of Feeling Depressed	0.30%	0.00%	0.00%
Total Variance Explained		14.76%	18.01%	19.77%

were willing to cross party lines to vote for a platform that favored blacks and who rarely missed work due to personal problems (87.80%).

Many of the characteristics that significantly predict attrition over the waves are demographic variables. Of those which are typically used to predict nonresponse in cross-sectional studies, region, urbanicity, gender, and age, only region was consistently nonsignificant. However, these variables clearly play a lesser role in predicting nonresponse across waves than other characteristics of the respondents.

The overwhelming predictor of nonresponse was whether or not the respondent or their family owned their home, with home owners consistently responding at a higher rate than renters and others. This was the strongest predictor of nonresponse in all three waves. The next most important predictor across all three waves was the condition of the unit in which the respondent lived, as rated by the interviewer. This is primarily due to nonresponse among people who did not own their homes and whose residence was in need of repair.

Age showed a nonlinear relationship to nonresponse with middle-aged respondents most likely to respond to later waves. The level of education a person achieved was strongly associated with nonresponse in the first two reinterviews, but not the third. Individuals with higher levels of education were more likely to respond than those with lower levels of education. This is the same pattern found in the ACL data at the bivariate level, though in that study it does not appear to have been significant in multivariate analyses (Kalton et al., 1990).

Not surprisingly, the nature of an individual's work, its steadiness, and their attitudes toward it were also important predictors of nonresponse. While we did not find a significant effect for employment status by itself, most of the items about working presume this information. Occupation for example, was only collected for those working at the time. It is a strong predictor of nonresponse with individuals who work in white collar jobs most likely to respond and people in service work, craft workers and the unemployed least likely to respond. Intriguingly, the elements of a job, other than pay, that an individual considered to be the most important factor of a job were important predictors of nonresponse. Individuals who valued job factors such as working at their own pace, having good promotion opportunities and good working conditions were most likely to be reinterviewed.

The effects of both household and personal income on response rates tend to be nonlinear in this sample. Those with incomes near the median appear to be most likely to respond. The interviewer's estimate of the household's income however acts in a linear fashion with respondents from households whose income appeared to be low reinterviewing less. In addition to the rate of pay, the manner in which one is paid has

a small effect on nonresponse as well. Respondents who receive a salary or are paid by the unit of work done are more likely to remain in the study.

The strong positive impact on response rates of a respondent's attending activities at a church other than regular services seems to bear out the importance of the respondent's ties to the community in maintaining a sample over time. Similarly, both voting and being either married or having a main romantic involvement are positively associated with being reinterviewed.

We found only minimal effects of the topic of the study on reinterview rates. Only two of the many items that relate to the centrality of being black to the individual were significant predictors of nonresponse. These were whether they were in favor of forming a black political party and whether they would cross party lines to vote for a candidate with the best platform for African Americans. Neither of these entered all three searches however, nor did they account for much of the variance.

The affective component, while weak, did show some impact despite the period of time between waves. Respondents who rated interviewer as open, interested, and friendly were more likely to continue in the study. Conversely, respondents who report feeling depressed frequently were less likely to be reinterviewed.

While the SEARCH technique has some advantages, the use of logistic regression to predict nonresponse is also common. Since SEARCH is designed to identify interaction effects, and logistic regression favors main effects, we have elected to present both here. The results of the stepwise logistic regressions are presented in Table 3. Only the variables from the SEARCH that were significant predictors of nonresponse for at least one of the three waves were entered into the analyses; variables not significant in any of the waves were deleted. Almost half of the variables identified by SEARCH as predictive appear to act only through interactive effects. Of those that remained, the most notable are the changes in the significance of gender and the amount of time during the year previous to the study that the respondent was employed. Their effects were weak and inconsistent in the SEARCH but strongly significant in all three waves in the logistic regression. The state of the respondent's job search, whether or not he or she was employed, unemployed but looking for work, not actively looking but interested in working or not interested in working, also increases in significance, but less dramatically.

Table 4 compares respondents reinterviewed with those who refused, and those we were unable to recontact. Presented are the column percents for the interviews and the three types of interview outcomes by some of the variables that accounted for the greatest variance in nonresponse. A clear pattern is visible in the effect of home ownership on the three types of outcomes. The proportion of home owners is similar

TABLE 3: LOGISTIC REGRESSION OF NONRESPONSE

	WAVE II		WAVE III		WAVE IV	
	B	*S. E.*	*B*	*S. E.*	*B*	*S. E.*
Owns Home vs Rental Condition	-.32***	.06	-.41***	.06	-.40***	.06
Gender	-.53***	.13	-.56***	.13	-.70***	.13
Employment Search Status	-.43***	.06	-.42***	.06	-.48***	.06
Voted - Presidential Election	.45***	.12	.44***	.12	.35***	.12
Living Three Generation Family	.36***	.11	.30***	.12	.39***	.12
Frequency of Other Church Activities	.11***	.03	.11***	.03	.12***	.03
Employed only part of 1978	-.49***	.17	-.54***	.18	-.60***	.18
IW'r Estimation of Household Income	.15***	.06	.13**	.06	.13**	.05
Support for Black Platform	.30**	.12	.26**	.13	.26**	.13
Number of Persons in Household	-.07**	.03	-.07**	.03	-.06**	.03
Most Important Factor in a Job	-.34**	.15	-.34**	.16	-.39**	.16
Education	.15**	.07	.14**	.07	.10	.06
Personal Income	-.08***	.03	-.05*	.03	ns	ns
Urbanicity	-.12*	.07	-.18***	.07	-.07	.07
Frequency of Absence from Work	-.10**	.04	-.02	.04	-.01	.04
Paid Salary / Hourly Wage	.39	.35	.44	.37	.68*	.38
*=<.10 **=<.05 ***=<.01						
Model Chi-Square	258.1***		323.9***		455.3***	
Percent Predicted Correctly	65.9%		69.6%		72.3%	

Table 4.

INTERVIEW OUTCOMES BY NPSBA WAVE II PREDICTORS OF NONRESPONSE			
	I'wed	Refused	Lost
Owns Home vs Rental Condition			
Owns Home	59.8%	59.5%	34.9%
Rents - Good Condition	17.6	20.2	20.0
Rents - Minor Repairs	15.2	13.1	26.7
Rents - Major Repairs	7.4	7.1	18.5
Education			
0-11 years	35.8	67.1	45.5
High School Graduate	33.1	24.4	31.8
Some College	18.5	4.9	16.8
College Graduate	12.6	3.7	6.0
Frequency Other Church Activities			
Frequent	42.6	41.2	30.2
Not Frequent	25.6	31.8	22.2
Not Church Member	31.8	27.1	47.6
Voted in Presidential Election			
Voted	64.2	68.3	43.8
Did not Vote	35.8	31.7	56.2
TOTAL N	935	85	851

in those interviewed versus those who refused, but dramatically different for those not found in the second wave. Though not shown here, this finding is repeated in the third and fourth waves. The same pattern is present for the state of the household unit. The percent of people living in units needing major repairs increases substantially among those who are lost in tracking.

In contrast to home ownership whose effect as a predictor of nonresponse is primarily due to respondents who could not be located, people with low levels of education refuse to be reinterviewed at a much higher rate than others. Conversely, those who have had at least some college education rarely refuse.

The effect of having strong ties to the community should act through a sense of general civic obligation, the means to track individuals who leave the community, and

the decreased likelihood that an individual will leave a community to which they have strong ties. The first dynamic should lead to a lower refusal rate while the second two should decrease the number of respondents lost in tracking. Surprisingly, Table 4 shows no impact for either frequency of involvement in church activities or voting on refusal rates. Both however substantially decrease the rates at which respondents are lost in tracking.

Concluding Remarks

Many of the findings presented here confirm the results reported by Kalton et al. (1990), and McArthur and Short (1985). Even though the sample is a considerably different subpopulation, home ownership remains extremely important to predicting nonresponse. Extending this further is the importance of the condition of the dwelling unit for non-homeowners. While this variable was not present in the other studies, it is reasonable to believe its impact reaches beyond the black population.

The level of education appears to be more critical for nonresponse among black Americans than in the American population as a whole. On the other hand, gender, urbanicity, and marital status are not as central to these analyses as they were to the analyses by Kalton et al. (1990) and McArthur and Short (1985).

The nature of the respondents' occupations and their attitudes toward and behaviors at work are predictors of nonresponse not addressed in previous studies. It may be that the unique employment conditions facing African Americans lead to very different dynamics on these dimensions. The importance to this population however is clear.

The strong relationship between household and personal income, and nonresponse is consistent with the findings of Kalton et al. (1990). However, the nature of the relationship between income and nonresponse is less clear in this study, tending toward nonlinearity, rather than the simple relationship found between low income and low response rates in the ACL data.

We also found variables related to the respondent's involvement in the community to be substantially related to nonresponse. The NPSBA tended to focus more on church-related and political activities than those reported for ACL (Kalton et al., 1990). This is, however, consistent with the broader, more political, and highly central role the church plays in the black community (Brown and Wolford, 1993; Lincoln and Mamiya, 1990).

While not major predictors in this study, variables central to the topic of the study did show some relationship to nonresponse in later waves. Despite the small size of the

variance they account for, their relation to the topic of the study makes them important to account for. Similarly, the impact of the items relating to the affective state of the respondent while small remain important. These confirm the findings from the ACL data, though clearly, the interviewer rating of the respondent's cooperativeness is more central to later nonresponse than the dimensions used here.

The relationships of the predictors to the three interview outcomes generally confirm the patterns seen in the work by Kalton et al. (1990). There are clearly different dynamics driving nonresponse due to refusal and nonresponse due to the difficulty of recontacting some respondents.

While this paper presents only the framework for a theory of longitudinal nonresponse, what we have discovered so far is useful for enhancing efforts to account for nonresponse. First, when developing a questionnaire, this framework identifies the range of important issues for nonresponse. Items should be developed to cover this range or modified to tap it more centrally. Second, we have now a framework for highlighting respondents who are most likely to be lost. This allows us to target tracking efforts more economically. Finally, the adoption of this framework allows the analyst creating nonresponse weights to restrict the focus of his or her search to those variables already identified as associated with longitudinal nonresponse.

Acknowledgement

This research has been supported by the National Institute of Mental Health (Grant #MH48666).

References

Brown, R. E. and Wolford, M. L. (1993). Religious resources and African American political action. *National Political Science Review.*

Kalton, G., Lepkowski, J., and Lin, T. K. (1985). Compensating for wave nonresponse in the 1979 ISDP research panel. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 372-377.

Kalton, G., Lepkowski, J., Montanari, G. E., and Maligalig, D. (1990). Characteristics of second wave nonrespondents in a panel survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 462-467.

Jackson, J. S. (Ed.). (1991). *Life in Black America*. Newbury Park, CA: Sage Publications.

Lincoln, C. E. and Mamiya, L. H. (1990). *The Black Church and the African American Experience.* Durham: Duke University Press.

McArthur, E. and Short, K. (1985). Characteristics of sample attrition in the Survey of Income and Program Participation (SIPP). *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 366-371.

O'Muircheartaigh, C. (1989). Sources of nonsampling error: Discussion. In *Panel Surveys*, Kasprzyk, D., Duncan G., Kalton, G., Singh, M. P. (Ed.), pp. 271-288. New York: Wiley and Sons.

Sonquist, J. A., Baker, E. L., and Morgan, J. N. (1973). *Searching for Structure.* Ann Arbor, Michigan: Institute for Social Research.