

Respondent Characteristics That Are Predictive of Financial Item Nonresponse

Laura Flicker, MS¹; Jill Ruppenkamp, MS¹; Barbara Bibb, MA¹, Sarah Riley, PhD²; Mark Lindblad, PhD²

¹RTI International Research Triangle Park, NC; ²University of North Carolina at Chapel Hill, NC



1. Statement of the Problem

Item nonresponse is a concern for researchers making inferences about target populations. Statistical imputation techniques assume that data are missing at random, yet that assumption is often violated. Item nonresponse has been shown to vary by subgroup for survey content related to sexual orientation (Gruskin, et al., 2001), sexual behavior (Kupek, 1998), drug use (Wolfe, 2003), health (Blaum and Liang, 1996; Slymen, et al., 1994), and income (Gruskin, et al., 2001; Turrell, 2000; Dengler, et al., 1997; Schröpfer, 2004).

Moreover, some of the same respondent characteristics that affect unit-level nonresponse also contribute to item nonresponse. That may compound problems for data quality, especially for income estimation (Turrell, 2000). Education, age, and gender have been shown to independently contribute to item nonresponse (e.g., Gruskin, et al., 2001; Slymen, et al., 1994; Dengler, et al., 1997; Blaum and Liang, 1996). Elsewhere, gender has been shown to interact with other characteristics to affect item nonresponse (Turrell, 2000) or to have no independent or interactive effects on item nonresponse (Wolfe, et al., 2008). Race in some cases contributes to nonresponse after controlling for socioeconomic status (Gruskin, et al., 2001; Kupek, 1998).

The goal of the present research is to determine whether specific respondent characteristics contribute to item nonresponse for sensitive data. For this analysis, the missingness of financial data is investigated.

2. Background to Panel Survey

- The Center for Community Capital at the University of North Carolina at Chapel Hill conducts the Community Advantage Panel Survey (CAPS) with low-to-moderate-income renters and owners, with funding from the Ford Foundation.
- The CAPS goals are to assess the social and financial impacts of homeownership and to track households' wealth over time.
- RTI International has conducted data collection by annual telephone and in-person interviews.
- Original respondents were identified at baseline as having signed the mortgage contract or rental lease.
- The wave analyzed for this work was the fourth year of participation among renters and fifth for owners. It was conducted by telephone during 2007

3. Methods

- One central goal of the renter and owner panels is to track changes to household wealth over time. The issue of nonresponse at the unit or item level is very important for analysis, particularly for financial questions.
- The 2007 computer-assisted telephone interview administered up to 19 questions as relevant for each respondent's situation (see Appendix). They included:
 - annual income
 - monthly rent or mortgage
 - monthly utilities
 - vehicle loan balances
 - credit card debt
 - money transfers outside household
 - other financial information
- Analytically our outcome measure was a flag indicating whether or not 19 financial questions administered during the telephone interview were answered completely. That flag equaled 1 if at least one question was answered Don't Know or Refused, or 0 for no missing data.

4. Findings

Our analysis included 2,982 respondents; only interviews with original respondents from baseline were included, since proxy interviews may know less about household finances.

The respondent profile in the 2007 wave was as follows:

- 58 percent female
- 31 percent currently renting
- 66 percent living in the South
- 55 percent under age 40
- 56 percent living with children
- 30 percent having a bachelor's degree or higher; 10 percent had not earned a high school diploma or general equivalency diploma (GED)

Crosstabulations suggested various possible contributors to item nonresponse. Figures 1–6 illustrate distributions of key predictors by response status.

However, successive logistic regression models revealed that marital status, presence of children in household, and region (not shown here) did not predict item nonresponse, nor did including them affect other predictors. Therefore they were dropped from our models. Race and ethnic categories were not predictive, with the exception of non-Hispanic Blacks. Therefore we dichotomized race into non-Hispanic Black and all others. We dichotomized age (above and below 40 years old) because other classification schemes did not show much effect.

Table 1. Prediction of financial item nonresponse by respondent characteristics

Respondent Characteristics	Model A			Model B		
	Estimate	Odds Ratio	Prob.	Estimate	Odds Ratio	Prob.
Intercept	-0.53	—	.0007	-0.41	—	.0115
Renter	-0.19	0.83	.0732	-0.29	0.75	.0079
Male	-0.28	0.76	.0058	-0.23	0.79	.0200
Non-Hispanic Black	0.48	1.62	<.0001	0.46	1.58	<.0001
Over 40 years old	0.35	1.41	.0003	0.29	1.34	.0029
High school only	-0.69	0.50	<.0001	-0.66	0.52	<.0001
Some education post high school	-0.78	0.46	<.0001	-0.70	0.50	<.0001
Bachelor's	-0.79	0.46	<.0001	-0.64	0.53	.0002
Some education post bachelor's	0.95	0.39	<.0001	0.76	0.47	.0002
Income \$40–59.9K ^a	—	—	—	0.76	0.76	.0273
Income \$60–74.9K ^a	—	—	—	0.73	0.73	.0743
Income \$75K and above ^a	—	—	—	0.62	0.62	.0059

^aThe income used in the logistic model came from the most recent wave of data collection for which the respondent provided an answer.

Both models shown in Table 1 obtained similar findings for predictor variables they shared:

- Males were about three-quarters as likely as females to have missing data.
- Non-Hispanic Blacks were 1.6 times as likely to have missing data.
- Respondents over age 40 were at least 1.3 times as likely to have missing data compared to younger respondents.
- Respondents with at least a high school education were about half as likely to have item nonresponse as those with less education.

Model A found that gender, race, age, and education were significant predictors of item nonresponse, but renter status was not. Recalling that other studies found low income to predict item nonresponse, we added household income to Model B. That income data came from an earlier wave of data collection. By controlling for income, Model B indicates that renter status was a significant predictor. Renters were three-quarters as likely to have missing data on the financial questions.

Figure 1. Frequency of Missing Data by Gender

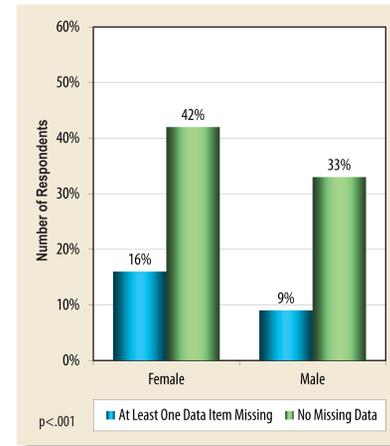


Figure 2. Frequency of Missing Data by Age

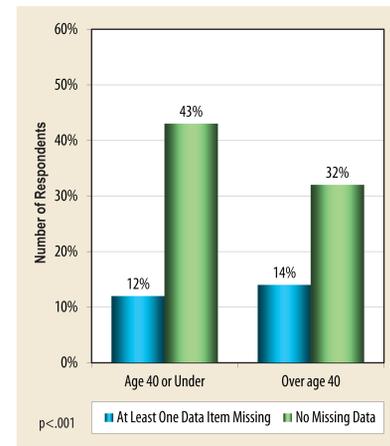


Figure 3. Frequency of Missing Data by Renter Status

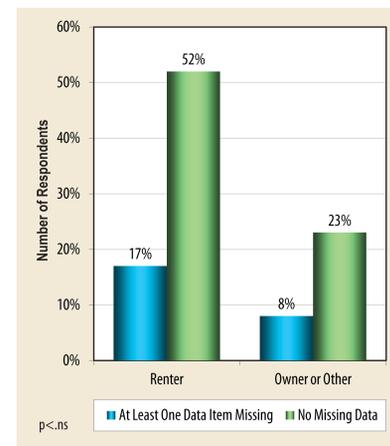


Figure 4. Frequency of Missing Data by Education

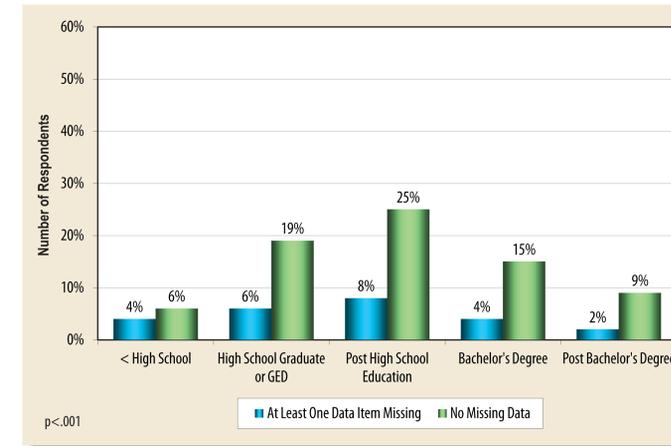


Figure 5. Frequency of Missing Data by Race

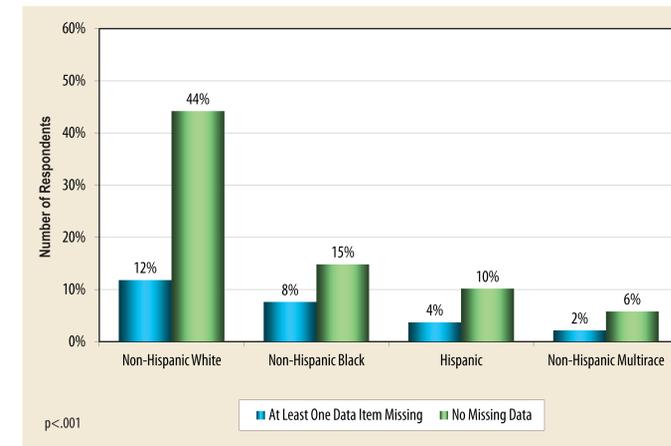
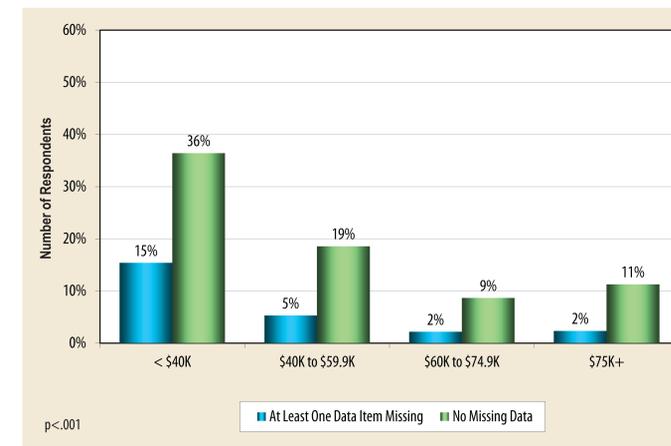


Figure 6. Frequency of Missing Data by Income, as Reported in Past Data Collection



5. Discussion

Item nonresponse was inversely related to income, so that respondents from higher income households were more likely to answer financial questions completely. This supports findings by other researchers. In accordance with other literature, males had lower rates of nonresponse, whereas non-Hispanic Blacks and those with lower educational attainment had higher item nonresponse. Of interest is the fact that renters had less missing data than owners. Perhaps the financial commitment of home ownership or its implied relationship to assets increases privacy concerns among those respondents.

References

Blaum, C., & Liang, J. (1996). Item non-response in the Longitudinal Study of Aging (LSOA). *Journal of the American Geriatrics Society* 44(9): P143.

Dengler, R., Roberts, H., & Rushton, L. (1997). Lifestyle surveys—the complete answer? *Journal of Epidemiology and Community Health* 51(1): 46–51.

Gruskin, E., Geiger, A., Gordon, N., & Ackerson, L. (2001). Characteristics of nonrespondents to questions on sexual orientation and income in a HMO survey. *Journal of the Gay and Lesbian Medical Association* 5(1): 21–24.

Kupek, E. (1998). Determinants of item nonresponse in a large national sex survey. *Archives of Sexual Behavior* 27(6): 581–594.

Schröpfer, J.P. (2004). Respondent behavior in panel studies: A case study for income nonresponse by means of the German Socio-Economic Panel (SOEP). *Sociological Methods & Research* 33(1): 118–156.

Slymen, D.J., Drew, J.A., Wright, B.L., Elder, J.P., & Williams, S.J. (1994). Item non-response to lifestyle assessment in an elderly cohort. *International Journal of Epidemiology* 23(3): 583–591.

Turrell, G. (2000). Income non-reporting: Implications for health inequalities research. *Journal of Epidemiology and Community Health* 54(3): 207–214.

Wolfe, E.W., Converse, P.D., & Oswald, F.L. (2008). Item-level non-response rates in an attitudinal survey of teachers delivered via mail and web. *Journal of Computer-Mediated Communication* 14: 35–66.

Wolfe, E.W. (2003). Using logistic regression to detect item-level non-response bias in surveys. *Journal of Applied Measurement* 4(3): 234–248.

Appendix: CATI Questions

- Household income
- Monthly rent payment
- Monthly mortgage payment
- Amount of rent paid out of pocket
- Debt at time of bankruptcy
- Filing cost for bankruptcy
- Money sent to family or friends
- Amount due on vehicles
- Vehicle worth if traded
- Credit debt
- Financed purchases debt
- Cost of unexpected emergency over \$500 (if applicable)
- Balance owed on unexpected emergency
- Monthly average winter heat and electricity
- Monthly average summer electricity
- Monthly vehicle loan payment
- Amount home sold for
- Money saved since previous interview
- Money received in prize/settlement (if applicable)

Presented at the American Association for Public Opinion Research 64th Annual Conference, May 14–17, 2009, Hollywood, Florida.

The authors gratefully acknowledge the assistance of Kevin Wang, PhD, and Rodney Baxter, PhD, of RTI International.