ASSESSING NONRESPONSE BIAS IN REPEATED TELEPHONE SURVEYS

Randal ZuWallack, Ronaldo Iachan, Leslyn Hall
ORC Macro, 11785 Beltsville Dr., Calverton, MD 20705

Key Words: Longitudinal survey, Nonresponse, Bias.

Introduction and Background

ORC Macro investigated the impact of nonresponse in surveys commissioned by the Department of Housing and Urban Development (HUD) as part of the Fair Market Rent (FMR) program.[1] Based mostly on the FMRs, the national payment standard for assisted rental housing units, HUD provides rent assistance to hundreds of thousands of individuals and families each year. The current measure is the 40th or 50th percentile gross market rent for eligible units. [2]

The FMR Regions Surveys are dual-mode longitudinal surveys measuring rent for metropolitan and non-metropolitan areas in 10 regions of the country. The longitudinal component of the Regions Surveys offers an opportunity to examine differential nonresponse across different geographic areas—possibly as large as regions but also as small as exchange areas—and across time (i.e., over repeated inclusions in the survey for sample households). This analysis is a first step towards further understanding the sources of nonresponse bias and characterizing areas that may have lower or higher response rates. The latter objective may help design telephone surveys that over-sample some exchange areas taking into account the differential response rates.

The purpose of the Regions Surveys is to measure annual rent change for large geographic areas, or regions of the United States. The Regions are defined in Table 1. The regional estimates are used to annually adjust rent levels for specific areas.[3] Fielding takes place over the course of approximately two months, usually in the fall. The survey is a multi-mode study, with data collected by telephone and by mail. In addition, each annual sample has two components, an RDD sample and a longitudinal, or re-contact sample. Sample members from the RDD component who complete an interview are selected for the longitudinal component in the following year. Housing units are eligible for the survey for a total of three years, one year as RDD and two years as a re-contact. Re-contact sample members are initially mailed a letter and survey instrument. Mail nonresponders are then contacted via telephone and are mailed two follow-ups.

Nonresponse Approach

As a general problem, survey nonresponse is a difficult topic to address, primarily since little is known about the nonrespondents. One approach to evaluating nonresponse targets individual nonresponding units directly, typically involving personal follow-up visits or re-interview surveys to establish patterns of nonresponse. This micro-level approach provides useful information concerning the nonrespondents but is very costly. A less expensive alternative is to compare the sample distributions to known demographic, socioeconomic, and geographic distributions. Assuming the selected sample is representative of the population, deviations from the population distributions may be attributed to differential nonresponse. This macro-level approach evaluates unit nonresponse for large domains but may not reveal patterns of nonresponse.

Specific to the FMR surveys, nonresponse evaluation is further complicated by the streamlined questionnaires including only survey eligibility and rent-related questions. The questionnaires do not include demographic and socioeconomic questions typically used in evaluating nonresponse. So not only do we know little about the nonrespondents, we know little more about the survey respondents. Compounding the situation are the extremely exclusive eligibility criteria: eligible units include only non-public housing, two-bedroom rentals. Refusals typically occur at the screening stage before establishing survey eligibility.

The design of the Regions Surveys provides the opportunity to evaluate nonresponse at the individual level without the need for costly recontacts. With two thirds of the sample in the current year sample consisting of the same respondents as the previous year, we have information about the current year nonrespondents.

[1] The research for this paper was funded in part by the Department of Housing and Urban Development. The authors would like to thank the Office of Policy Development in the Economic Market Analysis Division of HUD. We would also like to thank Joe Riley (Director, Office of Policy Development) for sharing his insight to the FMR program and his comments on this paper.

[2] Federal Register, Vol 68, No. 190, Wednesday, October 1, 2003, pp. 56704-56707.

[3] A second paper, *Analysis of Nonresponse in Cross-sectional RDD Survey*, evaluates nonresponse in the FMR Area specific surveys.

In fact, we have information to calculate gross rent values. As we previously stated, rent is the one purpose for this survey so we can evaluate the impact of survey attrition on the rent estimates.

Initial Bias Assessment

This section examines whether the FMR gross rent estimates compare to other established rent estimates. Before evaluating patterns of nonresponse, we first want to know if FMR rent estimates are potentially biased, and if so, to what extent. Large unexplained differences may indicate bias due to nonresponse or other potential bias-inducing survey errors.

For this assessment, we use Census 2000 as a comparison measure, or target. The Census Bureau does not publish median gross rent for two bedroom rental units, but they publish a gross rent frequency distribution for two bedroom rentals. We take two approaches in comparing the two sources of rent data:

1) Use linear interpolation to estimate the target (Census 2000) median and construct bias estimates by subtracting HUD gross rent estimates from the target rent estimates; and
2) Compare the frequency distributions between the HUD surveys and Census 2000 using goodness of fit tests.

As seen in Figure 1, the HUD Regions survey tends to underestimate both the 40th and 50th percentile rents, relative to the census. For all 20 strata, we see significant underestimates at the 5% significance level. Chi-squared tests are consistent with the rent comparisons results. With HUD as the observed distribution and the Census 2000 as the expected distribution, we see significant test statistics in all strata at the 5% level. All (20) survey strata exhibit a shortage of higher income respondents ($750.00 or $1000+), and we have a shortage of mid-high rent respondents ($500.00-749.99) in all but three strata.

This analysis indicates that there are differences between the HUD estimates and census estimates, but the difference may not be attributable to non-sampling or survey errors, but rather different measurement scope and procedures. The FMR surveys targets two-bedroom apartments that are not run by a Public Housing Authority (PHA): these units are screened out of the survey. On the other hand, the Census estimates include PHA units. In addition, HUD and Census calculate gross rents differently, likely a more significant contributor to the differences in the estimates. The Census calculates gross rent by adding tenant estimates of utility costs to contract rents. HUD RDD regional surveys construct gross rents by adding a computer-modeled regional utility cost updated with CPI utility changes to contract rent values. These differences in scope and measurement may account for much of the differences between HUD and Census estimates. In fact, over-reporting of utilities may be severe and varies from area to area as well as utility to utility (Chakrabarty, 1995; Tippett, 1984.) This over-reporting results in census estimates having an upward bias—likely explaining much of the differences.

Nonresponse Assessment

This analysis is focused on comparisons of nonresponse across different classes and the development of models for the response propensity. These models express the probability of a unit responding as a function of characteristics, or predictors, that are potentially related to the propensity to respond. This analysis not only advances the understanding of possible bias sources and components, but also supports the development of methods to compensate and adjust for biases.

Using a logistic regression model, we estimate a unit's likelihood to respond as a function of the known characteristics. We classify the survey outcomes into three broad response categories:

- *Complete interview.* The phone number is eligible for the survey and an interview was completed.
- *Known ineligible.* The phone number is determined to be ineligible for the survey.
- *Unknown eligibility*. Eligibility for the survey is not determined due to such things as refusals, hang-ups, or non-contacts.

To better understand differential nonresponse, we model the probability of a response and the probability of unknown eligibility. Theoretically, the probability of an ineligible case is obtained by subtracting the probability of a response and the probability of unknown eligibility from one.

For propensity modeling, we used a logistic regression model, $E(y \mid \underline{x}) = P(y = 1 \mid \underline{x}) = \frac{e^{\underline{x}'\underline{\beta}}}{1 + e^{\underline{x}'\underline{\beta}}}$, where y is a binomial variable, $\underline{x}' = (1, x_1, \ldots x_p)$ is a vector of known independent variables and $\beta = (\beta_0, \beta_1, \ldots, \beta_p)$ is a vector of regression

coefficients. The coefficientss are estimated through maximum likelihood estimation, where $\underline{x}'\underline{\beta} = \ln\left[\frac{P(y=1|\underline{x})}{P(y=0|\underline{x})}\right]$.

To explore the added longitudinal component in the Regions survey, we use design variables and variables from the previous survey iteration as explanatory variables. We focus on two models, one with both the RDD and longitudinal components, and the other a longitudinal model focusing on attrition of previous year respondents. The possible covariates for these models are:

1) Full sample model: Metro status, sample type (RDD vs. recontact) and regions
2) Longitudinal model: 2002 Metro status, 2002 Interview Type (mail vs. phone) and 2002 length of residence

Full sample models

Survey design variables are the independent variables in modeling response probability. These variables include region, metro status, and re-contact status. These models provide informative comparisons across design features. Any response differential is unlikely to effect the rent estimates, particularly change, since estimates are calculated within each stratum. However, these models play a role in explaining the degree of regional response differential.

Who are we not reaching? - Probability model for unknown eligibility status All covariates introduced into the model are significant in modeling unknown status probability. Phone numbers in metro areas are 13.5% more likely to result in unknown status as nonmetro phone numbers. Alternatively, but not unexpectedly, 2003 recontacts are 22.0 percent less likely to result in unknown status as the RDD portion of the sample. With region ten, Pacific Northwest and Alaska, as the base, the relative odds of resulting in unknown status range are highest in Regions 5, 7 and 8, where it nears 90% more likely.

Who are we reaching? – Probability model for complete interview. All covariates introduced into the model are significant in modeling completion probability. Phone numbers in metro areas are 24.7 percent more likely to result in a completion as nonmetro phone numbers. The 2003 recontacts are 48 times more likely to result in a completion than the RDD portion of the sample. This is not unexpected as the previous year acts a screening sample so eligibility rates are much higher. With region ten, Pacific Northwest and Alaska, as the base, the relative odds of completing are highest in regions 1-3.

Longitudinal models

In building the longitudinal models, we include only the 2002 respondents who are eligible for the 2003 survey. Completion rates are much higher and unknown rates are much lower since these 2003 respondents are essentially prescreened in the 2002 survey. Overall, there are 13,310 2002 respondents eligible for the 2003 survey. Of these, 38.6 percent resulted in a completed interview and 17.8 percent resulted in unknown status. Not only is the longitudinal component ideal for achieving the survey goals, which is estimating annual rent increases or decreases, but this design feature substantially increase the incidence of eligible cases.

Who are we not reaching? - Probability model for unknown eligibility status. The significant effects are 2002 interview type, 2002 length of residence, and 2002 metro status. Recontact status and rent have no significant effect on unknown status. The 2002 phone respondents are 55.8 percent more likely to result in unknown status as the 2002 mail respondents. Respondents living in a metro area are 13.2 percent more likely to result in unknown status. On the other hand, the 2002 respondents who moved within twelve months are 13.9 percent as likely less those who were living in the same residence for at least twelve months.

Overall, the covariate pattern with the highest 2003 unknown rates is the 2002 phone respondents in a metro area who were living at the same residence 12 months prior to the 2002 interview. This rate is 20.3 percent, with the model predicted probability of 20.1 percent. The covariate pattern with the lowest 2003 unknown eligibility rate is the 2002 mail respondents in a non-metro area who were living at the same residence 12 months prior to the 2002 interview. This rate is 12.5 percent, actual and predicted.

Who are we reaching? – Probability model for complete interview. The significant effects are 2002 interview type, 2002 length of residence, and 2002 re-contact status. Metro status and rent have no significant effect on completion status. The 2002 phone respondents are 29.7 percent less likely to result in a complete as the 2002 mail respondents. The 2002 respondents who moved within twelve months are 27.0 percent less likely as those were living in the same residence for at least twelve months. Conversely, those who are in

sample for the third time are 19.9 percent more likely to result in a completed interview than those in sample for the second time.

Class comparisons

The class comparisons focus on differences between the two survey components and on differences between the 2003 longitudinal respondents and nonrespondents

We see very little difference between the 2003 longitudinal respondents and RDD respondents in terms of rent. There are no statistically significant differences overall or by metro status but a few differences by regions and by stratum.

Response rates for eight of the ten regions are consistently around 39 to 40 percent with lower rates in two southern regions—two regions with larger than average differences with census estimates. Response rates by metro status are similar, 39.2 percent in non-metro areas and 38.0 percent in metro areas.

Response rates across levels of 2002 rent response are slightly different. The lowest response rates occur in the extreme categories for 2002 rent while the highest response occurs in the middle rent categories. There is no significant difference for mean 2002 rent between 2003 respondents and nonrespondents.

Discussion and Future Research

We set out to gain insight into the quality of rent estimates in the FMR Regions surveys. We found significant differences between FMR regions and Census 2000 estimates of gross rent, but overreporting of utilities in the census raises questions about gross rent accuracy.

There seems to be no differential attrition with respect to gross rent. This is a critical finding for the FMR program as differential response by rent levels could severely bias rent change estimates. There are some differences between regions and metro status, but this is of less concern considering rent change is estimated for each region and metro status.

For furthering this analysis, we identified the following topics for possible nonresponse research using the HUD survey data.

Polytomous response model. To evaluate nonresponse we created dichotomous variables for response and unknown eligibility. Then, we modeled each one and then solved for the third dichotomous variable known eligibility. Possibly a more natural fitting model would be to jointly model a polytomous response variable distinguishing the several response (status) categories.

Expanding over time. For this analysis, we analyzed attrition for one cycle of the survey. Many cycles of survey data exist and combining several years of data may strengthen the analysis. A next step is possibly to combine several years of data and analyze nonresponse at time t given the response data at time t-1.

A proactive approach to using the data concerning attrition is to develop a survival model using multiple cycles of survey data. Rather than analyzing nonresponse at time t given t-1, we can gain valuable insight by focusing on subsequent survey years given the response data at time t. Results that suggest differential response probabilities at time t+1 and t+2, given response data at time t may be used to evaluate current operations and possible develop methods for reducing attrition.

References

Chakrabarty, Rameswar P. and Georgina Torres, (1995) "American Housing Survey: A Quality Profile." Current Housing Reports H121/95-1, U.S. Department of Housing and Urban Development and U.S. Department of Commerce.

Federal Register, 68, 190, October 1, 2003, pp. 56704-56707.

Tippett, Janet A. (1984) "An Experimental Project to Improve the Reporting of Selected Costs in the 1980 Census." *Proceedings of Survey Research Methods Section, American Statistical Association*, Alexandria, VA, American Statistical Association, pp. 323-328.

Woodruff, Ralph S., (1952) Confidence Intervals for Medians and Other Position Measures, *Journal of the American Statistical Association*, Volume 47, Issue 47, pp. 635-646

ZuWallack, Randal, Leslyn Hall and Ronaldo Iachan, (2004) "Analysis of Nonresponse in Cross-Sectional RDD Surveys" *Proceedings of Survey Research Methods Section, American Statistical Association*, Alexandria, VA, American Statistical Association, to appear.

Region	HUD Regions	Metro status	Stratum
1	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont	Metro	1
		Non-metro	2
2	New York, New Jersey	Metro	3
		Non-metro	4
3	Delaware, Maryland, Pennsylvania, Virginia, West Virginia	Metro	5
		Non-metro	6
4	Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee	Metro	7
		Non-metro	8
5	Illinois, Indiana, Minnesota, Ohio, Michigan, Wisconsin	Metro	9
		Non-metro	10
6	Arkansas, Louisiana, New Mexico, Texas, Oklahoma	Metro	11
		Non-metro	12
7	Iowa, Kansas, Missouri, Nebraska	Metro	13
		Non-metro	14
8	Colorado, Montana, Wyoming, South Dakota, North Dakota, Utah	Metro	15
		Non-metro	16
9	Arizona, Nevada, California	Metro	17
		Non-metro	18
10	Alaska, Idaho, Oregon, Washington	Metro	19
		Non-metro	20

Table 1. HUD Regions stratum definitions

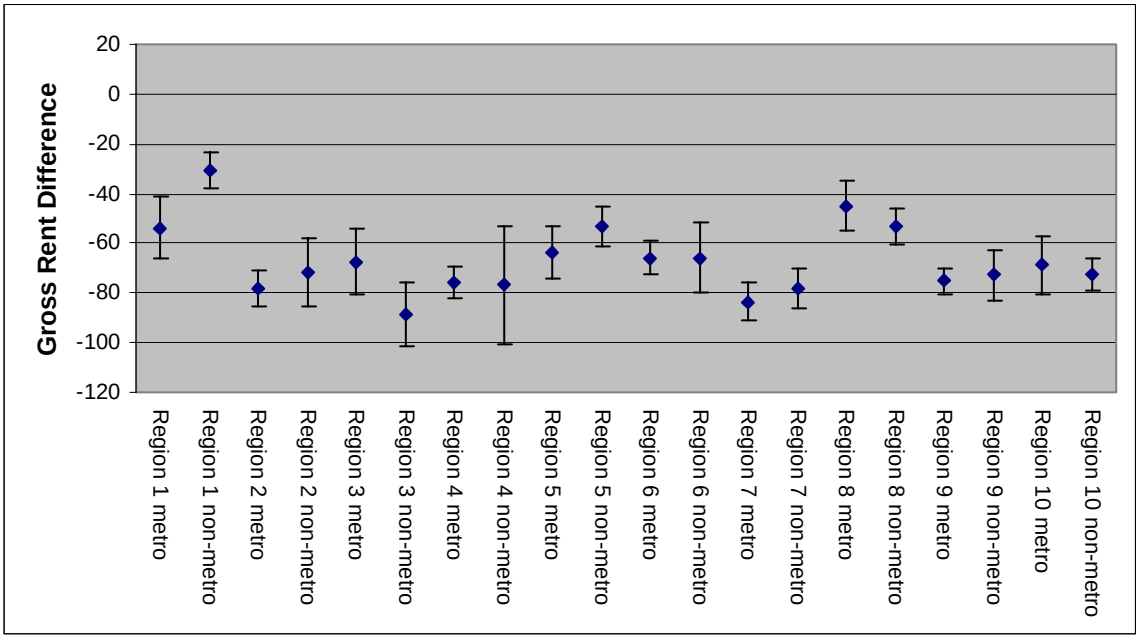
Table 2. Actual and predicted completion rates for longitudinal sample

2002 Recontact status	2002 Interview type	2002 Length of residence	2002 Respondents	2003 Completion Rate	
				Actual	Predicted
No	Phone	12+ months	6792	0.38	0.38
		0-12 months	3157	0.31	0.31
Yes	Mail	12+ months	1409	0.52	0.51
		0-12 months	398	0.39	0.43
	Phone	12+ months	1476	0.42	0.42
		0-12 months	78	0.41	0.35

Table 3. Actual and predicted unknown status rates for longitudinal sample

2002 Metro area	2002 Interview Type	2002 Length of Residence	2002 Respondents	2003 Unknown Rate	
				Actual	Predicted
No	Mail	12+ months	704	0.1250	0.1253
		0-12 months	184	0.1467	0.1098
	Phone	12+ months	4160	0.1829	0.1824
		0-12 months	1568	0.1556	0.1612
Yes	Mail	12+ months	705	0.1277	0.1395
		0-12 months	214	0.1308	0.1225
	Phone	12+ months	4108	0.2032	0.2017
		0-12 months	1667	0.1788	0.1787

Figure 1. Difference Between HUD and Census 2000
(a) 40th Percentile Gross Rent



(b) Median Gross Rent Difference