USING AUXILIARY INFORMATION TO INVESTIGATE NONRESPONSE BIAS

Louise Woodburn
Internal Revenue Service (R:S:P), P.O. Box 2608, Washington DC 20013-2608

Introduction

In sample surveys, adjustments to account for nonresponse have to be considered. Typically, the lack of information on the nonrespondents limits the extent of the adjustments which can be implemented. The introduction of auxiliary frame information on both respondents and nonrespondents makes more sophisticated adjustments feasible.

This paper describes the investigation into the possible bias due to unit nonresponse in the 1989 Survey of Consumer Finances (SCF). For the 1989 SCF, auxiliary frame information is available for both respondents and nonrespondents so that detailed comparisons can be made and more complex adjustments can be considered. Additionally, this paper extends the discussion from the use of auxiliary information for nonresponse adjustments to the use of such data for a post-survey weight examination for the 1989 SCF.

Organizationally, this paper is divided into seven sections. This introductory overview of the paper is the first section. In the second section, background information on the content and sample design of the 1989 SCF is provided as well as specifics on the constraints on the use of the auxiliary data. The response outcomes for the 1989 SCF are detailed next. In order to depict how the use of auxiliary data for the 1989 SCF fits into the big picture of adjustments for nonresponse, current adjustment techniques for differing levels of information are briefly summarized in the fourth section. The investigation into the potential bias due to unit nonresponse for the 1989 SCF is summarized next followed by the results of the use of the auxiliary data for a post-survey weight examination. The final section provides some concluding thoughts about use of auxiliary data for post-survey weight adjustments, in general, and some additional possibilities being considered for the 1989 SCF, in particular.

Background

The 1989 SCF is one in a series of surveys on the financial characteristics of households sponsored by the Federal Reserve Board (FRB) and supported by other federal agencies, including the Statistics of Income Division (SOI) of the Internal Revenue Service. The Survey Research Center (SRC) at the University of Michigan conducted the survey interviews and played a major role in the design of the sample. Data on the survey include detailed data on income, assets, liabilities, and employment. Additionally, pension data were collected from the pension providers of respondents who gave permission for their employer to be contacted.

The sample for the 1989 SCF consists of a new cross-section sample and a panel sample (Heeringa and Woodburn, 1991). The new cross-section sample is selected from a dual frame design which consists of a multi-stage area probability frame and a list frame based on a probability sample of tax filing units whose returns had been previously selected in the 1987 SOI Individual Statistics' program. The supplemental sample from the SOI list frame includes a higher proportion of wealthier individuals in order to provide a better representation of the upper tail of the wealth distribution. This was the approach used for FRB's landmark 1963 Survey of Financial Characteristics of Consumers (Projector and Weiss, 1966) and 1983 Survey of Consumer Finances (Avery et. al., 1988). The list sample for the 1989 SCF was drawn from the 1987 SOI Individual data file, which contains abstracted tax form data for a stratified random sample of approximately 108,000 1987 Form 1040 and Form 1040A tax returns selected from the over 100,000,000+ individual income tax filings for 1987 (Internal Revenue Service, 1990).

The selection of the list sample from the administrative data was quite complex. Among the issues were constraints due to disclosure concerns, processing constraints and sample design complexities. In order to address the disclosure concerns, high standards of disclosure provisions were developed for the 1989 SCF. Statistical and research uses of SOI data are closely regulated to guarantee that individuals (and other entities) will remain protected against any and all disclosure of their financial and tax data (e.g., Wilson and Smith,

1983). To this end, for the 1989 SCF, contractual agreements between SOI, FRB, and SRC have been written which clearly specify the limitations on the use of the administrative data. The terms of these agreements are written so as to guarantee the privacy rights of the individual taxpayers.

Processing constraints complicated the selection of the list sample and the preparation of the selected cases for release to the field. These constraints were mainly due to the availability of the data files of interest and the lead time necessary to access the files. Successively smaller subsets of the original 1987 SOI file of 108,000 were created during the process so that the final selection for the list sample was made from a file of about 10,000.

The complexities of the sample design are due mainly to the uncertain relationship of the administrative data to the survey data, the desire to most efficiently select a sample to estimate the wealth distribution, and the implementation of the dual frame approach (Heeringa, Juster, and Woodburn, 1991). In order to select a sample to efficiently estimate the wealth distribution, it was necessary to develop a measure of wealth, a wealth index, using the administrative data. Capitalized income flows from the SOI file were used to compute the wealth index. The sample design for the 1989 SCF incorporated the stratification of the SOI sample based on a measure of income with the final stratification based on the computed wealth index. Six wealth index strata were chosen with a cap on the top stratum of $250 million. This limit was set at $250 million in recognition of the difficulty of measuring the wealthiest households with a survey and that such data are available elsewhere (e. g., *Forbes,* 1989). The use of the dual frame approach adds a layer of complexity to the computation of the final sampling weights. The extent of this complexity depends on how difficult it is to compute selection probabilities in the list frame for the cases from the area probability frame.

The processing constraints, disclosure concerns and sample design issues all had to be addressed in order to put the 1989 SCF in the field. Now that the data are collected, it is the processing and sample design complexities which need to be taken into account in developing the sampling weights and computing variances; the disclosure constraints limit the extent of the feasible adjustments.

Response Outcomes for the 1989 Survey

The 1989 SCF is not unlike other surveys in that both unit and item nonresponse issues must be addressed. In order to account for the item nonresponse, a multiple imputation procedure based on the stochastic relaxation and Gibbs sampling techniques that condition on all of the possible data is being used to impute the missing items (Kennickell, 1991). Concerning the unit nonresponse, the topic addressed by this paper, the response rate for the area probability sample is 69%. For the list sample, the unit response rate is 34%. Most survey practitioners would agree that a response rate of 34% casts a shadow of doubt on the representativeness of the data. The setting for the 1989 SCF allows for investigations into the differences of the respondents and nonrespondents and adjustments to be developed to minimize such doubt.

The contact procedure used for the list sample cases is unique and thus a brief description of it is warranted. Sampled cases drawn from the list frame represent tax filing units and were approached with a two-step process. Initially, selected units were sent a project description and letters requesting cooperation from the Chairman of the Federal Reserve Board and the Director of the Institute for Social Research at University of Michigan (SRC's parent organization). In addition, the selected units were sent a postage paid postcard to return if they did not want to be contacted by an interviewer. This passive consent procedure resulted in the overall response rate of 34% for the list sample. This is an improvement over the response rate of 9% attained for the list sample in the 1983 Survey of Consumer Finances (Heeringa and Curtin, 1987). For the 1983 SCF, potential respondents for the list sample were asked to return the post-card if they wanted to participate in the survey. An investigation into the nonresponse issues was undertaken for the 1983 SCF and the results presented earlier (Woodburn 1989).

For the 1989 SCF, the resulting response rates for the list sample cases decreased, as expected, by increasing wealth index strata. The definition of the wealth index strata and the corresponding response rates for the list sample are shown in Table 1.

Table 1. -- Response Rates for 1989 SCF List Sample

Wealth Index Stratum ($ 000,000)	Overall Response Rate
under 0.1	48.4%
0.1 to 0.5	43.3%
0.5 to 1	39.6%
1 to 2.5	39.4%
2.5 to 10	30.6%
10 to 250	20.1%
above 250 - censored	

General Nonresponse Issues and Adjustment Techniques

Unit nonresponse is a common concern shared by all surveys and many techniques have been developed to address it. The possible adjustments are driven by the amount of auxiliary information available for the nonrespondents, the amount known about the relationships between the respondents and nonrespondents, and the type of frame information available. For the purposes of this paper, it will be useful to think about weighting adjustments for unit non-response in terms of quasi-randomization (QR) models of the response probabilities. Oh and Scheuren (1983) formulate the QR approach which basically treats the response mechanism as the final stage in the overall probability sample design. With this approach, the practitioner makes certain assumptions to model the probability of a response given selection to the sample.

A good overview set in the QR framework of some typical weighting adjustments for unit nonresponse is provided in Kalton and Maligalig (1991). In order to depict where the 1989 SCF fits into the big picture of nonresponse adjustments, two aspects of the Kalton-Maligalig paper will be reiterated here: (1) two different types of auxiliary information and (2) the spectrum of typical assumptions made in modeling nonresponse.

Kalton and Maligalig discuss two types of auxiliary information: Nonrespondent Data and External Data. Nonrespondent Data is information about individual nonrespondents. Typically this information is obtained from the sampling frame, other records, or reports from other informants. External Data is information about the population as can be obtained from censuses, such as the population age/sex/race distribution. For the 1989 SCF, information is available on nonrespondents from the 1987 SOI Individual data file. This nonrespondent data includes detailed sample design information as well as data on other financial characteristics which are likely to be correlated with data collected on the survey. External data in the form of census totals are also available; however these data are not used in making adjustments for the list frame by itself. (These external data will be used to adjust the combined sample of the list and area frame respondents in the computation of final sampling weights -- a topic not addressed by this paper.) Another source of data is available for the 1989 SCF which is an extension of nonrespondent data, yet is not exactly external data. Data from the 1987 SOI file is available for a good representation of, although, not the entire, tax filing universe. These data are used in the post-survey weight examination and computation of adjustments in preparation of the overall sampling weight for the list file.

Kalton and Maligalig also discuss the different assumptions typically made in modeling the nonresponse. The type of adjustment made and data necessary to implement it increase in complexity from making no adjustment at all (assume an equal probability of response) to using a probit model to estimate the response probability (assume probability of response can be modeled).

Although, the use of a probit or logistic model to estimate the propensity of response is probably the least commonly used adjustment technique, recent research indicates an increased interest (e.g. Ekholm and Laaksonen, 1991). Rubin (1985) proposes the use of propensity scores in applied bayesian inferences and discusses the application of propensity scores to adjusting for unit nonresponse. Rubin states that, generally, the propensity scores are the coarsest possible summary of the information in the covariates such that given the summary, the mechanism that selects units for inclusion in the sample can be ignored. Use of the propensity score method is necessarily limited to settings where detailed information is available on the nonrespondents, for example, in panel surveys to adjust for unit nonresponse due to attrition (Little and David,1983) or partial nonresponse (Lepkowski, Kalton, and Kaspryzk, 1989). A survey setting where frame data, such as administrative data,

are available is also a good candidate for the use of propensity score methods. Czajka et. al. (1987) investigate the use of propensity scores to account for late filers in the SOI Individual study.

Results -- Are the Respondents a Representative Subsample?

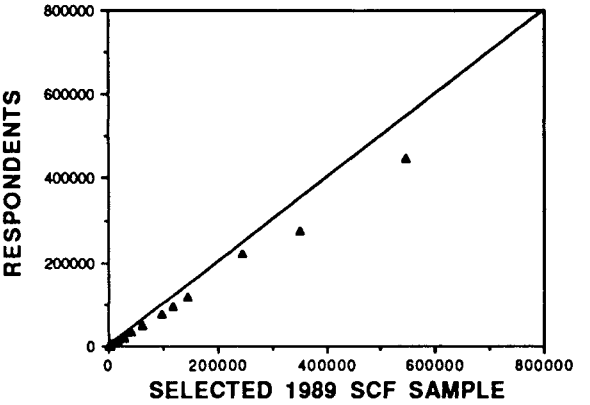
In order to use the auxiliary information to investigate the possible non-response bias and compute weight adjustments for the 1989 SCF, a special file has been made which contains only data from the SOI 1987 Individual tax file and an indicator of the response status for each case. The investigation into the possible bias due to the 34% response rate for the list sample was performed in wealth index strata to determine if the response mechanism is ignorable within sampling strata. The investigation began with a comparison of the univariate distribution of the respondents to that of the sample. Quantile-quantile (qq) plots were used in order to make the comparisons.

In a qq chart, the percentiles of one group are graphed against the percentiles of the other group. A 45 degree line represents the case where the two distributions are identical. A description of the use of qq charts can be found in Hoaglin et. al. (1985). (A normal probability plot is a qq chart where one group is assumed to have the standard normal distribution.) For the 1989 SCF, qq charts of the unweighted distributions of several auxiliary income variables from the SOI file were developed. The preliminary results are quite encouraging. In most cases, the respondent distribution is very similar to that of the entire sample. The income items investigated include adjusted gross income (AGI), gross business income, and income from interest, non-taxable interest, and dividends. The qq chart for Interest Income, one of the worst cases is shown in Figure 1.

Results -- Overall Post-Survey Weight Examination

In addition to investigating the respondent and nonrespondent characteristics as described above, it was desired to use the auxiliary information to validate the overall sampling weights computed for the list sample. In order to do this, weighted distributions using the entire 1987 SOI Individual file were computed as an estimate of the universe of tax filers. Weighted respondent distributions were also computed. Initially, uniform weights were assigned within wealth index strata. The resulting qq-plot comparison for AGI is shown in Figure 3. The uniformly weighted respondents greatly overstate the tails of the AGI distribution. Similar results were found for the other income variables investigated. These results were quite shocking given the similarity of the respondents and nonrespondents as discussed above. Two steps were taken to further investigate: (1) the sample selection process was revisited to determine if the assumption of equal probability of selection within wealth index strata was valid and (2) a probit adjustment was computed to account for the probabilities of selection into the sample as well as the response probability.

Figure 1. -- QQ Plot of 1987 Interest Income Respondents vs Selected SCF Sample Wealth Index 2.5 to 10 Million

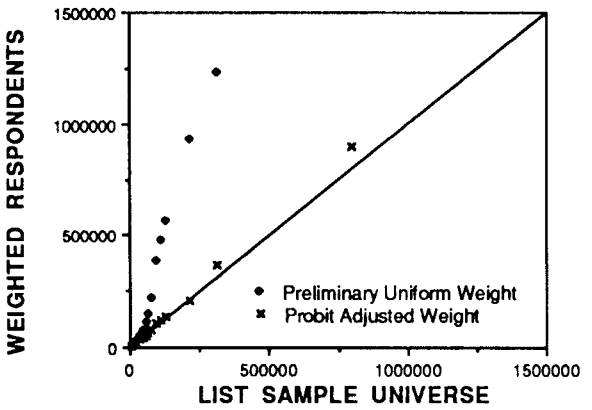


As mentioned in the second section, the list sample selection was complicated by several processing constraints. Upon reflection, these complexities and the introduction of additional pseudo-replicates into the field did invalidate the assumption of equal probability of selection within wealth index strata. Revised design weights which reflect these complexities have been computed and the resulting weighted distribution of the respondents is very similar to the base universe distribution for all of the variables considered. These results are not presented here.

The second approach to address the discrepancy was to compute model based sampling probabilities. At this point it is helpful to consider the chain of probabilities leading to

response in the 1989 SCF. Again referring to the QR framework, where the probability of response is explicitly included, the chain consists of the probabilities of: selection to the 1987 SOI Individual data file, selection to the subset of 10,000, selection to the sample released to the field, and response. A probit model was computed which reflects the later three probabilities of selection. According to the literature cited earlier, the use of a probit model to estimate the inclusion probability should remove all bias, be it design, processing, or response, for those variables included as covariates in the probit model. For the 1989 SCF, the probit model developed included every variable which could conceivably be of interest, transformations of these variables, and indicators of sampling strata. The model used the entire 125,000 in the 1987 SOI file to estimate the response probability for the 866 respondents. The probability of response predicted with the probit model was used to weight the respondents to represent the 1987 SOI Individual data file and the original SOI sampling weight completed the chain to weight up to the tax filing universe. The results using this model based weight are quite encouraging and are reflected in the qq chart in Figure 2 for 1987 AGI.

Figure 2. -- QQ Plot of 1987 AGI
Weighted Respondents vs List Sample Universe
Wealth Index .5 to 250 Million



Concluding Thoughts

The use of auxiliary information to investigate nonresponse issues, compute weight adjustments, and validate overall sampling weights as addressed here is still at a preliminary stage. Clearly our research has benefitted by using the data rich administrative records as described.

The results for the 1989 SCF presented here concerning the potential response bias are promising. In comparing the respondents to the units selected for the sample the response bias appears minimal. The use of auxiliary information to address the validity of the sampling weights proved helpful in this case, although the validation is limited, of course, to the auxiliary variables used and also depends on how well they correlate with the survey variables. The propensity score modelling approach to the computation of selection and response is promising. In our application for the 1989 SCF, not only will the model compensate for the selection process, but also for any response bias.

The main issue not addressed here is how such adjustments will affect the inferential uses of the data; this issue will be investigated. The current plan for the 1989 SCF is to compute bootstrap samples as a means to estimate the sampling and model uncertainty. It is also planned to explore the use of longitudinal data both in the comparison of the respondents to the sampled units as well as in the computation of weight adjustments. Additionally, the results of the comparisons of the respondents to the sampled units in wealth index strata will be used in preparation of the sample design for the proposed 1992 Survey of Consumer Finances in the attempts to improve the response rate for the list sample.

Acknowledgments

The presentation given at the 1991 ASA meetings in Atlanta compared the investigation described in this paper to climbing a mountain. When climbing a mountain, one does not always know where the top is, or what will be found there. However, along the way, there are many vistas where one can stop, reflect on the path already traveled, and map out plans for further climbing. The results presented in Atlanta were those of the first *Preliminary Results Scenic Overlook* for the 1989 SCF. This version reflects some progress up the hill to another

Scenic Overlook. Throughout this investigation there have been many contributors who have made the path to take more clear. Many of these are reflected formally in the reference section, however specific mention of some is warranted here. Many thanks are extended to Arthur Kennickell for his help with the probit models and his unending encouragement. Professors Don Rubin and Rod Little provided critical input along the way. Steve Heeringa provided support with the sample design issues. And finally, many thanks to Wendy Alvey and Beth Kilss for their help with the preparation of the presentation for the ASA meetings.

REFERENCES

AGRESTI, A. (1990). *Categorical Data Analysis*, John Wiley and Sons, Inc., pp. 102-104.

AVERY, R. B., ELLIEHAUSEN, G. E., and KENNICKELL, A. B. (1988). 'Measuring Wealth with Survey Data: An Evaluation of the 1983 Survey of Consumer Finances.' Review of Income and Wealth, Series 34, No. 4, pp. 339-369.

CZAJKA, J. L., HIRABAYASHI, S. M., LITTLE, R. J. A., and RUBIN, D. B. (1987). ' Evaluation of a New Procedure for Estimating Income and Tax Aggregates from Advance Data.' *Statistics of Income and Related Administrative Record Research: 1986-1987*, Internal Revenue Service, November 1987, pp. 109-135.

EKHOLM, A. and LAAKSONEN, S. (1991). 'Weighting via Response Modeling in the Finnish Household Budget Survey.' Journal of Official Statistics, 7, 3, 1991, pp.325 - 337.

FORBES (1989). 'The Four Hundred Richest People in America - The Forbes Four Hundred,' *Forbes*, October.

HEERINGA, S. G. and CURTIN, R. T., (1987). 'Household Income and Wealth: Sample Design and Estimation for the 1983 Survey of Consumer Finances,' *Statistics of Income and Related Record Research, 1986-1987*, Internal Revenue Service.

HEERINGA, S., JUSTER, F. T., and WOODBURN, R. L., (1991). 'The 1989 Survey of Consumer Finances: A Survey Design for Wealth Estimation', *Statistics of Income and Related Record Research, 1990-1991*, Internal Revenue Service (forthcoming).

HEERINGA S., and WOODBURN, R.L., (1991). 'The 1989 Surveys of Consumer Finances, Sample Design Documentation,' forthcoming.

HOAGLIN, D. C. et. al. (1985). *Exploring Data Tables, Trends, and Shape*, John Wiley and Sons, Inc., pp. 432-442.

INTERNAL REVENUE SERVICE, (1990). *Individual Income Tax Returns 1987*, Department of the Treasury, pp.13-17.

KALTON, G. and MALIGALIG, D. S. (1991). 'A Comparison of Methods of Weighting Adjustments for Nonresponse', paper presented at the 1991 Census Research Conference.

KENNICKELL, A. B. (1991). 'Imputation of the 1989 Survey of Consumer Finances: Stochastic Relaxation and Multiple Imputation,' *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

LEPKOWSKI J., KALTON, G. and KASPRZYK, D. (1989). 'Weighting Adjustments for Partial Nonresponse in the 1984 SIPP Panel.' *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 296-301.

LITTLE, R. J. A. and DAVID, M. H. (1983). 'Weighting Adjustments for Non-response in Panel Surveys.' Working Paper. Washington D.D.: U.S. Bureau of the Census.

OH, H. L. and SCHEUREN, F. J. (1983). 'Weighting Adjustments for Unit Nonresponse.' In *Incomplete Data in Sample Surveys, Volume 2: Theory and Bibliographies* (W. G. Madow, I. Olkin and D. Rubin, eds.), pp. 143 - 184. New York: Academic Press.

PROJECTOR, D. S. and WEISS, G. S. (1966). 'Survey of Financial Characteristics of Consumers', Board of Governors of the Federal Reserve System.

RUBIN, D. B. (1985). 'The Use of Propensity Scores in Applied Bayesian Inference.' *Bayesian Statistics 2*, pp 463-472.

WILSON, O. and SMITH, W. J. Jr. (1983). 'Access to Tax Records for Statistical Purposes.' *1983 Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp. 595-601.

WOODBURN, R. L. (1989). 'Using Supplementary Information to Investigate Nonresponse Bias', *1989 Proceedings of the Section on Survey Research Methods*, American Statistical Association