Journal of Official Statistics
Vol. 9. No. 1, 1993, pp. 233–243

Statistical Research in a Large-Scale Private Survey Research Organization

David Morganstein, J. Michael Brick, John Burke, David Cantor, and David Judkins[1]

Abstract: This article describes a number of recent research projects on survey issues conducted by the statistical group in a non-governmental statistical company in the United States. We begin with a brief overview of the group and its functions, in particular, the functions related to statistical research. Recent research is then presented in two categories: methodological variations in sample design; and, basic research on other aspects of survey methodology. In the final section, we discuss our efforts to improve the quality of our work using the contemporary approach to quality management.

Key words: Sample design; cognitive research; response errors.

1. Overview of Westat's Statistical Work

Westat's statistical group consists of two dozen masters and doctoral level statisticians who are involved in both Westat's survey operations and in research on statistical methods. Westat is a large survey organization that conducts surveys in many different areas mainly for the U.S. federal government. Survey topics include health, education, environment, energy, and a wide variety of social issues. Westat's statisticians participate in many aspects of survey research including: research design, sample design for an area, list and random digit dialing (RRD) surveys, instrument design, sample selection and monitoring, imputation, weighting and estimation, analysis and sampling error computations. A typical project will involve two or more statisticians, one being familiar with all aspects of the problem while a second, working in concert, can provide support at the same time he/she is learning.

A distinction can be observed between two categories of Westat survey research. Many of our surveys involve a traditional interviewing situation in which a respondent is administered a survey instrument. However, a significant number of our projects, particularly in the health and environmental areas, are direct measurement studies. Some examples of the kinds of measurements required are: medical examinations in portable clinics; engineering assessments of asbestos encapsulation; or laboratory testing of field collected samples. Both classes of surveys contain unique measurement problems which must be identified and resolved.

2. Methodological Variations on Sample Survey Designs

In the following section, we examine several variations of our basic function, designing

[1] Westat, Inc., 1650 Research Boulevard, Rockville, MD, 20850, U.S.A.

and implementing sample surveys. The first of these is an application to an increasingly larger number of case control studies in which a sample survey provides a more representative set of controls than would be obtained in a more traditional epidemiological study. The second variation addresses longitudinal surveys and our efforts to assess their efficiencies and biases in providing both measures of change and of point-in-time estimates. These are followed by two specific applications of survey redesign techniques. The first example is a redesign of the National Health Interview Survey and the second is the redesign of our own Master Sample area design based upon the 1990 census.

2.1. *Case control studies*

A recent development in the conduct of case-control studies in clinical research is the use of RDD samples to locate the control groups (Hartge et al. 1984). In traditional case control studies, financial resources generally limited the selection of controls to relatively small areas and frequently in a way that makes it impossible to define the control population. Westat has been a leader in refining the RDD approach to provide control groups that represent an entire population through use of probability samples, and in a cost effective manner.

As with any population survey, the first step in the decision process is to evaluate the telephone coverage of the control group. If telephone ownership for the group is relatively low (which essentially means that low income persons are of particular interest in the study), telephone-based sampling is inappropriate. Even if this is the case, two control groups can be considered, a telephone-based sample to be compared to those cases with telephone-ownership and a sample from households without telephones to obtain the controls for the non-telephone households. Such a dual approach may be cost-effective.

Epidemiological studies usually require that the subject sample be matched on a wide variety of variables. This matching can be accomplished more efficiently by telephone screening in several ways. A number of households can be contacted and then members enumerated. Aggregated into a single pool, the collection of household members can be sampled to obtain the desired sample sizes for each of the strata. If there is a concern that multiple contacts to the household will reduce response rates or lengthen the field period excessively, sampling rates can be applied at the time of enumeration to select respondents for immediate interview.

There are alternative methods for selecting control groups representative of the entire population. A variety of publicly available lists may be considered for such purposes. On the one hand, the list-based design is a single-stage design where the RDD approach is a two-stage design that selects households first and then respondents within the household. However, issues of coverage and timeliness can dramatically diminish the utility of the list-based approach. For these reasons, Westat has relied almost exclusively on the RDD method for the selection of population-wide control groups.

2.2. *Longitudinal studies*

Several Westat surveys involve rotation designs in which a sample of respondents is recontacted for interview at the same time a fresh sample of respondents is added to the existing pool of those who have been previously interviewed. The primary reason for choosing such designs is to efficiently assess change while retaining the ability to

make precise cross-sectional estimates. However, these designs also enable the assessment of panel biases by comparing the responses of the two groups. Response rates can also be compared to determine whether repeated contacts influence the likelihood of participation.

One survey using this design is the Youth Attitude Tracking Survey (YATS) conducted for the Defense Manpower Data Center. In this annual survey, men and women between the ages of 16 and 24 are interviewed as to their interest in joining the military. After an initial base-line survey conducted in the late fall of 1990, a summer 1991 supplement was selected. The supplement included both respondents from the initial survey as well as a fresh sample. In the late fall of 1991, a second annual survey was conducted. It contained a sample of respondents from both prior surveys, and again, a fresh sample for comparison. Comparison of the response rates indicated no important differences in willingness to participate in either a second or third interview. Comparison of respondents' propensity to enlist, however, did suggest some differences. These initially observed differences will be further analyzed with the aim of adjusting for possible panel differences.

Another series of repeated RDD survey series is being conducted for the State of California to assess changes in smoking prevalence in light of state mandated programs to inform the public of the dangers of smoking. This series also uses a rotation design in which respondents aged 16 and over are contacted to assess their knowledge, attitudes and behavior regarding smoking. The fresh sample of respondents is selected both within the same telephone clusters used in previous samples to design some positive correlation between estimates at two points in time, as well as, from an independent sample of clusters to avoid possible biases. As with the YATS, the design involves re-interviews of a sample of those previously contacted along with a fresh sample of respondents. Questions under study include the effect of re-interview upon responses and upon response rates.

2.3. *National health interview survey redesign*

The National Health Interview Survey (NHIS) carried out by the U.S. Bureau of the Census for the National Center of Health Statistics (NCHS) is one of the major data collection programs of the NCHS. The survey measures health characteristics of the civilian noninstitutionalized population of the United States. Measurements are taken by personal interviews of a probability sample of households. The NHIS sample is also used to integrate the designs for other sample surveys conducted at NCHS which use the NHIS Primary Sampling Units (PSUs).

The NHIS is redesigned each decade to coincide with new population information collected in the decennial census. At these times, the data requirements of the survey are re-assessed, and design changes are made accordingly. New requirements were established for 1995 specifying that the new design should provide reasonably precise estimates for a set of demographic and geographic subdomains. The design is also being revamped to include updated population information, new cost-saving features, and new and more efficient survey design techniques.

Under contract to NCHS, Westat is conducting research in support of the 1995 redesign. A number of specific research questions are under exploration. For example, NCHS wants to know whether the NHIS can be redesigned to provide adequate statistics for demographic subdomains defined

in terms of race/ethnicity, age, and sex and if so, how. They also want to know whether, at the same time, the NHIS can be redesigned to provide adequate state-level statistics. Such a design needs to be examined under a number of funding assumptions.

To answer these questions, Westat's research is following several lines. NHIS variances are being decomposed by stage of sampling to better understand contributions from different sampling units. Simultaneously, cost models are being developed to be used in conjunction with the variance components to identify optimal design alternatives. Multiple-frame and network sampling applications are being considered as alternative design choices. Westat is studying the effects of oversampling selected subdomains or states at each stage of sampling as well as the effect of stratification of PSUs within states.

2.4. *Master area sample redesign*

Upon the release of decennial census population counts, Westat selects a national probability sample of county clusters which serves as the first-stage sample for a variety of multistage area samples over the ensuing decade. Restricting these samples to a fixed set of areas enables our survey operations staff to recruit and train quality interviewers in each location and to call upon them repeatedly as the need arises. In a typical study there are additional stages of selection involving blocks or block groups, dwelling units, and finally individual respondents within households. An objective of the Master Sample is to build a first-stage sample that can be applied flexibly to a number of surveys throughout the decade. It is likely, for example, that some of the requirements for area samples will need to oversample subpopulations such as minorities. Below, we describe how we designed the Master Sample to meet these often divergent objectives.

The current Westat Master Sample consists of three nested samples of 100, 81, and 62 geographically defined primary sampling units, each sample representing the 50 states and the District of Columbia. Budget and precision requirements determine which of the three samples would be used for a particular study. In constructing the sampling frame, each Metropolitan Statistical Area (MSA) constituted a PSU. Outside MSA boundaries, a PSU was one or more contiguous counties with an aggregate 1990 population greater than 15,000. The 24 largest MSAs which contain roughly 30% of the U.S. population were designated as certainty selections for all three of the samples. The remaining PSUs were grouped into 38 strata of roughly equal size. The main stratification variables were region and MSA status.

A specific goal of the 1990 Master Sample design was to construct strata containing PSUs with high concentration of blacks or Hispanics to facilitate the oversampling of these subpopulations for studies that required it. Three urban strata and two rural strata with high concentrations of blacks were identified in the South. Three urban strata with high concentrations of Hispanics were constructed, one in the South and two in the West. The remaining stratification was based on 1988 per capita income (PCI) for non-MSA PSUs and PCI within population size class for MSA PSUs.

The 100 PSU sample consists of the 24 certainty selections and 76 noncertainties selected two per stratum using the Durbin method with probability proportional to 1990 population. The 62 PSU sample retains the 24 certainty selections and subsamples with equal probability, one noncertainty PSU of the two selected from each stratum. The 81 PSU sample is slightly more

complicated. The 24 certainty PSUs are retained as usual. Noncertainty strata are paired to form 19 superstrata. For each stratum pair, one stratum retains both noncertainty PSU selections and the other stratum retains only the PSU from the 62 PSU sample. Which stratum pair retains both PSUs was determined at random with a specified probability. The retention probabilities were chosen so that the three selections from each noncertainty superstratum have individual probabilities of selection equivalent to PPS selection from the superstratum. Standard errors will be computed using replicated estimation techniques.

3. Research Topics

In this section we describe four research topics under examination by Westat staff. Several of these areas explore sources of nonsampling error. The first topic examines the effect of cognitive research in instrument design in an effort to find ways of reducing the nonsampling component of total survey error. Next, we describe recent research into the relation between the length of the recall period and the accuracy of response. During the past decade, the use of the telephone as a data collection mode has increased substantially, and Westat has been an active participant in exploring this modality and finding efficient ways of using it. Several fertile research areas are described next. We end this section with a brief summary of work conducted on the Medicare Beneficiary Survey in which both response and nonresponse error are being quantified.

3.1. Role of cognitive research in instrument design

Under a task order agreement with the Bureau of Labor Statistics (BLS), Westat has conducted research associated with redesigns for surveys administered by the BLS and the Census Bureau. This research has involved extensive use of both traditional survey research methods and those drawn from the field of cognitive science.

For BLS, this research has focused on both household and establishment surveys. As part of the redesign for the Current Population Survey (CPS), these activities included recommending revisions to the questions on the "hours worked" (Edwards, Levine, and Cohany 1989) and "Industry/Occupation" questions (Gaertner, Cantor, Gay, and Shank 1989; Cantor and Esposito 1992). These recommendations led to significant changes in both of these items on the new design of the CPS. For the Consumer Expenditure Survey (CES), this research has focused on several aspects of the diary used to collect information on small or frequent expenditures. This included research on the development of a diary to collect expenditures on clothing and related apparel (Cantor, Keil, Greenless, Rose, and Bagin 1990) and detailed analysis of the decision making process respondents use when classifying food expenditures (Leven et al. 1991). For the Occupational Safety and Health (OSH) survey, a set of debriefing interviews were conducted to evaluate a new survey form designed to collect incident level information (Edwards, Levine, and Cohany 1989). This research also led to a preliminary response model adapted to establishment surveys (Edwards and Cantor 1992).

For the Census Bureau, this research has focused on two questionnaire design issues for the Survey of Income and Program Participation (SIPP). The first issue was related to the portion of the questionnaire that collects recipiency and amount information. The research involved a set of think-aloud interviews which were used to pinpoint major sources of measurement error. It led to a set of recommendations to redesign the

survey from structured questions to a set of open-ended items (Cantor, Brandt, Green, Moesinger and Rose 1991; Cantor, Brandt, Green, and Moesinger 1992). The second issue was related to improving coverage of the household population on the SIPP. It involved conducting a small field experiment which tested an alternative set of household roster questions that did not depend on ambiguous concepts such as "usual residence" (Cantor and Edwards 1992). It led to a recommendation for further research on the use of more straightforward roster questions that may not lead to substantial undercounts of minority populations.

3.2. *Recall error*

Several Westat staff have contributed to knowledge of respondent recall of events (Neter and Waksberg 1964; Waksberg and Valliant 1978; Eisenhower, Mathiowetz, and Morganstein 1991). In a recent experiment conducted for the U.S. Department of the Interior (DOI), fishing and hunting enthusiasts were studied to determine the optimal recall period for their recreational experiences (Chu et al. 1992).

Westat designed and conducted an experiment involving recall periods varying from two weeks to one year in length. Anglers and hunters residing in sites in two states were randomly assigned to recall length panels and interviewed over a one-year period. The survey instrument collected information on the frequency, length and cost of hunting and fishing trips. The responses from the various panels were compared under the assumption that the shortest recall period would provide the most accurate responses.

In addition to estimating the probable degree of recall error, Westat developed a variance model for the periodic DOI survey design. This variance model and the recall error estimates pointed to an optimal recall period. The optimal was selected as that recall period having smallest expected mean squared error. A period of four to six months was recommended, depending upon the choice of statistic.

3.3. *Random digit dialing*

The use of the telephone as a mode of data collection has increased dramatically over the last 10 to 15 years in the United States. Many surveys use the telephone as a means of primary data collection, and a very large proportion of those surveys using mail or in-person modes make use of the telephone for nonresponse follow-up or to resolve data collection problems. The lower cost for conducting interviews by telephone has been the major reason for this increased use of telephones.

With the increased demand for telephone surveys, Westat has been heavily involved in methodological research associated with this mode of data collection. One area of research is the efficiency of selecting a random sample of households with telephones, referred to as random digit dialing (RDD). A closely related area of research is the targeting of samples to achieve a higher proportion of the sampled households and persons with particular characteristics. A third area of methodological research involves efforts to reduce nonsampling errors in telephone surveys, in particular, errors arising from nonresponse and undercoverage. These three research areas are briefly discussed below to illustrate the practical problems which influence the research agenda at Westat.

The development of a sound basis for efficient RDD sampling was first presented by Waksberg (1978). The sampling method, called the Mitofsky–Waksberg procedure

after its two originators, is an efficient two-stage method of producing a self-weighting probability sample of telephone households. This method is the standard method of RDD sampling that alternative procedures are measured against.

The application of the Mitofsky–Waksberg procedure in different RDD surveys conducted by Westat revealed an awkward operational feature. The method requires a large number of callbacks to telephone numbers that repeatedly ring with no answer. Since these cases must be resolved before other numbers are released, this feature makes this procedure more difficult to implement, especially in surveys with tight time schedules.

Brick and Waksberg (1991) presented an alternative sampling procedure that eliminates the need to resolve these cases sequentially. This procedure is easier to implement than the standard Mitofsky–Waksberg procedure. However, it results in samples that are not self-weighting. Associated with the differential probabilities of selection, this procedure also results in slight increases in the variances of the estimates. The paper compared the two methods of sampling empirically and theoretically and presented the advantages and disadvantages of each.

A second area of telephone survey methodological research was prompted by the need to oversample rare segments of the population, most frequently black and Hispanic households. A common survey objective is to have a large enough sample size to make reliable comparisons among the black, white, and Hispanic populations of the United States. The use of a rejective screening method alone to obtain larger minority sample sizes can be very expensive, especially since a large fraction of the data collection cost is spent in the initial contact with the household.

Mohadjer (1988) investigated the use of a commercial list containing census characteristics for prefix (the area code and first three digits of the telephone number) areas as auxiliary data for sampling. She evaluated the quality of the data on the list by comparing the results empirically to a telephone survey and suggested procedures for using this data source. Subsequent surveys conducted by Westat have used this methodology and found it to be a very effective means of oversampling black and Hispanics compared to alternative methods.

A third area of continuing research in telephone surveys at Westat is methods to reduce nonsampling errors associated with undercoverage and nonresponse. Undercoverage occurs because some households do not have telephones. For general purpose surveys, the undercoverage is often not a serious problem since only slightly over 5% of persons reside in households without telephones and this does not result in large biases for estimates of the general population. However, in surveys targeted at groups that have much higher undercoverage rates, such as high school dropouts or persons living in poverty, the biases can be very substantial.

Various approaches to the problem of undercoverage have been studied at Westat. The most common approach has been to try to improve the post-survey adjustments that are made to reduce the undercoverage bias. Several investigations of the bias after adjustments have been conducted. One study used network or multiplicity sampling to improve the coverage of the population. In that survey, interviews were conducted with mothers of youths aged 14 to 21 years old. The youths were also interviewed. Without multiplicity methods, youths without telephones who no longer live with their mothers had no chance to be selected.

The other major source of nonsampling error in telephone surveys is nonresponse.

The study of nonresponse and methods to reduce it have also been investigated from many perspectives. Important topics that have been studied include scheduling of interviews, training methods, interviewer characteristics, and nonresponse follow-up strategies. Although we have not noticed an increase in nonresponse rates for Westat-conducted telephone surveys over time, we think that the introduction of new telephone technology could affect the future of telephone survey research.

3.4. *Medicare current beneficiary survey*

The Medicare Current Beneficiary Survey (MCBS) is a longitudinal sample of 12,000 elderly and disabled beneficiaries of the Medicare program. Some of the interesting statistical issues in this survey deal with nonresponse and response error. Sample persons are visited every four months to collect data about all of their medical care since the last visit. There is a heavy emphasis on the financing of that medical care and the interest in reducing errors in the reporting of these financial data is great.

As in other panel surveys, there are three types of nonresponse which must be handled in MCBS: complete, wave, and item (Kalton 1986; Lepkowski 1989). For complete nonresponse, we have fairly extensive administrative data to use in adjustment. We plan to use logistic regression modeling of nonresponse propensity to develop weighting adjustments for complete nonresponse. For wave nonresponse, we plan to use techniques similar to those on other panel surveys such as the Panel Survey of Income Dynamics and SIPP (Lepkowski, Kalton and Kasprzyk 1989). For item nonresponse, we are considering model-based imputation procedures with stochastic components, independent variables from both the same and prior waves, and multiple replications for the measurement of imputation variance. The administrative records on health care services partially financed by the government, "covered services," again provide unique opportunities for evaluation of nonresponse adjustments.

We will be able to measure the extent to which covered services are underreported by beneficiaries. Since we will be able to correct that underreporting, there may be a desire to correct in some way the assumed parallel underreporting of uncovered services. (If there is no adjustment of the uncovered services, there will obviously be a bias in the percentage of health care costs borne by the government.) At this point, the techniques for such an adjustment constitute an entirely open research area.

The well-known "seam effect" is also expected since three or four nonoverlapping recall periods will be used to cover a calendar year (Bailar 1989; Kalton and Miller 1991). However, it is not expected to pose a severe problem since the emphasis in MCBS is on annual patterns of health care consumption and financing rather than on transitions between various states.

4. The Application of Total Quality Management (TQM) Within a Statistics Department

Traditionally, Westat has valued the objective of continuous quality improvement (Morganstein and Hansen 1990). Westat is improving the quality of its statistical group principally in three ways: employee development; documentation of standard procedures; and improved technology. By employee development we mean formal and informal opportunities for statistical staff to gain breadth and depth of statistical knowledge. Our group provides informal training sessions as well as frequent lunch-time discussions of technical issues under explora-

tion within the department. Westat also provides financial support to staff in their efforts to upgrade their formal training at local universities as well as to attend or present papers at professional meetings. In this section we will elaborate on our efforts in documentation and improved technology.

The development of standard procedures is fundamental to Westat's ability to meet its clients' needs. We are required to complete a large number of varied projects, often in short time periods, each with their own unique characteristics. While we have only begun the process of documenting preferred procedures, we recognize that the resources that must be invested to accomplish this will reap significant rewards. We currently have five groups documenting procedures within the statistics group. Currently teams are addressing the following topics: single stage list-sampling, random digit dialing sample design, imputation, documentation, and weighting procedures. The teams consist of both very experienced and less experienced staff members, each bringing different strengths to the endeavor. The teams prepare either a checklist or a flow chart accompanied by detailed text on the way in which specific statistical functions are to be completed. The detailed text also provides some background on the task, placing it in perspective, and demonstrating how it serves the overall objective. Finally, these documents are reviewed and discussed among all staff members to arrive at an agreement on the approaches to be followed.

One aspect of improved quality is the use of better technology. Westat's statisticians are constantly seeking to develop and to implement novel and efficient approaches to survey designs. Drawing upon the experience of its most senior staff, new techniques for sampling and estimation have been developed and reported in professional journals and at professional meetings. In addition, automated software has been prepared which increases efficiency at the same time it reduces errors. Our software for variance estimation (WESVAR, WESREG and WESLOG) is routinely provided to federal clients to facilitate the computation of sampling errors so that the quality of federal reports may be enhanced through the inclusions of measures of precision along with the estimates themselves (Flyer, Rust, and Morganstein 1989). We also developed and use general purpose software for the selection of samples (WESSAMP). More recently, we have completed routines which perform the myriad of steps required for weighting (WESWGT) and for hot-deck imputation (WESDECK). Each of these routines is under continual scrutiny to identify opportunities for improvement.

5. References

Bailar, B.A. (1989). Information Needs, Surveys and Measurement Errors. In Panel Surveys, eds. Kasprzyk, D., Duncan, G., Kalton, G., and Singh, M.P. New York: John Wiley and Sons, 1–22.

Brick, J.M. and Waksberg, J. (1991). Avoiding Sequential Sampling with Random Digit Dialing. Survey Methodology, 17, 27–41.

Cantor, D. and Edwards, C. (1992). Testing an Alternative Household Rostering Procedure for the Survey of Income and Program Participation. Final Report submitted to the Census Bureau under BLS contract #J-9-J-8-0083.

Cantor, D. and Esposito, J. (1992). Evaluating Interviewer Style for Collecting Industry and Occupation Information. Proceedings of the Section on Survey Research Methods, American Statistical Association (in press).

Cantor, D., Brandt, S., Green, J., and Moesinger, K. (1992). Results from Wave 2. Memo submitted for partial fulfillment of contract with the Census Bureau under BLS contract #J-9-J-8-0083.

Cantor, D., Brandt, S., Green, J., Moesinger, K., and Rose, K. (1991). Results from First Wave Interviews. Memo submitted for partial fulfillment of contract with the Census Bureau under BLS contract #J-9-J-8-0083.

Cantor, D., Keil, L., Greenlees, J., Rose, A., and Bagin, C.B. (1990). Cognitive Research to Design an Apparel Diary. Paper presented at the annual meeting of the American Statistical Association.

Chu, A., Eisenhower, D., Hay, M., Morganstein, D., Neter, J., and Waksberg, J. (1992). Measuring Recall Error in Self-Reported Fishing and Hunting Activities. Journal of Official Statistics, 8, 19–39.

Eisenhower, D., Mathiowetz, N.A., and Morganstein, D. (1991). Recall Errors: Sources and Bias Reduction. In Measurement Errors in Surveys, eds. Biemer, P., Groves, R., Lyberg, L., Mathiowetz, N., and Sudman, S., New York: Wiley and Sons, 127–144.

Edwards, W.S., Levine, R., and Cohany, S. (1989). Procedures for Validating Reports of Hours Worked and for Classifying Discrepancies Between Questionnaire Reports and Validation Totals. Proceedings of the Section on Survey Research Methods, American Statistical Association, 496–501.

Edwards, W.S. and Cantor, D. (1992). Toward a Response Model in Establishment Surveys. In Measurement Errors in Surveys, eds. Biemer, P., Groves, R., Lyberg, L., Mathiowetz, N., Sudman, S., New York: John Wiley and Sons, 211–233.

Flyer, P., Rust, K., and Morganstein, D. (1989). Complex Survey Variance Estimation and Contingency Table Analysis Using Replication. Proceedings of the Section on Survey Research Methods, American Statistical Association, 110–119.

Gaertner, G., Cantor, D., Gay, N., and Shank, S. (1989). Research on Industry and Occupation Questions in the Current Population Survey. Proceedings of the Section on Survey Research Methods, American Statistical Association, 490–495.

Hartge, P., Brinton, L.A., Rosenthal, J.F., Cahill, J.I., Hoover, R.N., and Waksberg, J. (1984). Random Digit Dialing in Selecting a Population-Based Control Group. American Journal of Epidemiology, 120, 825–833.

Kalton, G. (1986). Handling Wave Nonresponse in Panel Surveys. Journal of Official Statistics, 2, 303–314.

Kalton, G. and Miller, M.E. (1991). The Seam Effect with Social Security Income in the Survey of Income and Program Participation. Journal of Official Statistics, 7, 235–245.

Lepkowski, J. (1989). The Treatment of Wave Nonresponse in Panel Surveys. In Panel Surveys, eds. Kasprzyk, D., Duncan, G., Kalton, G., and Singh, M.P., New York: Wiley and Sons, 348–374.

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.

Levin, K., Cantor, D., Lichtenstein, C., Brandt, S., Hosang, A., and Stunkard, J. (1991). Classifying Food Purchases on the Diary for the Consumer Expenditure Survey. Final report in partial fulfillment of contract with the Bureau of Labor statistics, contract #J-9-J-8-0083.

Mohadjer, L. (1988). Stratification of Prefix

Areas for Sampling Rare Populations. In Telephone Survey Methodology, eds. Groves, R., Biemer, P., Lyberg, L., Massey, J., Nicholls, W., and Waksberg, J., New York: John Wiley and Sons, 161–173.

Morganstein, D.R. and Hansen, M. (1990). Survey Operations Processes: The Key to Quality Improvement. In Data Quality Control, eds. Liepins, G.E. and Uppuluri, V.R.R., New York: Marcel Dekker, Inc. 91–104.

Neter, J. and Waksberg, J. (1964). A Study of Response Errors in Expenditures Data from Household Interviews. Journal of the American Statistical Association, 59, 18–55.

Waksberg, J. and Valliant, R. (1978). Final Report on the Evaluation and Calibration of NEISS. Westat, Inc. for Consumer Products Safety Commission.

Waksberg, J. (1978). Sampling Methods for Random Digit Dialing. Journal of the American Statistical Association, 73, 40–46.

Received April 1992
Revised July 1992