

# Survey Methods Research at the U.S. Bureau of Labor Statistics

*Cathryn S. Dippo, Clyde Tucker, and Richard Valliant<sup>1</sup>*

**Abstract:** Behavioral scientists and mathematical statisticians at the U.S. Bureau of Labor Statistics conduct survey methods research projects across a broad range of topics. In this article, we summarize recent research projects and plans in the areas of

sample design, data collection methodology, and data adjustment and analysis.

**Key words:** Sample design; data collection methodology; cognitive methods; questionnaire design; estimation.

## 1. Introduction

In the five years since “A Review of Statistical Research at the U.S. Bureau of Labor Statistics” appeared in the *Journal of Official Statistics* (Dippo 1987), the Bureau has broadened its commitment to improving its survey programs by establishing an Office of Survey Methods Research, which encompasses research in both mathematical statistics and the behavioral sciences. Behavioral scientists at the Bureau of Labor Statistics (BLS) include cognitive, educational, and social psychologists, sociologists, and an anthropologist. The addition of behavioral science researchers to the Bureau’s staff was prompted by the recognition that causes of nonsampling errors could be studied by several means. For example, on-going research attempts to map mental/cognitive processes employed by respondents and interviewers involved in completing and administering data collection instruments and procedures. For a sum-

mary of research conducted by behavioral scientists at BLS, see Dippo and Herrmann (1991) and Dippo and Norwood (1992).

In this article, the variety of survey research methods at BLS is highlighted. This research is conducted within four statistical units – the Office of Survey Methods Research and three Statistical Methods Divisions located within the Offices of Compensation and Working Conditions, Employment and Unemployment Statistics, and Prices and Living Conditions. Each of these units contains both behavioral scientists and mathematical statisticians, many of whom work on interdisciplinary teams associated with specific programs, (e.g., the Current Population Survey), or cross-programmatic issues, (e.g., nonresponse). The article is organized into sections corresponding to selected steps of survey design.

## 2. Sample Design

Many of our current survey designs place a heavy burden on respondents through long interviews. In the Point of Purchase Survey, respondents are currently asked about

<sup>1</sup> U.S. Bureau of Labor Statistics, 2 Massachusetts Avenue NE, Room 4915, Washington D.C. 20212, U.S.A.

purchases in almost 170 categories covering intervals varying from "last week" to "the last five years." The in-household interview averages about 75 minutes. In order to reduce the burden on sampled households and, at the same time, improve sample design efficiency within the same cost constraints, research has been conducted on converting the survey to telephone data collection using a random-digit dialing sample design.

Casady and Lepkowski (1991) developed a design for this survey which uses information from a commercial list of banks of ten consecutive numbers containing listed residential numbers to stratify the universe of ten-banks into a "high-density" stratum with one or more listed numbers and a "low density" stratum of all other banks. Tucker, Casady, and Lepkowski (1992) refined this design by substratifying the low density stratum and estimating the proportions of residential numbers in the substrata. They have developed various classes of designs which will optimize the allocation of the sample across the high and low density strata based on a researcher's needs. These designs are at least as efficient as those used in the past, and, in some cases, substantially more efficient. More general design and estimation issues were studied by Casady (1989).

In establishment surveys, a recurring design problem relates to defining the sampling unit versus the collection unit. The Bureau has recently created a new sampling frame, called the Business Establishment List. A multiple worksite report is used to break down large multi-unit establishments into individual worksite establishments. Such establishments are thought to provide better quality data, yet may be more expensive to sample and collect data. Principe and Sommers (1992) investigated alternative sampling procedures and found that, under

reasonable cost assumptions, the sampling of clusters of worksites belonging to a company results in a more efficient design.

### **3. Data Collection Methodology**

#### *3.1. Theoretical developments*

Several projects undertaken recently were designed to advance the conceptual understanding of the data collection process. This work is interdisciplinary, including perspectives from several of the social sciences. Esposito and Jobe (1991) proposed a general model of the survey interaction process to be used to structure and communicate research ideas and findings from a variety of disciplines. Their framework consists of three components: the survey context, the survey participants, and a descriptive model of the interviewer-respondent interaction process.

Mullin, Conrad, Sander, and Herrmann (1992) advanced a cognitive theory called the "information exchange theory of the survey interview." This theory goes beyond previous cognitive work in this area by simultaneously considering both the interviewer's and the respondent's behavior as well as their interaction. The authors specify the ways in which different kinds of cognitive processes are sequenced in order to produce the behavior of interest.

Many surveys require respondents to recall autobiographical information for varying reference periods. Dippo (1989) discusses the use of cognitive laboratory techniques for investigating memory retrieval errors and gives examples from several surveys. The use of cognitive techniques in establishment surveys is the subject of Phipps (1990).

As many of the BLS surveys ask respondents to report information, such as income, which could be considered sensitive or personal, some basic research into the concept

of privacy and confidentiality has been undertaken. One study considered the semantic basis of our notions of confidentiality (van Melis-Wright, Stone, and Herrmann 1990). Another (van Melis-Wright 1991) looked more specifically at social conventions for inquiring about personal topics.

### 3.2. *Questionnaire design*

As reported in 1987, in order to reduce the burden in the Point of Purchase Survey and facilitate the transition into computer-assisted telephone interviewing (CATI), the survey instrument was divided into two sets of questionnaires with either 20 or 40 categories. These questionnaires were field tested in 1988, and the results indicated that the shortened questionnaires could be successfully administered by telephone. In fact, the 20-category questionnaires produced 20% to 30% more respondent reports than the lengthy paper and pencil version (Tucker, Casady, and Lepkowski 1991). Given this enhanced productivity and the lower interviewing costs with CATI, the increase in sample size necessary to provide the same information for all 170 categories with the shortened questionnaires does not result in greater costs.

The extensive field test of new formats for the Consumer Expenditure Diary Survey described in 1987 demonstrated that more extensive cues increase expenditure reporting on the part of respondents (Tucker and Bennett 1988). In order to more fully understand the problems respondents might have with different formats, an experiment was conducted in the BLS Collection Procedures Research Laboratory. Subjects were asked to classify and record information about a specially constructed set of grocery items using one of two experimental formats. Accuracy rates were calculated and used,

along with think-aloud techniques and respondent debriefings, to further refine the diary formats (Tucker, Vitrano, Miller, and Doddy 1989). Based on the results of a final field test of two revised diary formats in 1991, a new diary instrument have been selected.

Tucker (1992) studied expenditure diaries by looking at the effects of different diary procedures on nonsampling errors. Measures of errors due to both response and nonresponse were created from information contained within the survey itself. Analysis of these measures indicated important interactions between the different diaries and respondent characteristics.

In addition to diary research, changes have been made to the Consumer Expenditure Interview Survey during the past few years. Among these changes were those designed to decrease nonsampling errors in reports of medical expenditures resulting from problems of comprehension and recall. Qualitative laboratory techniques as well as field research were used to assess the effects of the changes. The results suggested that the changes were successful, providing insight into the general nature of nonsampling errors (Miller and Downes-LeGuin 1989, 1990).

Other studies have attempted to estimate the extent of response error in the interview survey. Telescoping effects were researched using statistical comparisons of selected data from the first (unbounded) wave and subsequent waves of the 1984 Consumer Expenditure Interview Survey. Silberstein (1990) estimated that telescoping inflates apparel estimates by 18% and home furnishing estimates by 37% under certain assumptions about potential conditioning effects. These results underscore the need to use bounding methods in panel surveys.

Silberstein and Scott (1991) reviewed the type of errors prevalent in the data collection of household expenditures, both diary

and recall interview. The study provided estimates of diary day and week effects. Diary completion and the use of recall were analyzed as well and further studied in Silberstein (1991). Overall reporting levels also were investigated by Silberstein and Scott. Typical assumptions about data collection modes were generally supported. Small, routine purchases are captured better in a diary, but a recall interview provides better estimates of less frequent or more salient expenses.

The development of a new instrument for the Current Population Survey has been an integrated, multi-disciplinary project involving teams including statisticians, economists, and other social scientists. During the first phase, laboratory studies were conducted to investigate proxy reporting (Boehm 1989), interviewer coding of responses to open-ended questions (Fracasso 1989), and alternative questions for collecting hours worked as well as industry and occupation data (Bienias 1988; Edwards, Levine, and Cohany 1989; Gaertner, Cantor, Shank and Gay 1989). In addition, both laboratory and field studies in the form of respondent debriefings were conducted to investigate respondents' understanding of key terms such as business, work, and looking for work (Palmisano 1990; Campanelli, Martin, and Creighton 1989). Using data from these studies, two alternative questionnaires were developed and tested using a random-digit dialing (RDD) sample of 71,899 individuals over seven months (Rothgeb, Polivka, Creighton, and Cohany 1991). Responses to a field-based respondent debriefing questionnaire, interviewer debriefings, item-based response analysis, and systematically coded interviewer-respondent interactions were used to determine which questions resulted in more accurate data (Esposito, Campanelli, Rothgeb, and Polivka 1991). A revised question-

naire was developed and tested during a second CATI RDD test. In addition to these field-based tests, laboratory work continued with studies of interviewer style on collecting industry and occupation information (Cantor and Esposito 1992).

The methods used to develop the CPS questionnaire are also generally applicable to establishment surveys. Phipps, Butani, and Chun (1992) discuss the issues of respondent burden, use of records, and terminology as they relate to establishment surveys and major methods used to develop questionnaires for new programs or one-time special surveys to be conducted by mail, i.e., focus groups, document design analyses, pretests, and response analysis surveys. Focus groups are primarily useful in assessing survey concepts and indicators, definitions, instructions, and the availability of records; document design analyses in improving comprehension and readability of the questionnaire; pretestings, as dress rehearsals, in determining comprehension difficulties and records availability; and response-analysis surveys in assessing comprehension, records use, and respondent burden.

Some studies have been conducted to investigate the effects of physical aspects of the instrument design. For example, Phipps, Robertson, and Keel (1991) looked at the effects of color on mail survey response rates from establishments of different size. Small establishments were somewhat more likely to respond to green questionnaires than white ones. No differences were found for larger establishments receiving either yellow or white forms.

### *3.3. Data collection procedures*

Many BLS surveys which involve field representatives collecting data from business establishments do not use a structured, scripted instrument; instead, the field

economists are trained in the details of the occupational definitions of the jobs for which they are to collect wage and benefit data or the product definitions of the items for which they are to collect price data. The data collectors are given forms which indicate the types or level of detail of the data they are to collect, but not how to collect the data. For example, in surveys designed to collect wage and benefits information, field economists (the interviewers) generate their own questions and interpret responses from organizational representatives without the benefit of a structured instrument as part of the process of classifying employees of sampled organizations into BLS job categories.

The Bureau of Labor Statistics staff have recently begun to use cognitive research to explore the way this job-matching task is performed. One study attempts to depict field economists' mental representations of the job categories with which they work. Novice, intermediate and experienced field economists will rate various jobs for their similarity to one another. The results will be analyzed using multidimensional scaling to determine the way field economists organize knowledge about various occupations and whether the organizing principles change with greater experience. Another related study will investigate error associated with the use of particular job definitions. Accuracy rates will be subjected to factor analysis to find patterns across the different job definitions.

In household surveys, one respondent is often asked to report about all household members. Although there have been studies investigating the quality and reliability of proxy responses in the past, there have not been studies investigating what information related to survey topics is transferred among household members and how that information is acquired. As a first step to under-

standing knowledge transfer about autobiographical information among family members, a laboratory study was conducted with 70 families. In the study, two to five members of a household were interviewed simultaneously on individual computer terminals. Each household member reported employment and expenditure information for themselves and one or two other members of the household. Along with the proxy report for another household member, respondents answered questions concerning how they learned about that person's employment or expenditures. These data are now being analyzed to determine the accuracy of proxy knowledge, the most knowledgeable household members, the various types of communication patterns, and the important family and individual-level variables that affect the level of knowledge and communication patterns.

Information on self-proxy reporting practices of labor force characteristics in Hispanic households was obtained as part of an ethnographic field study of household roster reporting patterns (McKay 1991, 1992a). The study indicated long-term boarders are often not reported as members of the household. Further research using data from the U.S. Bureau of the Census' Alternative Enumeration project confirmed this finding (McKay 1992b).

Standardized surveys require interviewers to present questions to respondents verbatim. Non-verbatim presentation is generally considered undesirable because it may lead respondents to misinterpret the questions. Only limited research, however, has been conducted in this area. A recent analysis of tape-recorded interviews of the Consumer Expenditure Interview Survey (Miller, Herrmann, and Puskar 1992) revealed that the accuracy of question reading was related to the adequacy of respondents' answers.

### 3.4. *Mode of data collection*

Much of the research on data collection over the last few years has been focused on the incorporation of computer-assisted methods. In 1994, all data collection for the Current Population Survey will be through the use of a computer, either by field staff using notebook computers or from a centralized CATI facility. By 1994, the Current Employment Survey will use a fully integrated multi-mode collection system of mail, CATI, and touch-tone data entry. And the next revision of the Consumer Price Index in 1998 will be based upon collection of data by field staff using pen-based computers.

Data in many of BLS's periodic surveys must be collected with extremely restricted time schedules. One such survey is the Current Employment Statistics (CES) program in which employment counts are collected monthly from 380,000 establishments. BLS has conducted extensive research on computer-assisted telephone interviewing and computer assisted self interviewing (CASI) using touch-tone data entry and voice recognition self-response. Werking, Tupek, and Clayton (1988), Werking and Clayton (1990), Clayton and Winter (1990), Clayton and Harrell (1989), and Clayton, Rosen, and Winter (1989) discuss the findings of various aspects of a seven-year research program which led to large-scale implementation of CATI and CASI in the CES.

Phipps and Tupek (1991) examined the measurement errors associated with a touch-tone data entry system. Three sources of data were used to estimate the errors arising from certain aspects of the human-machine interface. The effects of instrument design on these errors were also addressed.

The success of telephone data entry for CES has prompted the prototyping of a CPS instrument for telephone self-response (McKay, Malik, and Tupek 1992). Such an instrument could be used to increase the effective sample size of CPS in order to provide more reliable state estimates of persons by labor force status on a monthly basis and more cost-effectively than the current collection methods. (Not all the information currently collected would be included in the abbreviated self-response instrument). The prototype includes novel features, such as a capacity for accepting and coding responses to open-ended questions. The next phase of the project involves cognitive research to investigate human-machine interface issues, such as the ability of respondents to use the system and the quality of data collected.

The development of computer-assisted data collection for the CPI differs from that for CPS and CES, in that it is forms-based and uses pen-based computer technology. Although much of the research to date has been technical in nature, it has utilized informal hands-on prototyping and focus groups with nearly all of the CPI field representatives (Uglove 1991, 1992).

As BLS moves, along with other U.S. Federal statistical agencies, towards the increased use of computer-assisted survey information collection, we recognize the increased need for basic research on a number of issues. To meet this need, four inter-agency work groups have been established to develop coordinated research agendas in the areas of human-machine interface, changing survey procedures, interviewer training, and interviewer work life. These groups will help prevent duplication of effort across programs and, at the same time, facilitate the sharing of resources and ideas among the agencies.

### 3.5. Interviewer training and evaluation

Interviewer trainers in the Consumer Price Index programs and the Occupation and Compensation Survey programs have used recent advances in training design coming from the social sciences to develop a new interviewer training process. This process is called the Survey Interviewer Certification Program. The goal of this program is to promote the consistent collection of high quality data through the application of performance standards for the interviewers (Fischer 1992).

The standards were developed by personnel in the Office of Field Operations in 1990–1991 based on a careful task analysis of the interviewer's job from the start of the interview to its conclusion. Next, trainers designed components of the training program to allow interviewers to achieve a minimum competency rating on each of the standards. The main components of the certification process are (1) a standard orientation for new workers, (2) observation of data collection prior to classroom training, (3) classroom training, (4) on-the-job training, and (5) evaluation of the interviewer's data collection skills.

The first three components enable the interviewers to achieve the initial stage of skill acquisition, declarative knowledge. The on-the-job training component assists interviewers with knowledge compilation through practice, applying their declarative knowledge in specific situations. Trainers observe and rate the interviewers on their ability to perform certain tasks, and they provide immediate feedback to them verbally and in writing after the interviews. To become certified, novice interviewers must demonstrate minimal competence on all performance standards for a series of interviews. The final training component assesses the experienced interviewer's performance

on the same standards as when initially certified. Moreover, it helps trainers maintain quality standards and teaches interviewers to diagnose and analyze their own performance on a regular basis.

## 4. Data Adjustment and Analysis

### 4.1. Editing and imputation methods

In many establishment surveys, considerable effort goes into reviewing and editing data. An initial effort to automate these processes is underway with the development of an expert system to assist commodity analysts judging the comparability of items priced in two successive periods for the Consumer Price Index (Conrad, Kamalich, Longacre, and Barry 1992). The system being developed simulates the reasoning of the commodity analysts, and, if it proves practical in actual production, will enable analysts to automate some of the more repetitious and time consuming aspects of their job. Moreover, it will lead to more consistent application of the knowledge involved in comparability judgements by codifying and making explicit that expertise.

### 4.2. Estimation

Many of the statistical research projects conducted are motivated by the need to compensate for known nonsampling errors such as undercoverage, nonresponse, and response bias. Others are motivated by changes in data collection procedures. With an emphasis on designing data collection procedures which aid respondents and interviewers, the form of the collected data may change, resulting in a need for different estimation methods.

For example, it was decided that the usefulness of the Occupational Establishment Survey would be improved if corresponding wage data were available. Yet, the reporting

of specific wage rates could become burdensome to respondents. Thus, a distribution of wages by categories was collected. Categories for the smallest and largest wage levels are open-ended, resulting in a need to model the tails of the wage distribution to estimate some finite population quantities. Research using micro-level wage data from other BLS surveys indicates a maximum likelihood method based on the Pareto distribution is more robust than other methods (West, Kratzke, and Butani 1992). In earlier work, West, Butani, and Witt (1990, 1991) and West, Butani, Witt, and Adkins (1989) explored various methods of imputing for missing wage and employment data.

As undercoverage in household surveys increases, the need for improved estimators to account for the undercoverage has increased. Post-stratification has historically been used to force the estimated numbers of persons in household surveys to be equal to age, sex, and race category totals based upon the most recent population census. Despite this use, theory to support the use has been incomplete.

Although the variance of a post-stratified estimator can be computed over all possible sample configurations, inferences made conditionally on the achieved sample configuration are desirable. Theory and a simulation study using data from the Current Population Survey have been used to study both the conditional bias and variance of the post-stratified estimator of a total (Valliant 1993). Results of the study indicate that the conditional variance can be consistently estimated by replication methods like balanced repeated replication or the jackknife, but that it is important to recompute the post stratification adjustments for each replicate. Use of full-sample adjustments can produce large over-estimates of variance.

A new post-stratified regression estimator, which is analogous to the linear regres-

sion estimator in simple random sampling, has been investigated by Casady and Valliant (1992), based upon the approximate multivariate normality of the vector of Horvitz-Thompson estimators of survey target variables and of post-stratum population sizes. Using both real and artificial populations to study empirically the conditional and unconditional properties of the estimators in multistage sampling, their results indicate that the usual post-stratified estimator can be conditionally biased in large samples while the regression estimator is not.

Establishment surveys are also subject to coverage error due to the creation of new businesses. Because of employment changes in new establishments not yet included in the sampling frame for the Current Employment Statistics (CES) survey, payroll employment estimates from the survey are biased. In a test project in New York state, Grzesiak and Lent (1988) identified new businesses (births) among new accounts in the Unemployment Insurance (UI) system and studied alternative estimators of birth employment. The extent of the bias due to missed births can also be measured using reports made by establishments to the UI system. The UI figures for a time period are not available when CES estimates must be published, but can be used for *post hoc* bias estimation. Alternative methods for correcting for this bias have been investigated by Sommers and Stamas (1991). Results indicate that Kalman filter methods can be used to model both employment estimates and their biases. Methods of bias adjustment have been developed that reduce mean squared errors.

Some of the major statistics published by BLS are price indexes. Valliant (1991) and Valliant and Miller (1989) have studied a class of estimators of Laspeyres price indexes that include ones used in practice and others that may have better statistical

properties. Their theoretical and empirical results show that if long-term indexes are calculated by chaining together short-term changes, then data from older time periods should be dampened out to avoid producing estimators with increasing variance. In the last five years, major efforts have been made at BLS to produce variances for published index series. Leaver, Johnstone, and Archer (1991) and Leaver and Swanson (1992) discuss research on variance estimation for the Consumer Price Index (CPI). The CPI uses a relatively complicated multi-stage area sample design in which separate surveys of households are used to derive expenditure weights and an establishment sampling frame. The papers by Leaver et al. describe the combination of Taylor series and replication methods used for variance estimation. Colliia (1988) and Spease (1989) discuss similar work done using replication variances for the U.S. Producer Price Index.

One of the active research areas in survey methods in recent years has been the application of time series methods to estimation problems. In 1989, BLS introduced a time series approach to statewide labor force estimation using data from the Current Population Survey (CPS). Tiller (1989, 1990, 1992) describes the Kalman filter methods used for estimation and the state space framework of labor force models which account for autocorrelations in model errors due to sample overlap between time periods. Initial research on mean squared error estimation for Kalman filter estimates derived from survey data has been completed by Miller (1989b).

One of the key comparisons in evaluating the Kalman filter models for state labor force estimates are variance estimates derived directly from CPS data. Lent (1991) described the replication methods used for the direct variance estimates and their limitations for subnational estimates.

Basic research has been done in the areas of variance estimation and use of computer intensive methods for finite population estimation. Valliant (1990) studied performance of ratio and regression estimators and associated variance estimators in systematic sampling, identifying types of populations where the regression estimator will outperform the ratio estimator. Estimators of distribution functions using non-parametric regression have been examined (Chambers, Dorfman, and Wehrly (1992) and Dorfman and Hall (1992)) and compared to more traditional design-based estimators.

#### *4.3. Analytical methods*

In an organization producing many of the major data series on the U.S. economy, analysis and interpretation by the Bureau of Labor Statistics' economists have fostered an increased need for research on analytical methods, especially for evaluating changes over time. Wootton, Tou, and Miller (1992) developed a method for comparing vectors of industry occupational employment distributions in order to assess whether staffing patterns vary over time. Valliant (1992) developed a method for generalizing variances for price indexes over time using non-parametric scatterplot smoothers. Scott (1989) used similar nonparametric techniques to decompose time series.

New and one-time surveys are being used to collect data on important issues such as employer-provided child care, employee turnover, and employer-provided training. Such surveys are conducted under tight time constraints, so that there is little time for extended research. Moreover, many of them are attempting to measure rare characteristics. Miller, Robertson, and Butani (1992) evaluated the stratified variance estimator and a random group variance estimator,

involving collapsing of strata. Differences were found in the distributions produced by the two variance estimation methods as the relative standard error increases, particularly for rare characteristics where the random group variance estimator overestimated the variance more than predicted by theory.

Distributions of the costs of consumer items are often truncated on the left in the sense that there is some minimum price for a commodity but the minimum may be unknown. Miller (1988) derived a maximum likelihood estimate of such a truncation point and of other model parameters, and illustrated his theory with data from the Consumer Expenditure Survey. Basic research on errors-in-variables models is also covered in Miller (1989a and 1990).

Seasonal adjustment of times series is standard practice in a number of BLS programs. The energy series in the Producer Price Index, in particular, have been subject to volatile changes in the last 15 years. Buszuwski and Scott (1988) tested X-11 seasonal adjustment methods with and without adjustment for structural changes and shocks. Ramp-style interventions along with adjustments for outliers and level shifts were found to improve seasonal adjustments compared to standard procedures.

## 5. References

- Bienias, J. (1988). Preliminary Laboratory Research on Current Population Survey Questionnaire with Employed Persons. American Statistical Association, Proceedings of the Section on Survey Research Methods, 185–190.
- Boehm, L. (1989). Reliability of Proxy Response in the Current Population Survey. American Statistical Association, Proceedings of the Section on Survey Research Methods, 486–489.
- Buszuwski, J. and Scott, S. (1988). On the User of Intervention Analysis in Seasonal Adjustment. American Statistical Association, Proceedings of the Business and Economic Statistics Section, 337–342.
- Campanelli, P., Martin, E., and Creighton, K. (1989). Respondents' Understanding of Labor Force Concepts: Insights from Debriefing Studies. ARC Proceedings, 361–374.
- Cantor, D. and Esposito, J. (1992). Evaluating Interviewer Style for Collecting Industry and Occupation Information. American Statistical Association, Proceedings of the Section on Survey Research Methods, forthcoming.
- Casady, R. (1989). Design and Estimation Methods for List Assisted Telephone Surveys. American Statistical Association, Proceedings of the Section on Survey Research Methods, 138–147.
- Casady, R. and Lepkowski, J. (1991). Hybrid Telephone Sample Design. American Statistical Association, Proceedings of the Section on Survey Research Methods, 111–116.
- Casady, R. and Valliant, R. (1992). Conditional Properties of Post-Stratified Estimators Under Normal Theory. American Statistical Association, Proceedings of the Section on Survey Research Methods, forthcoming.
- Chambers, R., Dorfman, A., and Wehrly, T. (1992). Bias Robust Estimation in Finite Populations Using Nonparametric Regression. Journal of the American Statistical Association, 87, forthcoming.
- Clayton, R. and Harrell, L. (1989). Developing a Cost Model of Alternative Data Collection Methods: Mail, CATI, and TDE. American Statistical Association, Proceedings of the Section on Survey Research Methods, 264–269.
- Clayton, R., Rosen, R., and Winter, D. (1989). Modifying Respondent Behavior:

- Mail vs. Telephone Collection. Paper presented at American Statistical Association Winter Conference.
- Clayton, R. and Winter, D. (1990). Speech Data Entry: Results of the First Test of Voice Recognition for Data Collection. American Statistical Association, Proceedings of the Section on Survey Research Methods, 387-393.
- Colia, D. (1988). Measuring Sample Variability in the Producer Price Index. American Statistical Association, Proceedings of the Section on Survey Research Methods, 710-715.
- Conrad, F., Kamalich, R., Longacre, J., and Barry, D. (1992). An Expert System for Reviewing Commodity Substitutions in the Consumer Price Index. Paper presented at the Conference on Computing for the Social Sciences, Ann Arbor, MI.
- Dippo, C. (1987). A Review of Statistical Research at the U.S. Bureau of Labor Statistics. *Journal of Official Statistics*, 3, 289-297.
- Dippo, C. (1989). The Use of Cognitive Laboratory Techniques for Investigating Memory Retrieval Errors in Restrospective Surveys. *Bulletin of the International Statistical Institute*, 53, 363-382.
- Dippo, C. and Herrmann, D. (1991). The Bureau of Labor Statistics Collection Procedure Research Laboratory: Accomplishments and Future Directions. Statistical Policy Working Paper 20, Office of Management and Budget, Washington, D.C., 253-267.
- Dippo, C. and Norwood, J. (1992). A Review of Research at the Bureau of Labor Statistics. In *Questions About Questions: Inquiries into the Cognitive Bases of Surveys*, edited by Tanur, J., New York: Russell Sage.
- Dorfman, A. and Hall, P. (1992). Estimators of the Finite Population Distribution Function Using Nonparametric Regression. *Annals of Statistics*, 21, forthcoming.
- Edwards, W., Levine, R., and Cohany, S. (1989). Procedures for Validating Report on Hours Worked and for Classifying Discrepancies Between Questionnaire Reports and Validation Totals. American Statistical Association, Proceedings of the Section on Survey Research Methods, 496-501.
- Esposito, J., Campanelli, P., Rothgeb, J., and Polivka, A. (1991). Determining Which Questions Are Best: Methodologies for Evaluating Survey Questions. American Statistical Association, Proceedings of the Section on Survey Research Methods, 46-55.
- Esposito, J. and Jobe, J. (1991). A General Model of the Survey Interaction Process. U.S. Bureau of the Census, Proceedings of the 1991 Annual Research Conference, 537-560.
- Fisher, D. (1992). Staff Development Models in Training at the Bureau of Labor Statistics. Paper presented at the Field Directors' Conference, St. Petersburg, FL.
- Fracasso, M. (1989). Reliability and Validity of Response Categories for Open-ended Questions in the Current Population Survey. American Statistical Association, Proceedings of the Section on Survey Research Methods, 481-485.
- Gaertner, G., Cantor, D., Shank, S., and Gay, N. (1989). Tests of Alternative Questions for Measuring Industry and Occupation in the CPS. American Statistical Association, Proceedings of the Section on Survey Research Methods, 490-495.
- Grzesiak, T. and Lent, J. (1988). Estimating Business Birth Employment in the Current Employment Statistics Program. American Statistical Association, Pro-

- ceedings of the Section on Survey Research Methods, 597–602.
- Leaver, S., Johnstone, J., and Archer, K. (1991). Estimating Unconditional Variances for the U.S. Consumer Price Index for 1978–1986. American Statistical Association, Proceedings of the Section on Survey Research Methods, 614–619.
- Leaver, S. and Swanson, D. (1992). Estimating Variances for the U.S. Consumer Price Index for 1987–1991. American Statistical Association, Proceedings of the Section on Survey Research Methods, forthcoming.
- Lent, J. (1991). Variance Estimation for the Current Population Survey. American Statistical Association, Proceedings of the Section on Survey Research Methods, 11–20.
- McKay, R. (1991). Effect of Central American Cultures on Household Surveys in the U.S. Paper presented at the American Anthropological Association Meetings, Nov 21–24, Chicago, IL.
- McKay, R. (1992a). Cultural Pattern Affecting Within Household Coverage and Proxy Reporting in Hispanic (Salvadoran) Households. Paper presented at the American Statistical Association Meetings, Boston, MA.
- McKay, R. (1992b). The Place of the Boarder in Hispanic Households. Paper presented at the American Sociological Association Meetings, Pittsburgh, PA.
- McKay, R., Malik, A., and Tupek, A. (1992). Telephone Data Entry for the Current Population Survey. Paper presented at the Field Technologies Conference, St. Petersburg, FL.
- Miller, L. and Downes-LeGuin, T. (1989). Improving Comprehension and Recall in the Consumer Expenditure Interview Survey: Discrepancies in Comprehension and Recall as a Source of Non-Sampling Error. American Statistical Association, Proceedings of the Section on Survey Research Methods, 502–507.
- Miller, L. and Downes-LeGuin, T. (1990). Reducing Response Error in Consumers' Reports of Medical Expenses: Application of Cognitive Theory to the Consumer Expenditure Interview Survey. In *Advances in Consumer Research*, Vol. 17, edited by Marvin E. Goldberg, Jerold Gorn, and Richard W. Pollay, 193–197, Association for Consumer Research, Provo, UT.
- Miller, L., Herrmann, D., and Puskar, M. (1992). The Accuracy of Question Presentation and the Adequacy of Response. Internal Bureau of Labor Statistics Memorandum.
- Miller, S. (1988). Truncated Regression Models Under a Monotone Transformation. American Statistical Association, Proceedings of the Business and Economic Statistics Section, 157–162.
- Miller, S. (1989a). Empirical Processes Based Upon Residuals from Errors-in-Variables Regressions. *Annals of Statistics*, 17, 282–292.
- Miller, S. (1989b). Mean Square Error Estimation of Kalman Filter Estimates from Aggregate Survey Data. American Statistical Association, Proceedings of the Section on Survey Research Methods, 718–723.
- Miller, S. (1990). Analysis of Residuals from Measurement Error Models. *Contemporary Mathematics*. Proceedings of a Summer Research Conference, June 1989, 59–71.
- Miller, S., Robertson, K., and Butani, S. (1992). Comparison of the Random Group Variance and the Stratified Variance Estimators. American Statistical Association, Proceedings of the Section on Survey Research Methods, forthcoming.

- Mullin, P., Conrad, F., Sander, J., and Herrmann, D. (1992). A Cognitive Theory of the Survey Interviews. American Statistical Association, Proceedings of the Section on Survey Research Methods, forthcoming.
- Palmisano, M. (1990). Respondent Understanding of Key Labor Force Concepts Used in the CPS. American Statistical Association, Proceedings of the Section on Survey Research Methods, in press.
- Phipps, P. (1990). Applying Cognitive Theory to an Establishment Mail Survey. American Statistical Association, Proceedings of the Section on Survey Research Methods, 608–612.
- Phipps, P., Butani, S., and Chun, Y. (1992). Designing Establishment Survey Questionnaires. Paper presented at the American Association for Public Opinion Research Conference, St. Petersburg, FL.
- Phipps, P., Robertson, K., and Keel, K. (1991). Does Questionnaire Color Affect Survey Response Rates? American Statistical Association, Proceedings of the Section on Survey Research Methods, 484–488.
- Phipps, P. and Tupek, A. (1991). Assessing Measurement Errors in a Touchtone Recognition Survey. *Survey Methodology*, 17, 15–26.
- Principe, J. and Sommers, J. (1992). Alternative Sampling Strategies of the New BLS Business Establishment List. American Statistical Association, Proceedings of the Section on Survey Research Methods, forthcoming.
- Rothgeb, J., Polivka, A., Creighton, K., and Cohany, S. (1991). Development of the Proposed Revised Current Population Survey. American Statistical Association, Proceedings of the Section on Survey Research Methods, 56–65.
- Scott, S. (1989). Applications of Locally-Weighted Regression to Time Series Decomposition. American Statistical Association, Proceedings of the Section on Survey Research Methods, 467–472.
- Silberstein, A. (1990). First Wave Effects in the U.S. Consumer Expenditure Interview Survey. *Survey Methodology*, 16, 293–304.
- Silberstein, A. (1991). Response Performance in the Consumer Expenditure Diary Survey. American Statistical Association, Proceedings of the Section on Survey Research Methods, 338–343.
- Silberstein, A. and Scott, S. (1991). Expenditure Diary Surveys and Their Associated Errors. In *Measurement Errors in Surveys*, edited by Biemer, Groves, Lyberg, Mathiowetz, and Sudman, 303–326, New York: John Wiley.
- Sommers, J. and Stamas, G. (1991). Correcting for Bias in Payroll Employment Estimates. American Statistical Association, Proceedings of the Business and Economic Statistics Section, 298–303.
- Spease, C. (1989). Comparison of Variance Estimators for Producer Price Index Data. American Statistical Association, Proceedings of the Section on Survey Research Methods, 176–183.
- Tiller, R. (1989). A Kalman Filter Approach to Labor Force Estimation Using Survey Data. American Statistical Association, Proceedings of the Section on Survey Research Methods, 16–25.
- Tiller, R. (1990). An Application of Times Series Methods to Labor Force Estimation Using CPS Data. American Statistical Association, Proceedings of the Section on Survey Research Methods, 284–289.
- Tiller, R. (1992). Time Series Modeling of Sample Survey Data from the U.S. Current Population Survey. *Journal of Official Statistics*, 8, 149–166.
- Tucker, C. (1992). The Estimation of Instrument Effects on Data Quality in the Con-

- sumer Expenditure Diary Survey. *Journal of Official Statistics*, 8, 41–61.
- Tucker, C. and Bennett, C. (1988). Procedural Effects in the Collection of Consumer Expenditure Information: The Diary Operational Test. American Statistical Association, Proceedings of the Section on Survey Research Methods, 256–261.
- Tucker, C., Casady, R., and Lepkowski, J. (1991). An Evaluation of the 1988 Current Point-of-Purchase CATI Feasibility Test. American Statistical Association, Proceedings of the Section on Survey Research Methods, 508–513.
- Tucker, C., Casady, R., and Lepkowski, J. (1992). Sample Allocation for Stratified Telephone Sample Designs. American Statistical Association, Proceedings of the Section on Survey Research Methods, forthcoming.
- Tucker, C., Vitrano, F., Miller, L., and Doddy, J. (1989). Cognitive Issues and Research in the Consumer Expenditure Diary Survey. Paper presented at the American Association for Public Opinion Research Conference, St. Petersburg, FL.
- Uglov, D. (1992). Pen-based Computers: Remodeling? Paper presented at the Field Technologies Conference, St. Petersburg, FL.
- Uglov, D. (1991). Design Issues in Forms-based CASIC Instruments. Paper presented at the Field Technologies Conference, San Diego, CA.
- Valliant, R. (1990). Comparisons of Variance Estimators in Stratified Random and Systematic Sampling. *Journal of Official Statistics*, 6, 115–132.
- Valliant, R. (1991). Variance Estimation for Price Indexes from a Sample with Rotating Panels. *Journal of Business and Economic Statistics*, 9, 409–422.
- Valliant, R. (1992). Smoothing Variance Estimates for Price Indexes Over Time. *Journal of Official Statistics*, 8, 433–444.
- Valliant, R. (1993). Post-Stratification and Conditional Variance Estimation. *Journal of the American Statistical Association*, 88, forthcoming.
- Valliant, R. and Miller, S. (1989). A Class of Multiplicative Estimators of Laspeyres Price Indexes. *Journal of Business and Economic Statistics*, 7, 387–394.
- van Melis-Wright, M., Stone, D., and Herrmann, D. (1990). The Semantic Basis of Confidentiality. BLS Statistical Note 34.
- van Melis-Wright, M. (1991). Confidentiality and Social Conventions for Inquiring About Personal Topics. Paper presented at the American Psychology Association Meetings, San Francisco, CA.
- Werking, G. and Clayton, R. (1990). Enhancing the Quality of Time Critical Estimates Through the Use of Mixed Mode CATI/CASI Collection. Statistics Canada, Proceedings of the Symposium on Management and Improvement on Data Quality.
- Werking, G., Tupek, A., and Clayton, R. (1988). CATI and Touch-tone Self-response Applications for Establishment Surveys. U.S. Bureau of the Census, Proceedings of the Fourth Annual Research Conference, 273–290.
- West, S., Butani, S., Witt, M., and Adkins, C. (1989). Alternative Imputation Methods for Employment Data. American Statistical Association, Proceedings of the Section on Survey Research Methods, 227–232.
- West, S., Butani, S., and Witt, M. (1990). Alternative Imputation Methods for Wage Data. American Statistical Association, Proceedings of the Section on Survey Research Methods, 254–259.
- West, S., Butani, S., and Witt, M. (1991). Alternative Imputation Methods for Employment, Wage, and Ratio of Wage to Employment Data. Proceedings of the

78th Indian Science Congress, forthcoming.

West, S., Kratzke, D., and Butani, S. (1992). Measures of Central Tendency for Censored Wage Data. American Statistical Association, Proceedings of the Section on Survey Research Methods, forthcoming.

Wootton, B., Tou, A., and Miller, S. (1992). Variations in Occupational Staffing Patterns. BLS Statistical Note 33.

Received October 1992  
Revised December 1992