

United States General Accounting Office

GAO

Report to the Chairman, Committee on
Government Operations
House of Representatives



LM129586

March 1986

STATISTICAL SURVEYS

Census Bureau Has Creditable Employment and Economic Data- Collection Procedures



RELEASED

RESTRICTED—Not to be released outside the General
Accounting Office except on the basis of specific
approval by the Office of Congressional Relations.

GAO/IMTEC-86-8

534819 / 129586



United States
General Accounting Office
Washington, D.C. 20548

Information Management and
Technology Division
B-222011

March 14, 1986

The Honorable Jack Brooks
Chairman, Committee on
Government Operations
House of Representatives

Dear Mr. Chairman:

In response to your request and subsequent discussions with your office, we are providing the results of our review of two statistical surveys, the Current Population Survey (CPS) and the Survey of Income and Program Participation (SIPP). Information for CPS and SIPP is compiled by the Department of Commerce's Bureau of the Census, a principal supplier of national statistical data. The Bureau is also responsible for collecting and publishing numerous reports on economic conditions, health, recreation, and, most notably, population (through the count performed every decade).

CPS reflects monthly employment, unemployment, and general labor-force characteristics; SIPP provides a continuing measure of individuals' and households' economic conditions, sources of income, labor-force activity, and participation in federal benefit programs. Data for CPS, SIPP, and other Bureau surveys are collected by part-time interviewers who contact the interviewees in person or by telephone. Participation in most surveys is voluntary. Responses are recorded on questionnaires and forwarded to regional offices for review, processing, and transmission to processing centers for final computation of the statistical estimates.

As a result of discussions with your office, our review focused on the procedures and controls of the CPS and SIPP data collection and review, but not on the entire survey process. We concentrated on questionnaire design; selection, training, and supervision of interviewers; and quality-control elements used in gathering and reviewing data. As agreed with your office, we selected 2 of 12 Bureau regional offices for on-site visits to assess controls for the data-collection and review activities. Detailed information on the surveys and our work is contained in appendix I, which follows this letter.

We have determined from our review that

- each survey's questionnaire design, though reliable for compiling statistical estimates and making needed comparisons, could be redesigned to improve the accuracy of certain items;
- selection of interviewers is based on reasonable criteria to perform the interview function;
- interviewer training is comprehensive and adequately covers the necessary interviewing skills;
- supervisory and administrative quality controls for the data-collection and review activities are in place and operating to reasonably ensure data integrity; and
- the regions visited are doing a creditable job of complying with the data-collection and review procedures.

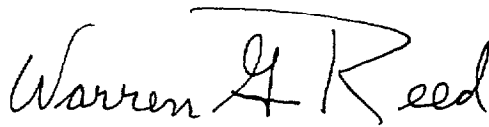
While we realize that our findings at 2 regional offices do not allow us to comment on performance at the other 10, we believe that if the others are complying with these procedures they also should be doing a creditable job of providing for the collection and reporting of timely CPS and SIPP estimates.

During our review we became aware of other issues that we believe can affect the quality of data. They are how interviewers are attracted and retained and how respondents' participation may influence data quality. In appendix II we include a limited discussion of these factors to more accurately present the environment surrounding the question of quality federal statistics. The factors can neither be measured with precision nor easily quantified; therefore, categorically stating the degree to which they enhance or diminish data quality is difficult.

We discussed key elements of this report with Bureau program officials and made changes as appropriate to reflect relevant factual information they provided. We did not request official agency comments on this report. As arranged with your office, unless you publicly announce its

contents earlier, we plan no further distribution for 30 days from the date of this report. We will then send copies to interested parties and make copies available to others upon request.

Sincerely yours,

A handwritten signature in black ink that reads "Warren G. Reed". The signature is written in a cursive style with a large, sweeping "R" and "E" at the end.

Warren G. Reed
Director

Procedures for Collecting CPS and SIPP Data Are Creditable

The Bureau of the Census is a principal supplier of federal statistical data, including national economic, income, employment and population estimates, that it compiles and reports through various statistical surveys and programs. Decision-makers in government use these data to formulate national policy; managers in private industry use them for developing business strategies. Because statistical estimates have far-reaching impact, the accuracy of data behind the published numbers is vital. Therefore, procedures and controls for data integrity and reliability are essential.

This report describes how statistical data are collected and processed for two Bureau surveys: CPS and SIPP. It discusses the controls the Bureau employs to reasonably ensure data quality and enhance the reliability of estimates reported for CPS and SIPP. Also discussed are other factors that can impact data quality but that, in our opinion, cannot be totally controlled by the Bureau's quality-control procedures.

Overview of CPS and SIPP

CPS and SIPP are monthly household surveys designed to provide employment and income statistics on individuals and families living in the United States. These estimates serve as indicators for measuring the current status and progression of the nation's economy and for deciding, monitoring, and understanding matters of national economic policy.

Since about 1940, CPS has been the primary source of data for government estimates on the level of employment, unemployment, and other key labor-force statistics. The Bureau of Labor Statistics, the primary sponsor for the information collected, releases Census' CPS data in the form of the monthly employment statistical estimates. CPS data are also used for establishing a basis for distributing federal funds to state and local areas. CPS data are obtained through personal interviews from a nationwide sample of about 75,000 households. In fiscal year 1984 the CPS program cost about \$23 million.

SIPP, also a monthly survey, provides estimates on individual and family income, financial condition, employment status, and participation in government benefit programs. Initiated by the Bureau in October 1983, SIPP should reach full operation during fiscal year 1986. In its first year, SIPP's sample size numbered approximately 26,000 households and cost about \$12 million. The Bureau planned to increase the sample size to 42,000 households in fiscal year 1985 and 60,000 in fiscal year 1986. However, due to a \$1-million budget reduction, the households to be surveyed totaled about 36,000 in fiscal year 1985.

SIPP data are intended to support government planning, policy, and program efforts by expanding the information available on income distribution, poverty, and government assistance. According to the Bureau, SIPP data are used specifically to assess the

- effect of changes in eligibility or government benefits on program recipients and federal spending levels;
- extent of benefits received by persons and families simultaneously participating in several federal-assistance programs and whether combined benefit levels are adequate or excessive; and
- reasons for changes in benefit status, employment, and household membership and ways of moderating undesirable changes in benefit status and employment.

Objective, Scope, and Methodology

Our objective was to assess the quality of two of the Bureau's statistical surveys, CPS and SIPP. As outlined in the letter of request and subsequent conversations with the Chairman's office, we focused on the adequacy of questionnaire design; interviewer selection, training, and supervision; and quality controls for the data-collection and review processes. As agreed upon, this review does not address the entire scope of the survey process. We did not examine CPS and SIPP sample designs, which identify the persons and/or households to be interviewed; nor did we interview survey participants to validate their responses. Also, we did not review data processing operations at Bureau headquarters or the formulas, processes, or computer data-editing and tabulation programs used.

We performed our review between November 1984 and August 1985 at the Bureau's headquarters in Suitland, Maryland, and two regional offices—New York and Kansas City. The regional offices were selected because they provide the greatest differences in performance and represent different environments with respect to data-gathering difficulties. For example, Kansas City traditionally has a rate of approximately 97.0 percent for completed CPS and SIPP interviews, whereas New York maintains a rate of around 90.3 percent. Kansas City is considered primarily rural, midwestern, and not an extremely difficult area in which to conduct interviews; New York is representative of urban areas, where it is considered more difficult to obtain data from respondents.

We examined Bureau survey procedures and manuals and conducted interviews with Bureau officials and employees. We developed flow charts for each survey, listing major functions and processes for data

collection and review. From the charts, we identified factors that we believe could potentially affect the data quality.

We then identified the Bureau's controls for minimizing potential problems and conducted on-site visits at the two regional offices to determine if the prescribed controls were in place, operating, and being followed. At each regional office we observed CPS and SIPP survey operations, participated in interviewer training sessions, and examined interviewers' personnel records to verify that training requirements and performance levels were being met.

To assess the training provided CPS and SIPP interviewers, we reviewed, in detail, each survey's home study courses, classroom training materials and manuals, and the interviewer's and instructor's manuals. We attended training sessions and interviewed new and experienced interviewers.

We conducted our review in accordance with generally accepted government auditing standards. Criteria for assessing data quality and controls were developed through a literature search; the General Accounting Office's methodology for assessing internal controls; review of Bureau policy and procedures manuals; and discussions with recognized statistical experts from Census, the Bureau of Labor Statistics, and private industry. To evaluate the CPS and SIPP questionnaires, we performed a technical review of questionnaire design. We sought the views of responsible officials during the course of our work, and we have incorporated their views in the report where appropriate. As the Chairman requested, we did not ask for official agency review and comment on this document.

CPS and SIPP Data- Collection Staff and Procedures

Various Bureau headquarters divisions are involved in the surveys' planning and development and in prescribing the data-collection processes. Headquarters' Field Division has the overall management responsibility for the survey process. Data collection is performed through a nationwide network of 12 regional offices, each headed by a director. Regional directors have the day-to-day responsibility of ensuring that collection of survey data is carried out according to headquarters' procedures. The survey collection process is managed in each region by a program coordinator, assisted by CPS and SIPP supervisors.

**Appendix I
Procedures for Collecting CPS and SIPP Data
Are Creditable**

Regional-office Staff

CPS and SIPP supervisors' day-to-day management responsibility primarily entails supervising the interviewers and overseeing the collection process. Interviewers usually collect data by conducting interviews at the respondent's home, completing standard survey questionnaires.

Assisting supervisors are Supervisory Field Representatives (SFR), seasoned interviewers who help to resolve day-to-day problems, and office clerks who perform processing functions when questionnaires are returned to the regional office. The following table illustrates the 1984 apportioning of CPS and SIPP staff for the two regional offices visited.

Table I.1: Regional-office Personnel

CPS	New York	Kansas City
Interviewers	92	129
SFRs	9	19
Supervisors	1	1
Total	102	149
SIPP		
Interviewers	20	23
SFRs	7	11
Supervisors	1	1
Total	28	35

Regional offices are also responsible for hiring, training, and supervising interviewers according to standards established by headquarters. Interviewers work part-time from their homes, and often receive minimal day-to-day supervision where distance separates them from the regional office. Communication between supervisors and interviewers, who may work hundreds of miles from the office, is usually by telephone or mail. For routine matters, this is generally sufficient. For nonroutine matters or serious problems, either a supervisor or SFR will visit the interviewer.

Survey Procedures

Households selected for CPS are interviewed monthly during the first and last 4 months of a 16-month cycle. SIPP interviews are conducted once every 4 months for 2 1/2 years. To make the work loads more manageable, approximately 25 percent of the SIPP sample households are contacted each month, rather than interviewing all participants in a single month. Most CPS interviews are completed during the third week of each month; SIPP interviews are completed within the first half of each month. Interviewers send completed questionnaires to the regional

office; there they are logged in and checked for completeness. CPS questionnaires are then shipped to the Jeffersonville, Indiana, data processing center, where the information is coded and subsequently transmitted to the CPS data processing system at headquarters. Due to time constraints in collecting, processing, and reporting data, the questionnaires containing computer-identified errors are not returned to the regional office for correction. Instead, the information enters an automated edit/imputation process that corrects errors and/or accounts for missing data.

SIPP data processing occurs in each regional office. Here the data is key entered and transmitted to the SIPP data processing system at headquarters. SIPP data are computer edited and are either accepted or rejected by the system. Data containing computer-identified errors are transmitted back to the originating regional office for correction and resubmission. This process continues until all SIPP data are accepted by the system.

The data-collection phase is virtually complete after the data's entry into the automated systems. Beyond this point, computer processes edit the data to correct errors and account for missing survey responses (nonresponses).

Quality Controls for Data Collection, Review, and Processing Are in Place and Operational

To achieve the CPS and SIPP survey objectives, the Bureau has established policies and procedures to ensure data quality through the survey processes. Producing CPS and SIPP estimates involves three phases—collecting, reviewing, and processing the necessary data—to produce the final estimates.

For data collection, controls are exercised through the questionnaire design, interviewer selection and training, and observations of interviewer performance. Regional-office controls over data collection include procedural and administrative checks on the data's completeness. Following these checks, questionnaires are either forwarded for processing to compile the survey estimates or resubmitted for resolution of questionable data.

We observed that established controls over data-collection activities and procedural and administrative reviews are being followed. We believe the policies and procedures in force at the two regions are sound.

**Controls Over Data-
Collection Activities**

Applying controls over the data-collection activities is a critical element in achieving data quality. Questionnaire design, interviewer selection and training, and periodic observations and evaluation of the interviewers' performance are the Bureau's primary controls to help ensure that data collection is uniformly and correctly accomplished.

Questionnaire Design

Survey questionnaire design can influence the data received. For example, the wording and order of questions can impact responses. Ideally, question design serves to produce the most accurate information possible; however, interviewers must be comfortable in asking the questions as worded, and respondents must understand what information they are being asked to provide. Also, variations in the way a question is asked and/or a respondent's failure to understand it can cause changes in the data.

We found that potential for bias may result from (1) use of an "outdated" CPS questionnaire and (2) how the CPS and SIPP questions are asked.

The basic CPS questionnaire has been employed since 1961 with only minor changes. Yet over the years decision-makers have desired more specific information, and the questionnaire has not been updated to accomplish this. To cite two examples, family relationships and racial classifications themselves have changed. The CPS questionnaire limits the family relationship choices to the traditional nuclear unit of children and parents. It does not provide for documenting other relationships possible in today's more expansive family structure. Also, the CPS questionnaire limits choices for race to white, black, American Indian or Aleut Eskimo, Asian or Pacific Islander, and other. For both examples, family relationships and race, respondents who do not exactly fit into the categories provided may be confused on how to best respond to the question and/or may be reluctant to provide the information. Consequently, the most accurate information might not be obtained.

Although racial classification may be more accurately depicted in the succeeding question on origin, responses to the race classification question are not altered to have the responses on race and origin agree. It would be more accurate to provide response choices that are more current with today's information needs.

In relation to work, there is, again, an insufficient range of alternatives. For example, categories for why a person was absent from work do not

include choices such as the unavailability of child care, dependent or family illness, important personal business, or lack of transportation. Also, questions about why a person started looking for work, or left a job, and what a person did to look for work do not have comprehensive choices. Limiting the choices may result in not obtaining the most useful information for the aggregate results.

For both surveys the interviewers ask the question, then fit the response into the closest possible choice listed (free-answer method) rather than reading the fixed-answer choices appearing on the questionnaire (aided-recall method). For example, a CPS question asks why the respondent is not looking for work. Respondents planning a business may not volunteer this reason and, under the free-answer method, would not be aided in recalling it. Using the aided-recall method of providing the respondent with a complete range of realistic choices could elicit more accurate answers than using the free-answer method.

The relative merits of each method have been debated for decades. The use of one or the other may result in different estimates. The preponderance of evidence favors the free-answer method for research and exploratory work rather than for ongoing surveys such as SIPP and CPS. We believe that, considering the nature of CPS and SIPP surveys, well-designed and tested aided-recall questionnaires would provide more accurate estimates than free-answer questionnaires.

According to Bureau officials, major changes have not been made because funding has not been available to do the research necessary to determine how changes in the questions would affect data collection.

Interviewer Selection

The extent to which interviewers are properly selected and trained impacts the effectiveness of any successful controls over the survey processes.

The Bureau has several basic requirements for interviewers' selection. Applicants must be U.S. citizens and at least 18 years old (or 16 with a high school diploma or equivalent). They must pass a test demonstrating their ability to read instructions and maps and to do accurate clerical work. Additionally, they must have an available automobile and good hearing, be in good physical condition, and be able to work in all types of weather. Applicants must be able to attend training sessions (often away from home), have a home telephone, be available for day, evening, and Saturday work, and reside in the county or city where most of the

assignment is located. Applicants passing the basic test must also complete a standard federal application form listing education, work or related experience, and references. Supervisors also interview the applicants to describe the job characteristics, duties, and responsibilities.

Both regional offices we visited use the same interviewer selection process, as prescribed by the Bureau. Our discussions with Bureau officials indicated satisfaction with the interviewer selection standards; in fact, they strongly endorse the selection criteria and process as being effective in ensuring that capable individuals are hired.

From our analysis of the tasks that an interviewer must perform, we believe the Bureau has adequately defined the job characteristics, employs reasonable selection techniques, and has a reasonable interviewer selection process.

Interviewer Training

Once an applicant is hired and assigned to a specific survey, training begins. Training explains the survey concepts and prescribed procedures the Bureau has determined that interviewers must know for collecting the most complete and accurate data possible. Additionally, the training attempts to develop the personal-relations skills essential for eliciting information.

Headquarters determines the training content and materials. Most training is actually conducted in the regional offices and is a continuous process consisting of initial, refresher, and special training. The Bureau requires that all interviewers on each survey receive the same initial and refresher training. This helps standardize how the interviews are conducted and information is obtained, and helps to reduce interviewer error or bias that might affect the interview results.

For both CPS and SIPP, the initial training combines home study with classroom training, followed by on-the-job training. The self-paced, pre-classroom home study course takes approximately 7 hours to complete for CPS and approximately 8 hours to complete for SIPP. Afterwards, the trainee should know the basic rules of interviewing and specific survey definitions and should be able to use the survey questionnaire. CPS and SIPP classroom training, which lasts 3 and 4 days respectively, reinforces the home study with lectures, written and oral quizzes, mock interviews, role-playing exercises, and testing of these job facets before any actual interviewing is done.

In reviewing the initial CPS and SIPP interviewer training, we found the manuals and training materials to be written, illustrated, and organized well. The manuals and classroom sessions included short quizzes followed by answers for immediate feedback. Classroom video presentations showed properly conducted interviews and examples of interviewer error, both of which were later discussed.

During the classroom sessions we observed role playing and mock interviews that exposed trainees to potential interviewing situations. Experienced interviewers are sometimes used to recreate situations they have encountered during actual interviews. Discussions of experiences by more seasoned interviewers help prepare trainees for their interviewing duties.

On-the-job training follows home study and classroom training. Here, the new interviewer is accompanied by a supervisor or SFR on the first assignment. This practice supports the new interviewer, helps to minimize errors of inexperience, provides immediate feedback, and ensures assignment completion and compliance with procedures. A second observation session is conducted later to further evaluate the new interviewer. CPS utilizes a third observation within 6 months of the first assignment. If deficiencies are repeatedly identified, the interviewer is provided additional training or coaching.

When a new interviewer's production, noninterview, and error rates fall within the satisfactory range established by the Bureau, the initial stage of training is considered to be complete. In the second stage of training, the Bureau conducts periodic and systematic refresher sessions and furnishes self-study exercises. Generally, these are offered in conjunction with the monthly CPS interview cycle, a change in the questionnaire or interviewing process or procedures, or the introduction of each new SIPP questionnaire. Such training can require several hours or days, depending on the amount of information to be covered. In addition, group training sessions, usually conducted two or three times yearly, address changes to the survey and/or procedures.

The third stage, special training, is remedial and is required if an interviewer's performance is found to be marginal or unsatisfactory: if an interviewer has a low production rate (e.g., takes too long), a high number of uncompleted interviews (the noninterview rate), and/or too many errors (such as improperly completed questionnaires). Supervisors can choose from several types of special training; a trainee could receive

feedback through “special-needs” observations or could take self-study courses.

From our review, observations, and discussions, we believe the training program for CPS and SIPP adequately covers the purpose, processes, requirements, expectations, and manner in which interviews should be conducted and recorded. Our review of the personnel folders showed that, during 1984, interviewers received the required training and the two regions visited had complied with Bureau training requirements.

**Performance Observations and
Evaluations**

Other quality-control measures exercised during data collection are observations and evaluations of interviewer performance. Observing interviewers provides a method for assessing actual performance and helps ensure mastery of skills such as probing a respondent for appropriate responses and timely completion of an interview.

Observers must be SFRS or regional-office supervisory staff with recent CPS or SIPP survey experience. The observer must spend at least 6 hours observing the interviewer’s activities. Observations occur annually for CPS and semiannually for SIPP unless cause for a special-needs observation surfaces. A special-needs observation that can occur at any time is conducted to determine the reasons for less-than-satisfactory performance rates. In all cases, observers note and later attempt to discuss and/or correct problems an interviewer may be having. An observation report is then prepared, reviewed by a supervisor, and provided to the interviewer.

Performance evaluations, prepared quarterly on all interviewers, help reassure both interviewers and the Bureau that performance standards are being met. Critical factors used for the evaluations are certain noninterview, error, and production rates. The evaluation is a composite of three successive monthly performance reports prepared by supervisors from computer or office edits or other documents.

To determine if interviewers in both regional offices received the required performance observations and evaluations during 1984, we reviewed a judgmental sample (about 15 percent) of CPS and SIPP interviewer personnel folders. We found the offices in compliance with the Bureau performance observations and evaluation requirements and the administrative records complete. Dates of observations, deficiencies noted, and recommended follow-up action, such as refresher courses, were recorded. We also examined interviewer performance, production,

and error rates and determined that unsatisfactory or marginal interviewers were receiving appropriate monitoring and additional training. During our interviews, both Bureau officials and current interviewers expressed the opinion that the training and performance evaluations were informative and beneficial.

Controls Over Data Review and Processing

Each regional office plays a critical role in ensuring data quality by executing certain procedures in a specified and uniform manner both prior to and after the CPS and SIPP interviews. The regional offices have assignment controls for accountability and for balancing work loads among the interviewers. When the completed questionnaires reach the regional office, they are logged in, reviewed for administrative completeness, and edited by the office staff. Further control is exercised by following up on questionnaires returned as noninterviews when a supervisor believes they can be converted into completed interviews. Also, control is exercised through reinterviews, wherein selected respondents are questioned again and their answers compared against their earlier responses. Finally, data processing controls are in place for data transmission and document shipping.

Assignment Control

The regional office must account for questionnaires for each person or household unit in the interviewer's survey assignment and must balance the work load among the interviewers. This is assignment control. Supervisors are responsible for preparing assignments and ensuring both their equitability among interviewers and their completion within required timeframes. When making assignments, supervisors also consider such factors as location of the interviewer's residence, difficulty in interviewing households in the assignment area, and the interviewer's ability. The regional offices provide interviewers with written assignments and the appropriate survey materials prior to the interview. Interviewers then review their assignments and plan their itinerary.

Interviewers return questionnaires to the regional office as they complete their interviews or after determining that it is not possible to conduct an interview. A respondent's refusing to give information or an interviewer's confirming that a respondent will not be available constitutes a noninterview. Both noninterview and completed questionnaires are returned to the regional office, where survey clerks log them in and supervisors monitor the work flow for timeliness. Should interviewers fall behind schedule, supervisors determine reasons for the delay and, if

needed, take such action as offering assistance or assigning another or more-experienced interviewer to assist in completing the assignment.

From our observations and interviews with responsible regional personnel, we found that the Bureau's procedures for controlling assignments are adhered to and that these controls assist in balancing the work loads at both regional offices. We also found that the daily surveillance of returned questionnaires facilitates a supervisor's detecting an interviewer's falling behind schedule and taking appropriate action. As a result, we believe the assignment-control procedures balance the work load and ensure accountability for the completed questionnaires.

Office Edits

Questionnaires are edited by regional-office survey clerks to detect errors in preparation (e.g., noncompliance with procedures, omissions, and inconsistencies). Correcting sloppy marking of questionnaires is especially important to CPS where the questionnaires are machine read.

Because editing CPS and SIPP questionnaires is costly and time-consuming, only certain questionnaires are edited. The Bureau requires that the questionnaires of new interviewers and interviewers failing to meet minimum error standards be edited in their entirety. Experienced interviewers continuously meeting or exceeding minimum error standards have only selected items edited. Editing standards are based on an interviewer's previous work as well as Bureau guidelines.

Survey clerks finding errors have several options. The mark for CPS answers can be neatly completed. Missing data, such as name or zip code, can be inserted on the basis of other available data. Where necessary, the survey clerk can contact the interviewer. Where errors cannot be corrected, the survey clerk is directed to leave the question unanswered. Later processing in the headquarters edit/imputation phase will account for an answer.

We observed the editing process in both regional offices, finding that it was being performed appropriately and that the regional offices are complying with Bureau procedures. We believe the edit function improves the data quality: administrative errors are corrected rather than the questionnaire data's not being entered because of minor problems.

Follow-up

Noninterview questionnaires are reviewed by regional-office supervisors who determine what can be done when respondents refuse to participate, are not at home, or are known to be temporarily absent, say, on vacation. If a supervisor judges that noninterviews can be converted to completed interviews, follow-up action is taken. If the interview cannot be performed, a noninterview is carried against the performance ratings of the interviewer and the regional office.

We observed that supervisory personnel at both regional offices review all noninterview questionnaires. Based on their experience, they decide if follow-up actions—calling the nonrespondents themselves, sending out a more-experienced interviewer, or sending an explanatory letter signed by the regional director to the nonrespondent—would be appropriate. A CPS follow-up method uses “telephone holds” in cases where an interviewer, believing an interview is possible in the near future, meets the Bureau’s accounting deadlines by mailing the blank questionnaire to the regional office. The interview information can then be communicated later by telephone, transcribed onto the questionnaire, and incorporated into the survey.

Generally, the supervisors we interviewed felt that follow-up actions often will not succeed in persuading unwilling respondents to cooperate. From the follow-up efforts we observed, we believe reasonable and sufficient actions are taken to successfully complete survey interviews.

Reinterview Program

Reinterviewing, a key quality control, is primarily used to detect falsification of data and ensure actual performance of interviews. The Bureau considers falsification to be the willful recording of incorrect information. Reinterviews also assist management in ensuring that the correct household unit was surveyed and that responses were properly recorded, and provide feedback on the interviewer’s performance.

The Bureau has designed a reinterview program for CPS and SIPP in which about a sixth of the interviewers are randomly selected for reinterview each month. Each interviewer undergoes a reinterview at least once, but not more than four times, yearly. Interviewers have a third of their assignments subject to reinterview.

Reinterviews are conducted by either a supervisor or SFR, provided that the SFR did not conduct the original interview. For CPS, the entire questionnaire is covered as if it were a first-time interview. Only selected SIPP questions are used for reinterview purposes. Reinterviewers for each

survey are experienced and familiar with the survey's concepts and current procedures.

Reinterviewers have the original responses so they may determine, as much as they can, why answers to the same question differ. Reinterviewers must (1) evaluate whether the original interviewer falsified data or erred, or whether the respondent provided inconsistent answers, and (2) prepare a report, which is reviewed by management personnel for appropriate action.

The Bureau does not tolerate falsification of data. If such is indicated and confirmed, the interviewer is dismissed. Bureau figures on proven incidents indicate that data falsification is not a significant contributor in reducing data quality. For instance, the Bureau confirmed 31 such incidents in the CPS for all regions from September 1982 to August 1984. An average of approximately 57,000 interviews is completed monthly. Based primarily on discussions with the supervisors and SFRs, we believe their experience and familiarity with the interviewer's commitment to the job is a significant control in preventing data falsification. Management personnel and interviewers we questioned and observed are very sensitive to the proper conduct for survey interviews. When supervisors detect potential problems, we found they take steps immediately to more closely monitor an interviewer's performance. When the disparity in responses can be attributed to interviewer error, additional training and closer supervision may be scheduled. In all cases, interviewers are informed of the reinterview results; a copy of the report goes into their personnel file.

The Bureau also uses the reinterview program to measure data reliability for CPS, but not for SIPP. In 20 percent of CPS' reinterview cases, the interview is conducted without referring to the original responses. The reinterview results are forwarded to headquarters and compared with the original responses; analysis of the comparison between the two provides a measure of data quality. For SIPP, the reinterview program is not used to assess data quality, partly because of the cost and respondent burden necessary to reconcile original and subsequent responses.

We examined two potential problem areas of this program: (1) interviewers' possible advance knowledge of forthcoming reinterviews and (2) providing the original responses to the reinterviewer.

In the first area, we determined that safeguards are sufficient to minimize an interviewer's foreknowledge of reinterview. For one, interviewers and their assignments are selected randomly each month. For another, only supervisory staff or SFRs conduct reinterviews. The Bureau relies on their integrity not to forewarn an interviewer of a planned reinterview.

In the second area, having the original responses before reinterviewing could influence a reinterviewer's judgment or could allow for copying of the original answers and not conducting the reinterview. We found that the Bureau has no direct check on reinterviewers.

Bureau officials interviewed on this matter generally agree that, ideally, all reinterviews should be conducted as if they were original interviews, with a third party comparing the original and reinterview responses. Should the two responses differ significantly, contact with the respondent would be initiated. We were informed that the Bureau has considered, but not implemented, this procedure, due primarily to additional cost, time delays, and respondent burden. In particular, the time required to compare original to reinterview responses and have the results included in the CPS estimates published a week later is difficult. We believe that current practices are reasonable, given the cost, time, and respondent burden. Overall, we believe the reinterview program is accomplishing its primary objective of discouraging data falsification.

Data Processing Controls

Data processing differs for CPS and SIPP. CPS questionnaires are shipped daily to Jeffersonville, Indiana, for entry into headquarters' data processing system. SIPP questionnaires are edited within the regional offices, then are electronically transmitted into headquarters' data processing system.

Each regional office counts and records the number of shipped CPS questionnaires. Each receives confirmation from Jeffersonville that the correct number of questionnaires has arrived. We observed that the prescribed procedures for accounting for the transactions were carried out. For each CPS interview cycle, monthly count forms are maintained at each regional office.

Jeffersonville's activities include converting CPS questionnaire data into electronic form for transmission to headquarters' processing system. There the data enter an acceptance/pre-edit routine; certain errors detected are returned to Jeffersonville for correction. The remaining

data processing—including a final edit and tabulation of results—continues at headquarters.

In contrast to CPS, SIPP's data processing activities occur in the regional offices. The questionnaires are batched and the data are entered by data entry clerks. These entries are checked by an electronic data processing (EDP) supervisor. The clerks are expected to operate at a minimum-standard error rate of less than .043 percent. If performance falls below this level, all the clerk's data batches are completely edited. In contrast, clerks continuously maintaining a satisfactory error rate have a sixth of their batches edited. After editing, data are electronically transmitted for entry into headquarters' data processing system. Any data rejected are transmitted back to the originator for correction and resubmission. Headquarters processes data until they are acceptable.

In both regional offices we found that controls over regional-office data processing—transmission and document shipping—are adequate, in place, and operating. We observed document shipping for CPS to be in compliance with Bureau procedures. We also observed that all SIPP document batching, delivery to the EDP unit, keying, supervisory checks, and edits are conducted as required. SIPP editing policy is followed: data are not transmitted to headquarters until the appropriate control steps have been completed. Rejected data are monitored until they are acceptable to headquarters' computer system. Since headquarters verifies receipt of the SIPP data from the regional offices and performs the final edit and tabulation, these data processing controls are not within the scope of this review. Neither are data processing controls for CPS at Jeffersonville and headquarters within the scope of this review.

Other Issues Potentially Affecting Federal Statistical Data Collection

As we examined data-collection, review, and processing controls for CPS and SIPP, we became aware of other factors that impact the overall quality of federal statistical data. Specifically, the Bureau's ability to attract and retain individuals to perform interviews, the job's being part-time, respondents' willingness to provide information, and general funding levels are factors that also affect data quality.

These issues are not unique to the CPS and SIPP but are factors in most federal statistical surveys. Because of congressional concern on preserving federal statistical quality, we include a general discussion of these factors.

Attracting and Retaining Qualified Interviewers

Interviewing requires certain skills; it is employment that cannot be readily performed by some, and its nature is not attractive to others. Recruiting and retaining qualified individuals for interviewing is a challenge. Bureau officials believe that the numbers of qualified individuals who desire an interviewer's job, who find the pay acceptable, and who will work part-time are limited, increasing the challenge of attracting and retaining interviewers.

The Bureau and other research-oriented organizations have done considerable study and analysis on desirable characteristics for interviewers. Generally, the following are considered best: middle-age; need for money or job; adequate and dependable child care for interviewers with children; desire for part-time, not full-time, work; neat appearance; good personality and physical health; and interest in the job. Bureau officials have stated that it is difficult to recruit individuals who have these characteristics and who also find the work and compensation acceptable.

Over the years, the typical interviewer can be described as a white female, 42-52 years old, middle-class, married, and having a high school education or better. While these individuals have traditionally made up the interviewing staff, today the Bureau finds the "typical interviewer" pursuing full-time, long-term, career-oriented employment and not seeking or interested in part-time Bureau employment.

Potential interviewers are subjected to the same selection criteria in all regions. New York constantly recruits to fill positions for all of its surveys. For example, from July to December 1984 that office held 39 testing sessions at various metropolitan locations for 797 individuals who had applied or had expressed interest in becoming an interviewer.

Of this number, 367 showed up for testing, and 214 passed. For various reasons, such as the applicant did not meet the other selection criteria or was no longer interested, only 93 of the original 797 individuals could be hired.

Kansas City does not need to devote as much effort to recruiting as does New York. In most cases, Kansas City has been able to fill its needs through current interviewers' knowing of qualified, interested individuals.

Bureau officials, asked about the differences in effort expended to attract potential candidates, said they believe it is in part due to regional demographic and economic differences.

Retaining the interviewer staff is another challenge the Bureau faces. For the past 20 years, it has been confronted with a turnover rate of approximately 30 percent. Turnover adds costs in terms of hiring and training new interviewers and can contribute to a loss of data accuracy. New interviewers generally tend to make more errors and are not as successful at probing for the data. Decreasing turnover rates, thus retaining more-experienced interviewers, might reduce costs and improve data quality. However, having an experienced staff, who accrue service time and pay increases, may add costs. In either case, it has not been clearly established whether reducing the present 30-percent turnover rate is possible or whether this is the best rate attainable. Kansas City does not yet consider interviewer turnover to be a critical issue. New York's rate is higher than the average 30 percent. Other metropolitan-area regional offices are beginning to experience increases in turnover rates.

The Bureau has tried to isolate the reason for high interviewer turnover. In its most recent study (January 1985), former interviewers were asked why they left. Most said the nature of interviewing, performance expectations, pay, personal/family situations, lack of supervisory support, cutbacks in work, relocations, and/or retirement caused their departure.

We believe that the Bureau is limited in what it actually can do. For example, Bureau officials believe it is difficult to dramatically change the job's characteristics. Performance expectations must be high to foster and maintain data quality. Cutbacks in hours worked depend somewhat on available funding or changes in survey practices. Officials have little control or influence over relocation and retirements. Improving pay may help, but will not single-handedly reduce the rate.

The Bureau recently surveyed interviewer attitudes and is studying the results. To date, no specific recommendations have evolved. Regardless, we believe turnover impacts data quality.

Part-Time vs. Full-Time Employment

There are advantages and disadvantages to part-time and full-time employment. Offering the interviewer's position as full-time employment might attract people who want a long-term career opportunity. It is possible that if interviewers' jobs were full-time their work would be more accurate and complete or that they would be less likely to quit. No available data support or refute these hypotheses. Hiring interviewers full-time would involve going through regular federal hiring and firing channels and paying benefits—sick leave, insurance, or retirement—not currently required. Additional factors are that surveys are short in duration and a full-time work load is difficult to guarantee. The Bureau has reached no definite resolution on these issues.

Respondents' Participation

The Bureau must carefully design its survey sample population and questionnaires to provide the basis for projecting national estimates with the desired high degree of statistical validity. Certainly the estimates are heavily dependent on the respondents' taking the time to provide the data. Participation in most federal surveys is strictly voluntary.

Respondents' willingness to participate in a particular survey can be influenced by various factors, such as disagreement with certain governmental activities or policies, timing, interview frequency (monthly for CPS, once every 4 months for SIPP), or a perception that questions are too personal or none of the government's business. If certain information, say, Social Security number, has already been given, a respondent may not repeat it to the interviewer.

If the Bureau cannot maintain a certain response rate within a prescribed sample size, the resulting lack of data will undermine the survey's statistical validity; larger population surveys can tolerate more nonresponses than small-sized ones, where nonresponses become more significant.

Funding Implications

The quality of statistical data collection requires a certain level of funding to balance the degree of accuracy desired and costs to be incurred. Areas of consideration are: raising interviewers' compensatio

and benefits, expanding a survey sample to possibly get better information, performing studies of alternative collection methods, and improving statistical methodologies. Funding is a significant factor when considering whether survey methodologies and questionnaires need change or updating. The necessary research and pretesting can be costly.

Such issues are difficult to evaluate since the effect of cost on data quality may not be clear or immediate. In recent years, the Bureau, experiencing budget cuts, has reduced the CPS and SIPP survey sample size. A reduced sample size gives nonresponses greater impact on data quality. Staff may perceive the need for high response rates as increased pressure to obtain interviews. Such pressure could, in turn, contribute to turnover rates, which might increase error probability and result in less reliable estimates.

We want to emphasize that our review covered only questionnaire design, interviewer selection and training, and data collection and review procedures at two regions. While we found these essential controls for data quality effectively operating, we believe that they, in and of themselves, are not sufficient to guarantee a high degree of statistical quality. The question of statistical quality must be answered with attention to numerous factors involved in preparing for and gathering the data as well as serious examinations of the trade-offs statisticians, policy-makers, and users are willing to make.

Requests for copies of GAO reports should be sent to:

U.S. General Accounting Office
Post Office Box 6015
Gaithersburg, Maryland 20877

Telephone 202-275-6241

The first five copies of each report are free. Additional copies are \$2.00 each.

There is a 25% discount on orders for 100 or more copies mailed to a single address.

Orders must be prepaid by cash or by check or money order made out to the Superintendent of Documents.

33797

United States
General Accounting Office
Washington, D.C. 20548

Bulk Rate
Postage & Fees Paid
GAO
Permit No. G100

Official Business
Penalty for Private Use \$300
