Computer-Assisted Personal Interviewing for the Consumer Expenditure Interview Survey

LINDA SIOUX GROVES

Beginning in April 2003, the Consumer Expenditure (CE) Interview survey will be conducted by computer-assisted personal interviewing (CAPI). The survey will continue to be conducted in person by U.S. Census Bureau interviewers in the respondent's home; however, the interviewer will administer the questions and record the answers on a laptop computer in place of the current paper-and-pencil questionnaire. This article describes the process whereby the CE CAPI questionnaire was designed and developed and discusses some of the benefits expected to be realized from CAPI data collection in the areas of data quality, operational efficiency, and opportunities for future improvements.

Design and development of CAPI

The Census Bureau collects the data for the CE Survey under contract with the Bureau of Labor Statistics (BLS). The administration of the survey is very much a collaborative effort between the two agencies. Discussions and planning regarding converting the CE Interview Survey to CAPI began in 1997. The Census Bureau was already collecting several other surveys in CAPI mode at that time, including the Current Population Survey and the Health Interview Survey. However, all of the CAPI surveys being collected by the Census Bureau, as well as all the peripheral systems that support collection activities, such as the Case Management System, had been developed in a DOS computing environment. The availability of new instrument-authoring software and more powerful laptops led to an early decision that CE CAPI would be developed in a Microsoft Windows® environment. The authoring software chosen was Blaise, which was developed by Statistics Netherlands and is in wide use in Europe and in other U.S. survey organizations.

The CE CAPI development project was an interagency effort, with management representatives from both BLS and the Census Bureau serving on the CAPI Steering Group (CSG). The steering group developed the strategic plan for the project and chartered numerous working teams that were then assigned to establish instrument design standards, write specifications, program and test the CAPI instrument and related systems, develop a new Case Management System, establish new-interviewer training, and plan a large "dress rehearsal" to assess the impact of CAPI on CE estimates. The steering group approved the project plans for each of the teams, facilitated communication among teams, and monitored progress throughout the project. Among the goals of the Census Bureau were (1) to use the CE CAPI development process to set Windows® stan-

Linda Sioux Groves is Chief, Division of Consumer Expenditure Surveys, Bureau of Labor Statistics.

dards for any future CAPI development of other surveys and (2) to create a Case Management System that all of the surveys that the Census Bureau administers could use. The latter aim was important because Census Bureau interviewers typically work on many different surveys.

A great emphasis was placed on testing during the CE CAPI development process, with numerous different kinds of tests designed to accomplish various goals. During the design standards stage, instrument prototypes were developed and examined by experts from several survey organizations. Once basic design features, such as colors and fonts, were established, development of the CE instrument began. The decision was made to program the questionnaire, which consists of 22 sections roughly corresponding to different topics, in three modules.

The first module consisted of sections of the questionnaire representing as many different design issues as possible. For example, it included sections in which the interviewer reads long lists of items to the respondent, sections wherein screening questions are used to skip the respondent to the correct set of detailed questions, and sections in which reported data from the previous interview are used extensively. Functionality testing was done to ensure that the module met the specifications. Following this testing, a panel of Census Bureau CE data collectors was brought in to test the usability of the module. The results from these two rounds of testing and the decisions made on program design issues were then applied to the development of the next module, and the process was repeated until the entire data collection instrument was programmed. The early input from data collectors during the development process resulted in a more "interviewer-friendly" collection instrument.

Concurrently with the development of the CAPI instrument, the new Case Management System and postcollection processing system were developed and tested. Once all of these pieces were complete, they were integrated with the CAPI questionnaire and a systems test was performed. The test was used to make final adjustments to the CAPI system for the dress rehearsal.

The CE CAPI dress rehearsal began in January 2002 in all 12 Census Bureau regional offices. Lasting 9 months, with three quarterly interviews per respondent in a sample of about 3,000 households, the purpose of the dress rehearsal was to analyze the impact of CAPI on response rates and on expenditure estimates. A secondary purpose of the test was to make final adjustments to training and procedures by involving a larger pool of data collectors.

The CAPI system was fully implemented in April 2003. Interview cases that began their five-interview cycle with the paper questionnaire were converted to CAPI at that time. In an effort to ease the transition of cases from paper to CAPI, changes that were anticipated in the CAPI questionnaire were largely incorporated into the paper questionnaire in advance (during 2001). Thus, the content of the paper instrument and that of the CAPI instrument are nearly identical.

CAPI and data quality

The CE interview is long and complex and usually takes from 60 to 90 minutes to complete. In addition to collecting information on expenditures for a wide range of items, the survey collects detailed demographic and income data pertaining to consumer unit members, data on assets and liabilities, and descriptive information about expenditures for classification and bounding purposes. The interviewer is required to navigate correctly through numerous screening questions and on to the detailed questions, all the while skipping inapplicable questions. In some cases, the interviewer is required to carry forward information from one part of the survey to another and make decisions about which subsequent parts to administer or questions to ask, all on the basis of a complex decision table.

Reviews of the collected data reveal that, because of the complexity involved, interviewers sometimes make mistakes in administering the paper-and-pencil interview, resulting in inconsistencies or gaps in the reported data. If these errors are detected early enough, the interviewer may recontact the respondent to fill in the missing data. Otherwise, the errors must be resolved through postcollection editing.

One of the advantages of a CAPI collection instrument is that many of these types of data problems can be eliminated. The logic programmed into the instrument forces the interviewer to stay on the correct path and does not allow questions to be inadvertently skipped. For example, in the section of the paper questionnaire dealing with properties owned by the respondent, interviewers ask (1) one set of questions for each new property reported, (2) different sets of questions, depending on what type of mortgage the respondent has and whether there are also home equity loans on the property, and (3) yet another set of questions if the property was disposed of or the mortgage payment amount changed from what was reported in the previous interview. The CAPI instrument will ensure a more seamless flow through all of the applicable questions for each property. In addition, the instrument is able to keep track of long lists of items and ensure that the correct set of detailed questions is asked for each item. As a result, there is much less postcollection editing and error resolution with CAPI.

Another way in which CAPI will improve the quality of the data is by requiring the interviewer to verify unusually high or low values with the respondent. Range edits are programmed into the CAPI instrument, based on previously reported data. When a value outside of the allowable range for a particular item is entered, an edit message is triggered, requiring the interviewer to explicitly accept the value or change it, thus checking for typos. The interviewer is also allowed, and even encouraged, to enter textual notes to explain unusual values. An unusually high expenditure for dresses or cut flowers, for example, could be accompanied by the note "Respondent is pre-

paring for daughter's wedding." The interviewer might also use the note field to indicate uncertainty about the classification of an item. A $45,000 expenditure under "Hobbies," for instance, might be accompanied by a note "Respondent collects antique cars." Notes such as these can prove useful to analysts who examine the data later, because items with notes associated with them are flagged in the data file and the text of the notes will be stored with the data. An outlier detection system will automatically display the notes to the data reviewer.

Another feature of the CAPI instrument is that help screens will be made readily available to interviewers as they administer the questionnaire, rather than in a separate collection manual that might be difficult to consult during an interview. The CE CAPI help screens include examples, such as what to include under "small household appliances," and definitions, such as those of "PPO" and "IPA" with regard to the type of health insurance that each offers.

Finally, another way in which CAPI may improve the quality of CE data is by allowing new items to be added to the questionnaire more quickly as they enter the marketplace. This feature is highly important to one major user of CE data—the Consumer Price Index program—in terms of keeping the index as current as possible, as well as being important to CE data users in the private sector.

Operational advantages of CAPI

From a survey operations perspective, a CAPI instrument has several advantages over paper-and-pencil data collection.

Currently, interviewers send their completed CE paper questionnaires to the Census Bureau's National Processing Center in Jeffersonville, IN. There, the clerical staff checks questionnaires against a master control list, applies codes to certain items (for example, on the basis of the interviewer's description, the make and model of a vehicle are coded), and keys in and verifies the data. With the implementation of CAPI, the data are input directly into the computer with no separate keying-in step. Coding is done as the data are entered. (In the case of a vehicle's make and model, the interviewer will select the correct description from an alphabetical popup list.) Instead of physically sending paper questionnaires to a central location for processing, the data are transmitted nightly from the interviewer's home via a modem. Consequently, CAPI data collection should make the data available for tabulation sooner.

Other survey operations also will be streamlined by the conversion to CAPI. Currently, at the National Processing Center, clerical staff transcribes certain information from each completed paper questionnaire onto the next quarter's blank questionnaire and mails both back to the Census Bureau regional office, which, in turn, mails them out to the appropriate interviewer in time for the next collection cycle. The transcribed data include inventoried items, which the interviewer does not recollect each time, but rather updates with current information (for example, on properties owned), as well as expenditure data collected in the previous period and now used for bounding in the current interview. These bounding procedures minimize telescoping errors that are common in retrospective interviews and result from a tendency to report past events in the reference period of the survey. With CAPI data collection, once these data are captured in electronic form and then transmitted, an input file is created for the next quarter's interview and is transmitted directly to the interviewer's laptop.

Certain survey management and control functions will also improve under CAPI. Field supervisors can easily reassign cases to a different interviewer, if needed, simply by retransmitting information about the case. Supervisors in the field, as well as headquarters staff, can get much more timely reports on the status of data collection activities than they could using paper questionnaires.

Future improvements

Respondent burden is a significant issue for the CE Interview survey, likely contributing to underreporting of expenditures and to refusals by respondents to participate in later waves of the survey. Unfortunately, CAPI will probably not make the interview any less burdensome to the respondent, and early indications are that the interview may even take slightly longer.

However, future research might permit CAPI's capabilities to be used to streamline the interview and reduce respondent burden. More customization of the interview could be possible, based, for example, on respondents' characteristics or previously reported data. Also, the added flexibility of CAPI might allow more experimentation with global questions and randomization of topics, so that not all parts of the questionnaire would need to be asked during each wave of the survey.

CAPI will certainly afford survey researchers much more quantitative information about the interview process itself. For example, each CAPI interview produces an audit trail that allows one to "replay" the interview. This can be used to diagnose trouble spots in the interview, detect whether the interviewer jumped around in the instrument or followed the default path, and count how many times help screens were invoked or warning messages were suppressed. Similarly, timing data from the CAPI instrument can be used not only to measure overall interview length, but also to access how revisions to questions affect timing in individual sections. These are valuable tools in the CAPI instrument that are not available in a paper interview. Through them, investigators can gain a much better understanding of some of the difficulties facing interviewers, and that increased understanding will lead to further improvements in the data collection process. ■