Research highlights of the Consumer Expenditure Survey redesign

Research is underway to help determine the best ways to improve the expenditure data generated by the Consumer Expenditure (CE) Survey, an ongoing survey that collects information on spending, income, and household characteristics. To determine how to optimally redesign the CE Survey, the Gemini Project was begun in 2009 with the goal of reducing measurement error, particularly error generally associated with underreporting, and with a secondary goal of halting or reversing the decline in response rates while also managing operational costs. The motivation, challenges, and independent expert recommendations for the redesign initiative are discussed in the article, along with data quality, respondent burden, and cost considerations associated with various potential CE Survey redesign features. The paper also discusses that an evaluation of the effectiveness of a survey redesign requires that a monitoring system identify the impact of each survey stage on measurement error and that it measure respondents’ perceived burden. The article concludes by pointing out how current research results inform and align with the next steps planned for the CE Survey redesign process.

This article highlights recent research conducted in support of the ongoing, multiyear redesign of the Consumer Expenditure (CE) Survey. In the first section, the CE research program is introduced. In the second section, the motivation, challenges, and accomplishments of the redesign initiative are discussed, along with next steps. Major research findings and implications are described in the third section. The article concludes with a discussion of the next steps in the CE survey redesign.
CE program overview

The Consumer Expenditure (CE) Survey is a federal government survey that collects information on a broad range of expenditures, income, and household characteristics. This information, which presents a statistical picture of consumer spending, is used by the Bureau of Labor Statistics (BLS) in computing the Consumer Price Index, as well as by economic policymakers, business, academia, and various federal agencies. Because of the important role of the CE Survey as the provider of this unique information, BLS continuously undertakes research into how best to collect high quality data.

The current survey consists of two component household surveys—the Consumer Expenditure Quarterly Interview (CEQ) and the Consumer Expenditure Diary (CED). The CEQ is carried out through five quarterly interviews, with data from the last four being used to produce official survey estimates. The CEQ data are collected primarily through face-to-face interviews, with each interview lasting approximately 1 hour. Respondents are asked to provide expenditure information for the prior 3 months. The overall response rate for the CEQ was 71 percent in 2011.

The CED component is composed of two independent diaries that each are used to collect 1 week of household expenditures. On consecutive weeks, interviewers provide diaries to the households and return one or two times to monitor data entry and retrieve completed forms. Respondents are instructed to record all household expenditures. The response rate for the CED also was 71 percent in 2011.

Over the past several years, the Branch of Research and Program Development (BRPD) of the Bureau of Labor Statistics has overseen research to assist in the revision of the CE Survey. Past improvements in the design of the CE survey have included the introduction of computer-assisted personal interviewing in 2003 and, in 2005, a revised diary form and an instrument to track interviewer contacts. Also, to keep pace with changes in both society and consumer products, the CEQ undergoes biennial updates. The BRPD is currently investigating ways to improve the quality of expenditure data while considering the possible impacts of these changes on survey budgets and respondent burden; the CE program defines respondent burden as encompassing not only the time needed to complete the survey but also the cognitive effort that the response process entails. Research projects have focused on ways to reduce the survey length and burden, improve respondent reporting behavior, and minimize or manage current survey costs. This research supports the Gemini Project, a multiyear initiative to redesign the CE Survey.

Redesign initiative

The CE program initiated the Gemini Project in early 2009. The mission of the project is to improve expenditure estimates in the CE Survey by reducing measurement error, particularly error generally associated with underreporting. Underreporting in the CE Survey occurs when respondents fail to report expenditures that they were instructed to report upon. A secondary goal of the redesign is to halt or reverse the decline in response rates while maintaining current production costs. The redesign project is motivated by concerns about both the quality of the current survey and the need to adapt to the changing circumstances in which the survey operates. In terms of data quality, evidence of measurement error has provided impetus for a new design. The ratio of aggregate expenditures measured by the CE Survey to the Personal Consumption Expenditures data from the Bureau of Economic Analysis has been declining, suggesting the presence of possible CE Survey underreporting.
Examination of subgroup expenditures has revealed differences in amounts reported based on the mode of interview and whether records were used. Several other elements of the survey—complexity, potential conditioning in respondent behavior, proxy reporting, interview length, and recall error—also present obstacles to the accurate reporting of expenditures. Changes in the social and technological environment complicate the task of reducing measurement error, as these changes are increasingly important mechanisms through which measurement error may be introduced into the survey. For example, purchases made online or recurring bill payments made by automatic debit may be less salient to respondents. Flexibility in the CE Survey’s ability to modify data collection strategies to incorporate new questionnaire designs and multimode collection better positions the program to respond to such changes over the long term.

In terms of response rates, the CE program has noted a gradual decline in response to both the CEQ and CED in recent years. In 2000, the response rate was 81 percent for the CEQ and 77 percent for the CED (compared with 71 percent for both the CEQ and CED in 2011, as noted earlier). As a result, CE program staff are developing, testing, and evaluating design changes with the goals of improving overall data quality, increasing the analytic value of the data to users, and supporting greater operational flexibility to respond to changes in the data collection environment.

The changes being pursued through the Gemini Project will ensure that the CE Survey satisfies its primary purpose: maintaining the integrity of the expenditure weights used in the Consumer Price Index (CPI). Consumer expenditure data supplied by the CE Survey are a critical component of the CPI, as they are used to estimate weights for the CPI’s consumer goods and services classification structure used in the calculation of the CPI. In the construction of the CPI, four distinct functional uses of CE Survey data are made: (1) to estimate biennial expenditure weights, (2) to estimate monthly expenditure weights, (3) to calculate the probability that an item’s price will be included in the CPI calculations, and (4) to allocate expenditure estimates between more-broadly defined expenditure categories from other survey sources. Improved data quality also enhances the usefulness of the CE Survey data to other major data users, both public and private, in addressing topics such as consumption/inequality analyses and the measurement of poverty. Therefore, data quality improvements refer to aggregate estimates of total and mean expenditures as well as to the distribution of expenditures at the microdata level. Finally, increasing the flexibility of survey operations allows the program to meet new challenges associated with data quality in a timely fashion.

Challenges. The process of redesigning the CE Survey poses a number of challenges. Defining survey requirements is a complex task, particularly given the CE Survey’s diverse user community. Identifying the needs of varied users and then reconciling competing interests is likely to leave some users with unmet needs. Additionally, the process of gathering, responding to, and acting on stakeholder input must be integrated into the redesign process such that progress on the project is maintained. Ultimately, however, the redesign process is constrained by two high-level factors: (1) the final survey design must produce the estimates required by the CPI and other major data users, and (2) long-term operational survey costs must not exceed specified budget levels.

History. Numerous tasks in the last 4 years on behalf of the Gemini Project’s mission have been completed. Beginning in 2009 and continuing through 2010, project planning involved the establishment of a survey research database and a multidimensional data quality definition to guide survey monitoring. In 2010, the CE program signed a contract with the Committee on National Statistics (CNSTAT) to convene an expert panel for the purpose
of producing a set of CE Survey redesign recommendations. That same year, efforts were made to identify the needs and priorities of CE Survey stakeholders through a forum convening a broad range of data users. The following year, panels and workshops were held to examine current practices, new technologies, and other survey methodologies that could inform the redesign process.

In 2012, the CE program started evaluating the design options and available research in order to develop a design proposal. This effort was informed by the release of the final report of CNSTAT’s expert panel. The CNSTAT report reflected high-level thinking and insights from nationally recognized experts in economics, statistics, and survey methodology on factors affecting the quality of CE Survey data and included options for improving the way the data are collected. The report concurred with the CE program on issues affecting the survey, providing broad recommendations as well as specific design proposals and cost estimates. The report also served as an important advocate for additional resources required to improve the CE Survey. The report outlined promising design features, including a one-sample design, flexible recall periods and interview structure, increased use of technology to encourage real-time reporting, increased reliance on self-administration, increased use of records, reduced proxy reporting, mixed-mode data collection, and large respondent incentives.

Current status and timeline of major milestones. The Gemini Project is currently moving from the investigative stage to the decisionmaking stage. The Gemini design team, tasked with recommending a design for the new CE Survey, developed a full redesign proposal in the spring of 2013. Currently, feedback is being solicited from BLS stakeholders. The proposal will be revised and comments from outside data users on the potential impact of the new design will be solicited and evaluated. Concurrently, a roadmap for transitioning the survey to its new design will be developed. The roadmap will be completed by 2014, and then a 5-year testing and evaluation period will commence.

Research highlights

Although the CE survey has not been fully redesigned since the late 1970s, the CE program has a longstanding history of conducting research that addresses challenges faced by the survey. From experimenting with recall periods in the early 1970s to large scale field testing in the 1990s to developing a user-friendly diary, the CE Survey has sought to base key design decisions on research. A key component of the Gemini Project’s redesign objectives has been evaluation of the features of various designs. This section of the article identifies some of the design decisions that have been researched and discusses the findings and the implications of the work, such as their impact on respondent burden or cost savings. The appendix provides a summary of the design options and associated findings that are described below.

Reduce the number of interviews. One option investigated for its potential to reduce measurement error and respondent burden in the CEQ is to reduce the number of interviews from the five interviews currently conducted. One concern with the present design is that it may result in panel conditioning, whereby survey respondents’ participation in multiple interviews changes their actual behaviors or those they report. Research into the presence of panel conditioning in the CEQ has suggested that although conditioning may be present in some categories, overall there is limited evidence that panel conditioning is a source of measurement error in later rounds of interviewing. Jennifer Shields and Nhien To examined expenditures in the trips and vacations section of the CEQ for respondents participating in all five interviews between April 2001 and March 2002. The researchers found evidence of curtailed reporting across interviews, thereby suggesting the presence of panel conditioning within this
expenditure category. Ting Yan and Kennon Copeland studied changes in mean expenditure amounts and the number of reported expenditures across all interview sections for interviews conducted during the April–June 2008 period. Comparing the expenditure reports of respondents completing later interviews against reports from earlier interviews, they did not find a statistically significant decrease in either the amount or the number of expenditures. An examination of reported expenditures by the size of expenditures and by different respondent subgroups also did not reveal a decline in reports for respondents who completed later interviews. Therefore, while respondents may be burdened by participating in five separate interviews, conducting fewer than five interviews with each household is not seen as a redesign option that would reduce measurement error.

Although there is only limited evidence that reducing the number of survey waves will improve data quality, a reduction in the number of interviews conducted would be expected to cut costs. The CEQ currently uses the first interview wave to collect both roster information and an inventory of housing characteristics and goods as well as to serve as a bounding interview to prevent telescoping (for example, respondents reporting purchases in the second interview that had occurred prior to that interview’s reference period). Studies comparing reporting levels from households that received a bounding interview with those that did not found no statistically significant differences between the two types of households, thus limiting the utility of the first interview as a bounding interview. Ian Elkin looked at CEQ timing and data collection costs to calculate the effect of transferring roster and inventory questions from wave one into wave two. He found that while shortening the wave-one bounding interview resulted in only marginal cost savings, much larger savings were associated with eliminating the bounding interview entirely. He also reported that the addition of roster and inventory questions to wave two could be accomplished with a manageable, incremental increase in interview length. This, combined with evidence that the National Crime Victimization Survey did not report any adverse effects on data quality after adjustment for its elimination of the bounding interview, has led to the recommendation to eliminate the wave-one CEQ interview.

Reduce the interview reference period. The CE program has researched the impact of having respondents report their expenditures monthly as opposed to quarterly. A shorter reference period is valued by survey methodologists: memory studies consistently demonstrate that recent events are recalled more accurately than events occurring further in the past and that memory decay increases with longer recall periods. Further, shortening the reference period may reduce the cognitive burden of recalling expenditures over the reference period; a shorter reference period reduces the overall respondent burden associated with completing the interview. Past research suggests that reducing the CEQ to a monthly instead of a quarterly interval may improve reporting; however, this change may actually result in greater respondent burden because the interviews would take place more frequently to yield 12 months of data from each household.

Reduce the interview length. Research has been undertaken to identify the role of interview length in measurement error and respondent burden. The CEQ currently averages approximately an hour to complete, which has led to concerns about the impact of survey length on data quality. An experiment with changing the order in which sections were administered within the first interview did not conclusively prove that reports early in the interview provide higher quality data than do later reports. Studies on the use of a split-questionnaire design to reduce the length of the interview have shown promise. Split-questionnaire designs involve splitting the full interview into subsections and administering only select sections to each respondent. It was shown that when responses from the first wave of interviews were used to predict whether a respondent made purchases in certain expenditure categories in wave two, fewer questions would need to be
administered in the second wave. This technique would reduce the total interview time for expenditure sections by 69 percent, with minimal impact on the precision of the estimates for many expenditure categories. Using split-questionnaire designs that are responsive to respondent information collected earlier may also improve the data quality of expenditure reports.

Because interview costs are dependent on the amount of time interviewers spend obtaining a completed survey, research was undertaken to identify how reducing the interview sections would cut back on the time needed to conduct an interview, thereby lowering overall wave-one CEQ costs. The research determined that the reduction in time depended heavily on the section from which questions are removed. A large portion of the costs associated with wave-one interviews apparently are incurred in the process of contacting respondents to participate. Therefore, reducing the length of the wave-one interview by piecemeal reductions in content would result in only minimal cost savings.

In addition, studies have investigated administering “global questions,” which are questions asked at a more aggregated level, as this could reduce the burden imposed upon respondents. One study examined whether data quality was affected by asking global questions versus asking a series of questions at a more detailed level. It was assumed that higher expenditure amounts were a measure of better data quality in this comparison of question types. Across the 10 expenditure sections tested, global questions resulted in higher expenditure amounts than detailed questions, findings consistent with comparisons of CED detailed and CEQ global food-at-home expenditures. Qualitative research has shown that global questions resulted in some higher expenditure amounts, but researchers also raised questions about the accuracy of the responses for these types of questions. These findings suggest that caution is called for when using global questions in place of more detailed questions or when attempting to attribute better data quality to the administration of global questions.

Reduce proxy reporting. Research has identified proxy reporting as a source of the underreporting of expenditures. A study of a computer-administered survey found that, across four expenditure categories, approximately 60 percent of the reports by the target person were in agreement with the proxy respondent’s report that an expenditure had been made. The target person’s reports were taken to be the “truth” measure. The underreporting of other household members’ expenditures was accompanied by underreporting of their expenditure amounts. Furthermore, underreporting of amounts was more pronounced among proxy respondents who were not related to the target person. In light of this finding, the CE program has explored providing a diary to each individual within a household. A National Opinion Research Council (NORC) review of relevant literature found that other surveys have successfully used individual diaries to augment expenditure reports, and this finding encouraged research in this area. The CE program has conducted two feasibility studies testing this approach. One found that, while response rates were lower in study households that were given individual paper diaries, both the number and the amount of expenditures reported increased. This study also suggested that the greater number of visits that would likely be necessary to collect diaries from all household members could cause an increase in field costs. Another study had similar findings.

Increase the use of records and minimize respondent recall burden. Reducing the number and length of interviews and reducing the reference period used in interviews are redesign options that have been explored for their potential to improve data quality and reduce respondent burden. Research has pointed to the increase in reporting that occurs when respondents make the effort to consult records of expenditures. A model of determinants of reported expenditure amounts found that respondents’ use of records (as well as use of the Information Booklet, a
respondent interview aid) heavily influenced quality differences between in-person and telephone interviews. However, anecdotal data provided by field interviewers discussing face-to-face interviews indicate that, because of insufficient motivation, respondents rely mainly on recall instead of records and also do not use the Information Booklet. Researchers have found that respondents may have trouble using records because records for prior months are unavailable, record detail is insufficient, or billing periods do not match reporting periods. In addition to records not being available, respondents may not take the time to locate records. A study which requested that households collect records for the purchases they had reported found that households only produced records for 36 percent of their reported expenditure items. Additional studies continue to explore the feasibility of increasing the use of respondent records in the CEQ, but initial conclusions suggest that although the use of respondent records can improve data quality, any system relying on using records to answer CEQ questions needs to be flexible enough to accommodate the variety of records that might be provided. Additionally, the time required to use records to complete an interview needs to be carefully monitored to ensure that the benefit to data quality is not outweighed by an increase in respondent burden.

*Increase the use of new technologies.* Another way to reduce measurement error is to facilitate the reporting of respondent expenditures by incorporating new technologies into the data collection process. According to CE program research, 10 percent of respondents report using financial software at home. Encouraging interview respondents to consult these records could result in information being retrieved from a source that tends to be more reliable than memory. In addition, with more and more Americans using smartphones on a daily basis and with data scanning and processing technologies advancing rapidly, the CE program has explored options for integrating these technologies into the surveys. The use of website or smartphone platforms for entering expenditure amounts is being studied as a way to provide respondents with greater flexibility in tracking expenditures. Receipt and bar code scanners have been reviewed, and while they can be time consuming and bulky to use, their ability to capture some of the information required by the survey without the respondent having to answer survey questions gives them great appeal. Unfortunately, software and technology that can reliably capture and record information remains limited. Furthermore, having people code scanned records is a time-consuming process, estimated to be 8 hours for 2 weeks’ worth of records from a single household. Use of bar code scanning has been found to be most feasible if done through a respondent’s own app or device instead of a dedicated bar code scanner, and bar code scanners do not currently collect price information, something that respondents would need to provide separately. Methods to ensure that respondents’ digital and online records are captured securely would also need to be developed.

*Incorporate multimode interviewing.* In addition to incorporating new technologies to assist respondents in recording their expenses in a timelier manner, giving respondents more options for providing data may reduce measurement error. Although the CEQ was designed as a face-to-face interview survey, interviewers choose to conduct more than one-third of CEQ interviews over the telephone. The findings of the model of determinants of reported expenditure amounts referenced earlier, as well as recognition that telephone interviews are less expensive to conduct and may be more appealing to some respondents, led the CE program to permit telephone interviews for data collection. In contrast, the CED is a pen-and-paper survey by design, with respondents recording their daily expenses in a paper diary. The CE program is exploring use of an Internet diary, which would allow expenditures to be entered from any computer as an alternative to filling out the paper diary.
Measuring redesign effectiveness

The goal of the CE redesign is a reduction of measurement error. Each of the design decisions has been or will be made on the basis of evidence that data collected with that design would have a lower level of measurement error than the current CEQ. Measurement error is difficult to quantify, however, so data quality and respondent burden will be evaluated to determine the success of the redesign.

Data quality. To assess the effectiveness of redesigned versions of the CE Survey, it is necessary to have a monitoring system to measure data quality. This system would produce metrics from the various survey processes that indicate the effectiveness of these processes before and after the redesign. Scott Fricker and Lucilla Tan have created a framework for developing and maintaining this system, which involves identifying major activities that support the CE Survey, potential issues with each stage of interviewing, and specific ways to monitor these issues. The proposed monitoring system would cover all stages of the survey life cycle, because each stage contributes to the quality of the final survey products. The comprehensive nature of the proposed monitoring system raises the question of whether the system’s metrics should mirror the CE program’s focus on underreporting of incurred expenses—that is, respondent failure to include expenditures that they were instructed to report—or on underestimates—that is, identifying why incurred expenditures are reported at a lower-than-actual cost. Whether identifying sources of underreports, underestimates, or both, having a system in place will be critical for tracking the effect of the redesign on data quality. Beyond monitoring the redesign’s effectiveness, this system can also be used to provide information for a continuous improvement process in the current survey, regardless of the scale of the survey redesign or when it is implemented.

Carrying out continuous quality improvement efforts on the CE Survey is an important goal independent of the survey redesign. These efforts necessitate having a full understanding of the various survey error sources and their relative impacts on the current survey. For this reason, the CE program is conducting research specifically focused on measurement error. This investigation involves multiple phases. The first phase will result in a report, based on past studies, about the current state of knowledge of measurement error in the CE. This phase will also result in a proposal to address gaps in understanding CE measurement error, as well as addressing how to monitor and evaluate this source of error. Subsequent phases of this investigation will involve evaluating this proposal, then developing and testing selected recommendations.

Respondent burden. Another goal of the CE survey redesign monitoring is the reduction of respondents’ perceived burden. Although the Gemini Project is primarily focused on reducing measurement error, respondent burden can impact how completely respondents participate in the survey and how accurate their responses are. The CE program is developing a summary index of perceived respondent burden on the basis of post-survey questions. This index differentiates between respondents who had given an initial survey refusal and those who had not, and between respondents who had completed all five interviews and those who did not. As with the monitoring framework, this index can be used to proactively evaluate redesign options for their impact on burden. The index can also be used to study how perceived respondent burden relates to other measures of survey quality. Use of these monitoring tools can guide the redesign toward a more accurate and less burdensome CE Survey.

Next steps
The CE program has presented the final design proposal to users and key stakeholders, whose feedback may result in additional modifications. The CE research program is now developing a research agenda that combines small scale studies and larger field experiments oriented around the testing and development of the new design through 2018, with training and implementation completed by 2023. The goal of these studies is to acquire knowledge to inform design and implementation decisions so that the Gemini Project will result in a CE Survey that collects higher quality expenditure data.

Appendix. Summary of major findings by design option and redesign objective

<table>
<thead>
<tr>
<th>Design option</th>
<th>Impact on data quality</th>
<th>Impact on respondent burden</th>
<th>Impact on cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduce the number of interviews</td>
<td>Inconclusive. One study found evidence of panel conditioning within one section, but another study suggested no deterioration in overall data quality by having multiple interview waves.</td>
<td>No study</td>
<td>Cost reduction. A study suggested that eliminating the bounding interview and shifting questions to later waves would reduce overall CE cost.</td>
</tr>
<tr>
<td>Reduce the length of interview</td>
<td>Mixed. A study that switched the order of sections found earlier administration of a section increased item reports but not expenditure amounts (with no measure of “true” expenditure amounts available).</td>
<td>No study</td>
<td>Minimal cost reduction. One study (using wave 1) found removal of questions would lead to a minor decrease in overall costs (which are mostly driven by the cost of making contact).</td>
</tr>
<tr>
<td>Increase the use of global questions</td>
<td>Mixed. Comparison of global and detailed questions found global questions resulted in some higher expenditure amounts (but no indication of true amounts are available). Another study suggested items were missed with global questions.</td>
<td>No study</td>
<td>No study</td>
</tr>
<tr>
<td>Reduce the interview reference period</td>
<td>Inconclusive. Study comparing monthly versus quarterly reporting periods found a nonsignificant increase in expenditures reported with monthly reporting (with no “truth” measure available).</td>
<td>Mixed. Respondent feedback to quarterly and monthly reporting indicated more respondents with monthly reporting said questions were easy to answer, but overall survey still was perceived as burdensome and too long.</td>
<td>No study</td>
</tr>
<tr>
<td>Reduce proxy reporting</td>
<td>Beneficial. Study indicated only 60 percent of categories are correctly reported by proxy. Individual diary placement resulted in a significant increase in amount and number of expenditures reported.</td>
<td>Harmful. Most studies have indicated lower response rates resulting from individual diary placement, which suggests a perception of increased burden.</td>
<td>Cost increase. Study found that placement costs for individual diaries might increase substantially because of a greater number of contact visits associated with this design.</td>
</tr>
<tr>
<td>Increase the use of records to decrease recall problems</td>
<td>Beneficial. Study has indicated that the use of records is more important than the mode of interview in determining expenditure amounts (with no “truth” measure available).</td>
<td>No study</td>
<td>No study</td>
</tr>
<tr>
<td>Increase the use of new technologies</td>
<td>No study</td>
<td>Mixed. Research suggests the use of technology may reduce reporting burden, but some types of technology may increase burden (e.g., the time required to scan receipts).</td>
<td>No study</td>
</tr>
<tr>
<td>Provide flexibility in mode of interviewing</td>
<td>Mixed. Study indicated that the use of records is more important than the mode of interview in expenditure amounts, though the use of records is affected by mode.</td>
<td>No study</td>
<td>Cost reduction. Telephone interviewing is less expensive than in-person interviewing.</td>
</tr>
</tbody>
</table>


Related Articles


Related Subjects

- Survey methods
- BLS Programs and surveys
- Consumer price index
- Consumer expenditures
- Statistical programs and methods
- Prices
- Survey procedures