

Global Questions in the Consumer Expenditure Surveys Program

1. Background Statement

Averaging 65 minutes, the length of the CEQ is thought to be one of the primary factors related to measurement error, respondent burden and response rates. The length is a result of both the wide range of expenditure categories and the number of specific questions included in each category. For example, within the clothing section there are 17 questions about specific types of clothing (e.g., pants, dresses, footwear), and as many as 6 detailed questions for each reported purchase (e.g., the age and gender of the person for whom the item was purchased, and whether sales tax was included).

When considering alternate CEQ designs, a common suggestion is to ‘break up the interview’ by asking each respondent only a subset of the expenditure categories. This option is limited by the requirement from CE data users, documented at the 2010 Data Users Forum, for the CE to produce total expenditure estimates for each household. One way to meet this user requirement while reducing questionnaire length is to replace some of the detailed questions with broader questions to capture total expenditures for a given expenditure category. These questions, often referred to as global questions, would ask about an expenditure category from a cumulative standpoint, rather than as a detailed series of questions about specific expenditure types within the larger category. For example, rather than asking individual questions about pants, dresses and footwear, a global question would ask for the total amount spent on clothing.

Although there are several ways that global questions could be incorporated into the CEQ, the most promising option is to use a global question to summarize expenditure categories or logical subsets of the category (e.g. clothing, footwear). We will assume for this discussion that global questions would be used within a redesigned CE that meets CPI and other data user requirements.

2. Relevant Work

Internal Research

Global questions have been a topic of interest within the CE for many years. Though there has not been any large-scale testing to evaluate the efficiency of globals, several smaller scale projects have explored the impact of globals on reporting rates and interview length. The most substantial effort included research on the cognitive processes respondents use to answer global questions (RTI, 2000). This work concluded that there are several complex cognitive steps respondents must perform to respond to global questions: decomposing the global category into subcategories, searching memory for and retrieving expenditures, determining whether the expenditures fit the specified time frame, and then mentally summing the expenditure amounts for each subcategory. RTI recommended a heavy reliance on the CEQ information book, which provides respondents with lists of the subcategories in each expenditure category, and having interviewers help respondents with their mathematical calculations. These suggestions imply that without aid, respondents may not be able to arrive at accurate responses in response to global questions on their own. While the RTI research noted cognitive difficulties experienced by some respondents in computing responses to global questions, they did not attempt to validate the accuracy of those responses, nor did they collect any data outside of the laboratory environment.

Although the CEQ does not currently include global questions, there are several expense pattern questions (i.e. “What has been your usual weekly expense on groceries?”) which can provide useful information on the topic of global questions. Since the CEQ expense pattern items ask about expenses also collected in the CED (food at home, food away from home), we can compare responses with the detailed CED data. Weber (2002) and Henderson (2009) compared the CEQ expense pattern questions and associated detailed CED data and found that the CEQ expense pattern questions were consistently higher than the detailed CED data.¹ Despite conceptual differences in the between the surveys (e.g. expense patterns versus individual expenditures), the differences raise significant concerns about the quality of data collected using global questions.

¹ In both the CEQ and CED interview portions, the question about food for home consumption is followed by a question asking about how much of the reported grocery expense was for nonfood items. A possible explanation for the CEQ-CED discrepancy is that respondents remember the total purchase but never "encode" the amount of that purchase going toward non-food items.

Battistin (2003) cites the methodological differences of asking detailed, rather than global, questions as one possible explanation for discrepancies between the CEQ and CED surveys in reported expenditures on non-durables (generally defined as items used once or with a lifespan of less than three years). While the CED collects detailed data about non-durable expenditures such as tobacco products and food and nonalcoholic beverages consumed at home, and then sums these up to get total spending, the CEQ asks about "usual expenses" for the last quarter, a global retrospective question about total spending on these items. When generating age profiles of log non-durable expenditures, Battistin found that the profile obtained using data from the CED detailed questions is always below the profile obtained using data from the corresponding CEQ global question. In other words, the expenditures reported using detailed questions are always below those obtained using a global question, paralleling the results reported at BLS (Henderson 2009; Weber 2002). However, it should be noted that in the CEQ, global responses to Food Away from Home have usually been lower than the specific responses from the Diary.

Recent research by Edgar (2009) reinforces the concerns about the quality of data collected from CEQ questions. The author found that respondents used generalizations, inferences from a single episode, or even guesses, when providing their answers, providing reason to suspect substantial, and possibly systematic, measurement error. This research lends support to the conclusion that the expenditure patterns (and perhaps globals by extension) are not a reliable method for obtaining accurate expenditure information.

In 2007, a small telephone study asked respondents one of two sets of global questions about clothing purchases, and followed those questions with the detailed items currently used in the CEQ (Goldenberg & Steinberg, forthcoming). One global question set asked about the entire clothing category in three parts (questions on clothing, on footwear, and on accessories) and the other divided the category into eight topics (e.g., outerwear, business clothing, athletic wear, footwear, etc.). Responses to both sets of global questions were consistently higher than the detailed questions, but the difference was only statistically significant for the eight-part global questions. Although this research cannot indicate whether the global questions are over-estimates or the detailed items are underestimates of the true values, the results do support concern about the quality of data produced by global questions, especially when considered in the context of using them in replacement of, or even conjunction with, detailed questions.

One application of global questions is as filter question which ask respondents about a broad category of items to determine eligibility for the items that follow it. Kreuter (2009)

explored the accuracy of filter questions when presented in different formats and concluded that the format of the filter question impacted the measurement of items, with some formats leading to relatively inaccurate responses. The CEQ currently includes screener questions in some expenditure categories. Safir and Goldenberg (2008) found that there were significant differences in expenditure reporting between telephone and face-to-face respondents for filter questions, leading to further concerns about accuracy. These findings about filter questions may generalize to the broader category of global questions, suggesting that we need to think carefully about not only the use of global questions, but also how we present them to respondents.

External Research

The literature presents a mixed picture of the advantages and disadvantages of using global questions over more detailed questions. In general, studies show global questions to result in less reporting, but not necessarily less accurate responses than more detailed questions. For example, from a reporting accuracy standpoint, Belli et al. (2000) examined behavioral frequency reports, and noted that detailed questions may be more susceptible than global questions to double-counting behaviors. Supporting this broad conclusion, Beatty and Maitland (2008) observed a general trend for global questions to be associated with lower frequency reports than their detailed counterparts, but not in every case, leading the authors to conclude that detailed questions may be preferable under some, but not all, circumstances. They did not provide guidance on the criteria that should be considered when determining when to use global questions versus detailed ones. Additionally, the finding of less accurate reporting (i.e., due to more double counting) in response to detailed questions is not consistently seen across studies, or even across different interview topics. While Menon (1997) finds that for regular behaviors, detailed questions may interfere with the response process, the author also cites evidence that detailed questions have a positive effect for reports of irregular behaviors.

The Panel Study of Income Dynamics (PSID) asks global questions on several topics also collected in the CEQ. Charles et al. (2006) compared PSID results with the more detailed items in the CEQ and found that some categories of expenditures (total food, health care, education, child care) were higher than CEQ reports, with food away from home much higher than CEQ. It should be noted, however, that total food and food away from home are not very detailed items in the CEQ, so they aren't particularly good cases for assessing differences between global and

detailed questions. On the other hand, some PSID questions yielded reports that were lower (transportation) or the same (housing) as the CEQ.

When looking at assessments of activity survey, Sternfeld et al. (2000) note relatively low correlations between global questions and more detailed assessments, suggesting that respondents may use different constructs for responding to these two types of measures. A recent paper by Xu and Schwarz (2009) suggests that global memories are consistent with expectations rather than actual experiences. Additionally, although global questions have been found to require less administration time, Beatty, Fowler and Cosenza (2006) report that global questions may be more difficult than detailed questions for the interviewer to administer.

3. Key Issues

Global questions are appealing in that they may allow a redesigned CEQ which meets data users' needs by providing estimates of total expenditures while allowing respondents to receive only a subset of the interview, but the quality of the data may be questionable, particularly in comparison with those collected using the current detailed questions.

4. Discussion Points

We seek recommendations for both for research that could be conducted to explore the potential use of global questions in the CEQ and alternative designs for the CE Surveys. In the discussion of a redesign model for CE, the following questions related to global questions might be addressed:

1. How accurate are responses to global questions?
2. Does accuracy vary across expenditure categories?
3. Can global questions obtain data that are accurate and reliable enough for use by the CE?
4. Detailed questions are known to have measurement error, particularly associated with recall. Do global questions have different sources of measurement error? What issues should be considered when combining data from detailed questions and global questions to produce estimates of total expenditures?
5. What should the next steps be for research on this issue for a possible change in CE design?

6. How would you design a survey, or set of surveys, to collect detailed monthly data on a comprehensive range of expenditures? The proposed design should address global questions as well as other relevant survey design issues, while meeting the requirements specified in the CPI Requirements Document.

5. References

- Battistin, E. (2003) Errors in Survey Reports of Consumption Expenditures. Working Paper W03/07, Institute for Fiscal Studies, London.
- Beatty, P., Cosenza, C. & Fowler, J. (2006). Experiments on the Structure and Specificity of Complex Survey Questions. American Statistical Association: Proceedings of the Section on Survey Research Methods: 4034-4040.
- Beatty, P. & Maitland, A. (2008). The accuracy of Global vs. Decomposed Behavioral Frequency Questions. Presented at the 2008 American Association for Public Opinion Research Conference, New Orleans, LA.
- Belli, R. F., Schwarz, N., Singer, E., & Talarico, J. (2000). Decomposition can harm the accuracy of behavioral frequency reports. *Applied Cognitive Psychology*, 14, 295-308.
- Charles, K. K., Danziger, S., Li, G., & Schoeni, R. (2006). Studying Consumption with the Panel Study of Income Dynamics. Population Studies Center Research Report 06-590. Ann Arbor: University of Michigan.
- Edgar, J. (2009). What does 'usual' usually mean? Presented at the 2009 American Associate for Public Opinion Research Conference, Hollywood FL.
- Goldenberg, K., & Steinberg, B. (forthcoming). Analysis of 2007 JPSM practicum project. Internal report, Bureau of Labor Statistics.
- Henderson, S. (2009). Comparisons of CEQ and CED data. Internal email, September 28, 2009. Bureau of Labor Statistics.
- Kreuter F., McCulloch S., and Presser S. (2009): Filter Questions in Interleafed versus Grouped Format: Effects on Respondents and Interviewers. Paper presented at the Annual Conference of the American Association for Public Opinion Research.
- Menon, G. (1997). Are the Parts Better than the Whole? The Effects of Decompositional Questions on Judgments of Frequent Behaviors. *Journal of Marketing Research* , 32, 335-346.
- RTI (2000). CE Interview Global Questions Research: BLS Conference Presentation Report. RTI Project No. 07624.000.
- Safir, A. & Goldenberg, K. (2008). Telephone Effects in the Consumer Expenditure Quarterly Interview Survey. Internal BLS report, August 29, 2008.
- Sternfeld, B., Cauley, J., Harlow, S., Liu, G., & Lee, M. (2000). Assessment of Physical Activity with a Single Global Question in a Large, Multiethnic Sample of Midlife Women. *American Journal of Epidemiology*, 152 (7).
- Weber, W. (2002). CE Diary food-at-home "global-detailed" comparisons: Interim report. Internal BLS Report., May 31, 2002.
- Xu, J. & Schwarz, N. (2009). Do We Really Need a Reason to Indulge? *Journal of Marketing Research* 46, 25-36.