Discussion of Methodological/Cognitive Issues and the Proposed Redesigns

Richard A. Kulka, Consultant
Statistical, Survey and Social Research

Consumer Expenditure Survey Redesign Options Workshop
Committee on National Statistics
The National Academies
20F Street Conference Center
Washington, DC
October 26, 2011

Introduction

My charge as part of this workshop was to serve as one of two formal discussants, focusing in particular on how well the proposed redesigns address the key methodological and cognitive concerns expressed in various forums with regard to the Consumer Expenditure Surveys. I greatly appreciate the opportunity to review many of the papers and presentations on the methodological concerns about the current CE, to study the two proposed redesigns in some detail in advance of the workshop, and to reflect on (1) what was expected by BLS, the CNSTAT Panel and other stakeholders of this important statistical program and (2) what has been achieved.

Background: Current Workshop

First, let me provide a little background, although much of this is quite familiar to most of you, so I will be brief. As noted in the introductory remarks, this workshop builds on a very extensive ongoing program of both internal and external research, advice and review dating back to the mid-1980s. These include in recent years a series of carefully planned and coordinated workshops, forums and expert panels, under the general aegis of the "Gemini Project." All of these focused in various ways and with different emphases on the development of a *fundamental* redesign of the Consumer Expenditure Surveys, which in turn evolved as a major component of the charge and work of the CNSTAT Panel.

CNSTAT Panel Charge

In brief, the CNSTAT Panel's charge is to:

- Review the discussion, output and conclusions of this prior research, and the other workshops and forums, including especially a "data user needs forum" and a "methods workshop" previously convened by BLS.
- Convene a "household survey data producer workshop" to review the experience of other leading survey research organizations in addressing similar challenges.
- Contract with one or more experienced data collection organization to:
  - Develop papers on concrete design options,
  - Provide a critical discussion of their relative merits, and
  - Convene a workshop based on results of this work.
- Based on these workshops and its review and deliberations, produce a consensus report of findings and recommendations to BLS on what characteristics and options should be considered for a redesigned CE survey.

Contractor Charge

In turn, the Panel's charge to the contractors was to:

- Provide a balanced evaluation of the benefits, costs and risks of proposed options compared with current CE surveys.
- Provide a comprehensive proposal for a survey design and/or other data collection/capture process that collects the required data.
- Provide a clear plan and recommendations on how they would evaluate their proposed design in the CE context.

- Focus their proposed designs in particular on:
    - Measurement error,
    - Changes in the survey, consumer and social environments, and
    - Increased flexibility in modes of collecting data.

The key question here, of course, is have they indeed met this charge? Overall, I believe that each has done so, and in several instances their reports exceed these requirements by providing additional insights, ideas and even some new questions not yet fully recognized to date—many of which should prove to be of great value to the Panel and BLS as they move toward developing closure on a set of key redesign options. However, there are also several more implicit questions embedded in this charge which--while not formally part of the charge and thereby not addressed directly--also have potential value for the Panel and BLS as they review and act on the review, discussion and recommendations contained in these reports. That said, given the very brief time period available for this work, these reports are extraordinarily rich in innovative ideas and insights, including valuable advice not only on what might be done, but also appropriate cautions on things that probably should *not* be done. However, consistent with a longstanding cliché that nevertheless often proves to be true, it is very much the case here that the "devil is in the details."

Discussion Outline

To address these observations in more detail, I have organized my discussion around a serious of basic questions:

1. What are the key features of each proposed redesign? Are these well justified relative to other possible alternatives and their feasibility? Which proposed aspects of these designs (if any) merit some additional thought and clarification?

2. Are these two proposed designs fundamentally (or radically) different from the current CE surveys? And, in turn, are they fundamentally (or radically) different from one another?

3. Are there some additional insights, questions or themes emerging from these efforts that merit careful consideration or serious attention independent of the specific designs proposed by each contractor?

For the first set of questions, I will comment on each contractor report in turn, beginning with that from the Westat team, which I received and reviewed first, followed by that developed by the University of Wisconsin-Milwaukee /University of Nebraska-Lincoln/Abt SRBI team. Questions 2 and 3 are de facto more integrative and interdependent so I will discuss them after the two separate reports.

WESTAT

Key Features

Key features of the redesign proposed by the Westat team include the following:

- Maintaining separate diary and quarterly interview surveys, as in the current CE approach.
- Both, however, provide several alternative methods for reporting data, including
  - Choice of either electronic or paper reporting;
  - A strong emphasis on the use of receipts;
  - Downloading electronic data files with information on expenditures; and
  - Scanning of receipts and other paper for electronic submission.

- In addition, both the diary and quarterly interview surveys rely on regular submission of such data to a *central repository*,
    - where data are received, extracted, translated, coded and evaluated ,
    - followed by "smart" prompting based on these data, either by web survey or telephone,
    - to obtain additional information required by the CE program.
- Both also involve intensive monitoring and prompting by interviewers through multiple scheduled contacts to stimulate complete and timely reporting by respondents.
- And, both surveys are explicitly designed to significantly reduce dependence on recall and recall interviews.
- In the diary component only, all CU members aged 14 and over would self-respond by household-selected method of data collection (electronic or paper).
- Similar to the current CE survey, quarterly interviews would maintain a three-month reference period (albeit with recall interviews minimized), but with only two waves of data collection conducted 12 months apart.

Observations/Questions

An innovative, complex central data depository is a particularly attractive feature of this proposed design. However, the burden on this depository system to receive and dynamically manage multiple types of data and, in turn, provide real-time feedback and tailored queries is considerable.

Detailed descriptions provided in the report on how various types and forms of data submitted to the repository will be read and interpreted, and how appropriate follow-up queries will be formulated and generated in real time, are quite impressive. But questions about the feasibility of how all of these will work in practice, especially for the diary component (with one-week, real-time data collection periods), seems problematic.

Assuming that the detailed data requirements of the CE program are maintained, it would appear than many (if not most) of the expenditures reported will require web (or telephone) survey prompting after submission to the repository, which could well serve as a significant perceived burden and barrier to complete reporting (analogous to "panel conditioning") .

For this and some other reasons, it is not entirely clear whether *actual* or "objective" reductions in respondent burden afforded by the proposed design will be *perceived as real*. Downloading and submitting data daily in the diary survey (especially banking and retail data) may in fact be more burdensome.

Similarly, the learning curve required for effective and timely electronic reporting of the quarterly interview data may increase perceived burden with 12 months of "downtime" between waves of data collection. Twelve months between quarterly interviews may also not be optimal for other reasons (e.g., significant changes in CU composition).

While the literature and experience cited from other studies suggests that development and use of an Event History Calendar will very likely improve recall (when a recall interview is necessary) in the quarterly interview survey, its value in a one week diary seems more limited, where appropriate marker events (other than expenditures themselves) may not be all that prevalent or useful.

Other than the panel component, the essential data requirements described by the CE analysts are mostly met by the proposed design. However, the costs of the proposed design are substantially greater than the current CE surveys, thereby requiring a careful analysis of cost-benefit tradeoffs/choices and some difficult priority setting.

As noted in the report, when viewed as a single survey program, the redesigned procedures address flexibility in several ways, most notably by using a variety of different data collection/capture methods that can be tailored to each respondent's (or CU's) preferences or daily routines. The challenge under this redesign would be to develop an array of integrated procedures that take full advantage of this flexibility without making reporting tasks overly complicated or to give respondents a way or motivation to avoid completing these.

Summary

- The proposed redesign focuses heavily on:
  - Increased reporting flexibility, along with extensive use of monitoring, prompting and help, to reduce respondent burden.
  - Greater use of personal electronic data, receipts and real time recording to significantly reduce reliance on recall.
  - A sophisticated, central "data repository" platform designed to capture, manage and interpret such information in real time; to integrate and control the flow of information; and generally to hold it all together.
- Feasibility testing of all these components and systems, and at all levels, will obviously be critical. The detailed evaluation protocol and plan described, though well thought through and conceived for the purposes suggested, may not be sufficient.
- Unfortunately, some of the most attractive features of the proposed redesign make it significantly more costly.

UNIVERSITY OF WISCONSIN-MILWAUKEE/UNIVERSITY OF NEBRASKA-LINCOLN/ABT SRBI

Key Features

- A key feature of the proposed UWM consortium's redesign is a single sample that integrates both the diary and retrospective reports of expenditures, thereby reducing redundancies between the CEI and CED surveys.
- The proposed single-sample design has two components:
  - A cross-sectional component, where each selected CU participates for one month, and
  - A panel component, for which each CU participates three times (also for one month) at six month intervals.
- In both components, the design proposes to:
  - Minimize reliance on retrospective reporting,
  - Promote real-time recording of all expenditures, and
  - Emphasize self-reporting for all CU members.
- The reporting task and methods are the same for both components:
  - Maintaining a diary for a one month period in which all CU members record all purchases and expenditures.
  - Use of shared computer tablets as the primary "diary" recording device, either by (1) a web-based data collection or (b) a computer-assisted self-interview (CASI) using the same questionnaire platform, with these data then downloaded to the web.

- Capture and recording of both paper and electronic receipts.
- Use of portable diaries, cell and smartphones, and other technology (e.g., scanners, barcode readers and financial software) for real-time data capture and for use as "memory triggers."

Observations/Questions

In principle, what is proposed is a fully-integrated design that emphasizes prospective, "real-time" collection of expenditures and payments. It also takes full advantage of technology, broadly conceived, including a common platform for integrating input and information. Data are captured by a flexible array of tools, serving either as memory or recording aids, or as a means for direct data input.

In practice, however, there are some significant questions about the feasibility of this design. Some of these involve the design and integration of the technology, including their support, maintenance and training. On these issues, there is some relevant experience and "lessons learned" from recent work by Knowledge Networks, Stanford University and Abt SRBI, Nielsen and the USDA, but the requirements of the proposed redesign are significantly different and greater. Of greater uncertainty, however, is the capacity to recruit and motivate CU members to maintain a diary for a full month rather than two weeks (three separate times over and 18 month period for the panel component!), even with all of the new tools and methods to be leveraged in this effort.

The preferred requirement to use a computer tablet for reporting is both a strength and weakness of the design. For example, the greater flexibility afforded by use of the array of other technologies and methods proposed is undermined somewhat where their predominant role is relegated largely to use as "memory triggers" or aides rather than direct electronic data capture and transmission. Use and integration of these various other technology- and paper-based methods as memory aids or triggers, versus direct sources for data input, requires significantly more thought and design work.

Recognition by the team that nonresponse is a key threat to this design, and the corresponding need for a rigorous “responsive design” approach--along with some initial creative thoughts on how one might carry this out in the proposed CE context--is a major strength of the report.

Most data requirements for CE analysts are met by this proposed redesign (one notable exception being generation of quarterly panel change estimates for a CU), but it explicitly gives highest priority to improvement of data for the CPI.

Costs of the proposed design are substantially greater than current CE surveys, but since some of these (e.g., technology and development) would be amortized over multiple years, a more detailed analysis is necessary.

Summary

- A bold, creative redesign that reflects a number of the key concerns discussed in previous workshops and forums of both methodologists and users.
- The design proposes to take maximum advantage of virtually all recent and future technological and cultural changes identified in the survey and consumer environments.
- A clear focus on significantly, and perhaps dramatically, reducing our historical dependence and reliance on retrospection and recall for reporting expenditure data.
- However, in spite of its considerable appeal, there is little empirical evidence available to date to make one confident that such a design is really feasible in the CE context.

- Achieving the full degree of flexibility and impact envisioned requires considerable more refined design thinking and work.
- While the proposed evaluation needs and plan address some of these issues, it is likely that an adequate assessment of feasibility and impact would require studies significantly beyond the types and sizes of evaluations proposed.
- The real costs and specific cost estimates and drivers associated with fully developing, testing, implementing and sustaining this very substantial redesign need to be carefully evaluated and estimated.

How Different Are These Designs?

Both of these proposed redesigns are different in fundamental ways from the current CE surveys. These differences clearly reflect insights and concerns from prior research, workshops and forums sponsored by BLS. The UWM team's design is self-described as a "radical" departure from the current approach, which is certainly true. The Westat team's design is somewhat less "radical," maintaining more features of the current design, but with some important fundamental design changes as well.

And, at least on first reflection, they are also quite different from one another. However, while the differences are certainly significant and important, these are also in part somewhat illusory. The basic philosophies underlying both are to design a consumer expenditure survey that is far more prospective (less dependent on recall), flexible, and technological.

One key difference is:

- Westat's proposed use of a central data depository as a major element to support these features, versus

- The UWM team's reliance on shared computer tablets as the primary "diary" recording device and platform.

- Both possibilities, however, were discussed as options in a previous Westat report on data collection technologies for BLS. A second key difference is the UWM team's proposal to integrate the diary and interview surveys, an option also considered by the Westat team in its current report. Thus, one can certainly envision deriving various "hybrid" approaches that take advantage of the desirable features of both designs, some of which overlap in any case. That said, for future planning purposes, each of these two redesigns merits careful attention and thought in their own right, including a careful evaluation of their key features not only in terms of their meeting CES user needs as currently stated, but also potential modifications of those needs and priorities. Moving too quickly (or at all) toward one or more blended or hybrid designs could hinder that process.

Additional Insights, Questions or Themes

In their reviews of factors driving these two redesign proposals, two observations seem especially significant and worthy of more detailed attention:

- The Westat report in particular does an excellent job of demonstrating and highlighting that *over 70 percent of consumer purchases are documented in electronic format.*

- Both reports clearly show and emphasize, based on experiences across a diverse array of surveys, that the *use of records significantly improves the quality of data obtained from respondent reports.*

While each proposed design incorporates some important elements that recognize and attempt to address and leverage these factors, these observations seem dramatic and significant enough to merit more detailed attention and investigation. Let me briefly discuss each in turn.

Electronic records of consumer purchases. Encouraging respondents

1) To provide personal expenditure data by retrieving their information from these databases, and then

2) obtaining their consent to seek such information from retailers and utility companies

are certainly sound ideas as far as they go. Far bolder, however, would be for BLS to use its credibility and clout (likely to require new legislation) to obtain such information directly for all CE households (with their consent) from holders of these electronic records (as suggested by in part by the Westat report). In addition, for retail purchases, the possibility of creating a universal CE-branded "loyalty card" issued in a "public-private partnership" of BLS and retail vendors might also be explored. Even if these specific suggestions prove not to be fully feasible, it just makes sense to explore such options to the fullest extent possible.

Use of records and data quality. While both redesigns aggressively promote the use of receipts and records far beyond the current methodology used in the CE surveys, the consistently demonstrated value of using records is such a significant factor for the likely success of any future design that it merits even more direct and sustained attention.

A multi-faceted and multi-phase program of both qualitative and quantitative research based on the theories and methods developed by Don Dillman and others, for example. At their most basic level, these studies (a) systematically explore ways of highlighting the critical importance and value of participating in given survey and responding as accurately as possible, while also (b) persistently seeking to minimize the burden of doing so. Implementing such a program would seem essential to assuring the success of either of these two designs (or any other that would meet the requirements of the CE redesign effort).

Recent work on the Survey of Income and Program Participation (SIPP), described in both reports, would be a good starting point for such an effort. Of particular note was their insistence in communications with respondents that use of records was the "norm," the usual, expected part of the respondent's task. In effect, not using such records would make the response task far more difficult and less accurate. While these SIPP experiments showed that use of records resulted in both greater costs and higher nonresponse, this does not necessarily mean that their use within the CE context, after incorporating some changes and new design features to counter these limitations, would necessarily experience the same fate. For example, using respondent incentives with a major (or even predominant) focus on encouraging respondents to use records, rather than just responding to the survey per se, might well be one tool or approach to address these problems.

Conclusions

- Both proposed redesigns provide significant, detailed advice, guidance and insights to the both the CNSTAT Panel and BLS.
- Neither, however, provides a "turn-key" design package that could be implemented as proposed without significant testing, a critical examination of costs (and benefits), and (unfortunately, given the BLS charge) some substantial additional research.
- But it was probably unrealistic at the outset to assume that they would do so.

- Nevertheless, the exemplary thinking and timely work of these two teams clearly represent very significant progress in providing the elements and a concrete roadmap for how a specific, well-conceived and well-tested redesign for the CE surveys can be achieved.
- Both contractor teams are to be congratulated for their outstanding efforts and products.