Volume Title: Improving the Measurement of Consumer Expenditures

Volume Author/Editor: Christopher D. Carroll, Thomas F. Crossley, and John Sabelhaus, editors

Series: Studies in Income and Wealth, volume 74

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-12665-X, 978-0-226-12665-4

Volume URL: http://www.nber.org/books/carr11-1

Conference Date: December 2-3, 2011

Publication Date: May 2015

Chapter Title: Exploring a Balance Edit Approach in the Consumer Expenditure Quarterly Interview Survey

Chapter Author(s): Scott Fricker, Brandon Kopp, Nhien To

Chapter URL: http://www.nber.org/chapters/c12671

Chapter pages in book: (p. 347 – 364)

12

Exploring a Balance Edit Approach in the Consumer Expenditure Quarterly Interview Survey

Scott Fricker, Brandon Kopp, and Nhien To

12.1 Introduction

The US Consumer Expenditure Survey is an ongoing monthly survey conducted by the US Bureau of Labor Statistics (BLS) that provides current and continuous information on the buying habits of American consumers. The Consumer Expenditure Survey consists of two independent components: the quarterly interview (CEQ) survey and the diary (CED) survey. For the CEQ, interviewers visit sample households five times over the course of thirteen consecutive months. Each interview is conducted with a single household respondent who reports for the entire household. The first interview establishes cooperation, collects demographic information, and bounds the interview by collecting expenditure data for the previous month. This "bounding" interview is designed to limit forward telescoping, which is the process by which respondents remember and report events or purchases as taking place more recently than they actually occurred. The four remaining interviews are administered quarterly and ask about expenditures in the three-month period that just ended. In the second and fifth interview, respondents are asked additional detailed questions about household income, assets, and liabilities.

The CEQ survey presents a number of challenges for both interviewers and respondents. The interview is long, the questions detailed, and the experience can be perceived as burdensome. In part because of these challenges,

Scott Fricker is a research psychologist in the Office of Survey Methods Research of the Bureau of Labor Statistics, US Department of Labor. Brandon Kopp is a research psychologist in the Office of Survey Methods Research of the Bureau of Labor Statistics, US Department of Labor. Nhien To is an economist in the Division of Consumer Expenditure Surveys of the Bureau of Labor Statistics, US Department of Labor.

For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see http://www.nber.org/chapters/c12671.ack.

there is concern that some CEQ data are underreported (e.g., Shields and To 2005; Bosworth, Burtless, and Sabelhaus 1991). Underreporting has been variously attributed to recall error, panel conditioning, respondent fatigue, and other causes.

To combat response errors like underreporting, some expenditure surveys have used a data quality control measure known as a "balance edit" check. Early expenditure surveys conducted by the BLS (e.g., 1935–1936 Study of Consumer Purchases; the 1950 Survey of Consumer Expenditures) used a balance edit to check for consistency among reported expenditures, income, and asset and liability totals; sample units whose expenditures exceeded income by more than 10 percent were followed up and reinterviewed or removed from the sample. This balancing procedure was eliminated prior to the establishment of the Consumer Expenditure Survey in 1972 because it was deemed "judgmental" and not operationally feasible for a (precomputer-assisted) quarterly survey (Jacobs and Shipp 1993). More recent implementations of this measure have differed across surveys, but the basic process is one in which respondents are given the opportunity to review and revise their reported expenditures, income, and changes in assets and liabilities. For example, in the 2009 Survey of Household Spending (SHS) conducted by Statistics Canada, households that had expenditures that were out of balance with the reported cash flow (i.e., spending that was significantly above or below income plus net assets and liabilities) were probed to identify and reconcile possible sources of error. There is evidence from two empirical studies that the data resulting from use of a balance edit are of higher quality than those collected by the alternative methods (Brzozowski and Crossley 2011; Hurd and Rohwedder 2010; section 12.3).

12.2 Study Objectives and Design Considerations

The current CEQ uses built-in range edit checks to flag reports that exceed normative thresholds for an expenditure category and consistency edits to flag reports that are inconsistent with data in related fields. The purpose of this study was to explore the feasibility of including a balance edit check based on a household's computed total spending relative to its income, assets, and liabilities. We conducted a small-scale lab study that addressed three basic areas of inquiry: (1) the effects of a balance edit in identifying and correcting reporting errors, (2) participants' reactions to the balance edit process, and (3) the factors that impact the quality of participants' reports or their reactions to the reconciliation process.

The study is qualitative in nature and exploratory, and there were a number of design considerations that impacted its scope and analytic objectives. One consideration was the potential increase in respondent burden that might accompany a balance edit process. The current CEQ asks an extensive battery of questions that takes an average of sixty-five minutes to admin-

ister. We were concerned that incorporating balance edits at the end of the existing CEQ would lengthen an already long and burdensome interview. Moreover, for some households the number of reports respondents would need to review could be unduly large, therefore making a balance edit process impractical to implement. In addition, in this study we wanted to be able to conduct real-time balancing checks and immediately follow up with participants during the interview. This necessitated the development of an electronic instrument that would record and tally participants' expenditure, income, asset, and liability totals, and it was not feasible to do this for the full CEQ questionnaire given the project timeline and available resources.

On the basis of these considerations, we decided to test a balance edit measure using a modified CEQ in which the detailed expenditure questions were replaced with a fewer number of global items that asked participants to report their total household spending in each of the CEQ section categories (see section 12.4 for details). We acknowledge that the use of global questions and (therefore) a shorter interview are significant departures from the current CEQ procedures, that both factors may affect the nature of response errors, and that these effects have the potential to interact with the balance edit response process. Nevertheless, we view this study as a useful first step in investigating the feasibility of implementing a balance edit, and in gathering some initial information about factors that may affect its outcome (e.g., frequency of purchase, topic sensitivity, household size and participant knowledge, conceptual clarity, cognitive difficulty, etc.).

12.3 Previous Research

Two recent surveys have implemented some form of a balance edit procedure in an effort to improve the quality of expenditure estimates—the Survey of Household Spending (SHS) conducted by Statistics Canada, and the American Life Panel (ALP) maintained by RAND Labor and Population. We briefly summarize these surveys and their use of this approach below, and discuss the studies that have examined the impact of reconciliation procedures on data quality.

12.3.1 Survey of Household Spending (SHS)

The SHS is a face-to-face survey that collects household expenditure and income data for the previous calendar year. To help combat recall errors arising from the long recall period, Statistics Canada allows respondents to report expenses for smaller and more frequently purchased items (e.g., food) on a weekly or monthly basis, and encourages respondents' use of records during the interview. In addition, the SHS has implemented a balance edit check in which respondents' reported expenditures are compared against the sum of reported income and net change of assets in the household. When expenditures were more than 20 percent different from reported income/

assets, the interviewer attempted to collect additional information to bring the two into better balance (i.e., 15 percent or less). According to Statistics Canada, most of the changes respondents made to their reports during this process were to reported income and assets (personal communication, May 2011); this is likely due to the wording of the structured probes, which focused on these areas. The SHS households that remained unbalanced were deemed unusable and excluded from estimates.

In 2006, SHS data collection moved to computer-assisted personal interviewing (CAPI) and the balance edit was not used in the collection phase that year. Instead, balancing was applied at the processing phase (with no active involvement or reconciliation by respondents), and this had the unexpected consequence of significantly increasing the number of records that were deemed "out of balance" (from 546 in 2005 to 4,300 in 2006). This comprised an unacceptably large percentage of completed SHS questionnaires for 2006 (29.4 percent), so a decision was made to reinstitute a field balance edit feature for 2007 data collection. To assess the effect of the balance edit on data quality, Brzozowski and Crossley (2011) compared SHS data from 2006 (no interview-based balance edit procedure) with data from 2005 and 2007. They found no differences in income or expenditure reporting across the three years for the top fifteen income vigtiles. However, respondents in the bottom of the income distribution (lowest five vigtiles) underreported income when there was no field balance edit.

12.3.2 American Life Panel (ALP)

The ALP is an Internet panel of approximately 1,500 respondents who are solicited once a month to participate in surveys typically taking less than thirty minutes to complete. From June 2009 through December 2010, a cohort of ALP respondents was asked to complete a monthly questionnaire that collected information about household spending in twenty-five medium- to high-frequency purchase categories and a quarterly questionnaire that collected data on spending in eleven less frequently purchased categories. Because outliers are a problem in self-administered surveys (where there is no interviewer to probe unusual reports), the ALP presented these respondents with a "reconciliation" screen at the end of each survey that provided a summary table listing the individual reported expenditures and the spending total for the household. Respondents then were asked to review this information and correct any items, but no automatic edit checks were used. Examining data from this panel, Hurd and Rohwedder (2010) found that ALP respondents corrected about 2–3 percent of entries per interview wave, and that there were significant reductions in item nonresponse and in the frequency and magnitude of outliers due to the reconciliation process. There also was good agreement between the total annual spending estimates derived from the reconciliation-aided ALP and those from CE over this time period (i.e., ALP spending was 96 percent of CE spending).

12.3.3 Gaps in Existing Research

Although the studies by Brzozowski and Crossley (2011) and Hurd and Rohwedder (2010) provide evidence that offering respondents an efficient means of reviewing and making appropriate changes to their prior responses improves the quality of survey estimates, there is no published empirical work that has examined the cognitive underpinnings of this effect. How do respondents interpret their task? What changes do they make to their reports? (The SHS work suggests that they are more likely to change reported income and assets, but this may be an artifact of the types of probes used by Statistics Canada, which tended to focus on missed income or sources of financing for larger purchases, not expenditures.) What are respondents' reactions to being questioned about previous reports, or to seeing their household spending totals? The answers to these questions are important for understanding the quality and consequences of implementing this kind of procedure, and the present study was designed to begin to fill a gap in the literature on these and related issues.

12.4 Method

12.4.1 Design and Procedure

The study was a small-scale, lab-based test that presented participants with a modified version of the CEQ survey and a balance edit procedure for reconciling expenditure-cash flow disparities. The test sessions were conducted in the Office of Survey Methods Research (OSMR) laboratory in the Bureau of Labor Statistics (BLS). At the start of each session, a researcher explained the study's purpose and procedures and obtained informed consent from the participants. Study participants then took part in a CAPI interview that asked a brief set of demographic questions about the household, global expenditure questions for thirty-four categories (covering all of the CEQ section topics), and questions about household income and changes in assets and liabilities for the reference month.[1] All participants were given a modified version of the CEQ information booklet to refer to throughout the survey. The booklet contained a set of flashcards that provided a list of examples of the kinds of expenditures asked about in each category. All three authors served as interviewers.

Expenditures, assets, and liabilities were collected for the preceding month (one-month recall period). We used a one-month recall period rather than a three-month period (as the CEQ does) or a flexible reference period (as is done in the SHS) for two reasons. First, we needed to measure the change in

1. For a full set of study materials see appendixes I–III in the working paper (http://www.bls.gov/cex/cesrvymethsfricker.pdf).

assets and liabilities over the reference period (e.g., depletions/additions to savings, changes to credit card balances) to get a full picture of the household cash flow. We felt that asking participants to recall the relative change in their accounts over a longer recall period would be very difficult, especially in a lab setting in which participants had limited or no access to their household records. The second reason was practical—it was not feasible to program and administer an electronic collection instrument that allowed flexible recall periods and tailored question fills; the resulting database tracking required for the balance edit would have been too complex for this study. The drawback of using a one-month recall period is that it likely misses larger and less frequent purchases, and it is possible that these types of purchases could affect both the need for reconciliation and participants' response processes during reconciliation. We attempted to probe for this information during a postinterview participant debriefing.

For household income, we allowed participants to report using a flexible reference period (weekly, biweekly, monthly, quarterly, or annual) because there were relatively few income questions (so programming was manageable), and we felt that participants would provide more accurate information if they could choose their preferred time period. One potential problem with this approach was that our unit of measure for the balance edit check was one month. We had to aggregate up from weekly and biweekly income reports (by multiplying by four or two, respectively) or disaggregate quarterly and annual reports (by dividing by three or twelve, respectively) in order to get a one-month income value.[2] Income can naturally fluctuate for some people over time, however. The extent to which the (dis)aggregated income values differ from participants' true income for the reference period impacts the likelihood of triggering the balance edit, with potentially more effort focused on "fixing" the derived income values. Again, we examined the effect of this issue on the survey outcomes and participants' reactions to the survey.

Balance Edit. Data were recorded in an Excel workbook that calculated the ratio of participants' income-to-spending over the reference period.[3] We calculated the ratio as follows:

$$\text{Income-to-Spending Ratio} = (I_i - C_{Ai} + C_{Li}) / S_{i,}$$

where I_i is the (derived) monthly, after-tax household income for household i, C_{Ai} and C_{Li} are the change in assets and liabilities, respectively, for the household in the reference month, and S_i is the total spending for the household in

2. In addition to giving participants the flexibility to choose their preferred recall period for income, they could report pre-tax or after-tax income. When pre-tax income was reported, the instrument automatically calculated an estimate of the after-tax value based on current federal, state, and county tax schedules.

3. Throughout the remainder of the report we use the phrase "income-to-spending ratio" for convenience to refer to the "income-plus-net-assets-and-liabilities-to-spending ratio" unless otherwise stated.

the month. A ratio less than 1.0 indicates that the household *spent more* than its reported available income (plus net assets and liabilities). A ratio greater than 1.0 indicates that the household *spent less* than its total available income.

After all of the basic expenditure, income, asset, and liability questions were asked, the interviewer showed participants a graph depicting the ratio of their income-to-spending and read the following text:

> Thank you very much for your time so far. I'd like to take a look now at the overall picture of your household finances last month based on the information we've collected from you. This simple chart plots your reported household expenditures and your income plus any assets and liabilities. Ideally, we'd expect to see that these two figures match up pretty closely. [IF RATIO EXCEEDS THRESHOLD, READ:] However, sometimes when there is a big difference between these two amounts in a given month, it's because we missed some of your HH's expenditures or income, or need to make other changes to bring these in line.

The balance edit check was triggered by one of two income-to-spending ratios. For households with at least $30,000 in annual income, the balance edit was triggered if this ratio deviated by 15 percent or more from 1.0. For households with annual income below $30,000, the balance edit was triggered if the ratio differed by 20 percent or more from 1.0. When participants' income-to-spending ratios indicated acceptable balance at this stage of the interview (Phase 1), the interviewer terminated the interview and proceeded to the debriefing portion of the study session.

If households were out of balance (based on the criteria above), participants were given the opportunity to review a summary page of their reported expenditures, income, assets, and liabilities (individual reports and summed totals) and to make changes to their earlier reports in order to bring their ratio closer to 1.0. We recorded any changes made by participants during this review and revision phase (Phase 2), and then showed them a revised graph of their updated household income-to-spending figure. If the household was still unbalanced at that point in the interview, the interviewer administered a brief set of probes designed to capture additional sources of income and expenditures that might have been missed (Phase 3). The balance edit procedure was terminated in Phase 2 or Phase 3 when expenditures were within 10 percent of reported income plus net assets and liabilities, or when all of the CEQ items and follow-up probes had been administered.

Debriefing. Following administration of the modified CEQ (with any balance edit checks required), participants filled out a short paper-and-pencil questionnaire and then participated in a semistructured debriefing session with the interviewer. The purpose of the debriefing was to assess the following topics:

- participants' general reaction to the survey and the balance edit procedure;

- participants' perceptions of the accuracy of their reported data;
- sources of confusion or conceptual difficulty (e.g., global items, reference period, proxy reporting); and
- participants' perceptions of survey burden.

12.4.2 Participants

Participants for this study were recruited from an OSMR-maintained database of individuals who responded to advertisements for research studies placed in Washington, DC-area newspapers. We used a nonprobability-based convenience sample, but attempted to recruit participants who varied in their family size, educational attainment, household income, and employment status.[4] We interviewed twenty participants but report findings for only nineteen; one individual provided insufficient data during the survey and debriefing session. All twenty participants were paid $40 for their time. Interviews lasted approximately one hour on average—thirty minutes to administer the modified CEQ survey and balance edit, and thirty minutes to conduct the participant debriefing session.

12.4.3 Data Quality Indicators

In the next section of the chapter we present our study findings. We focus on several key results. We do not have a benchmark for "true" spending and income, so we examine the level of expenditure reporting as one commonly accepted measure of data quality. Balance edit check measures have been developed primary to reduce the likelihood of response errors due to underreporting, and thus more reporting is taken as evidence of better reporting. Another one of our primary measures of interest is the change in the household income-to-spending ratio across the interview. If the balance edit procedure implemented in this study was effective, we would expect to see those ratios improve (i.e., get closer to 1.0). Finally, we explore the qualitative responses obtained in our debriefing session to give us some additional insight into the quality of the data reported in the interview, the factors that may impact the effectiveness of a balance edit, and participants' reactions to their survey experience.

12.5 Results

12.5.1 Effect of Balance Edit on Income-to-Spending Ratios

Table 12.1 shows the distribution of participants' income-to-spending ratios at different phases of the survey. The first column shows the ratio cri-

4. We achieved reasonable balance on participant education, gender, and family size, but higher-income households and employed individuals were overrepresented in our sample. We had only two respondents with household incomes under $30,000 (sample mean: $67,800) and only three who were unemployed or retired.

Table 12.1 **Number of participants in each income-to-spending ratio group after each interview phase**

Deviation from income-spending ratio of 1.0 (%)	End of Phase 1 (main questionnaire)	End of Phase 2 (review & revise)	End of Phase 3 (additional probing)
0.0–10.0	1	3	5
10.1–15.0	0	*3*	*3*
15.1–20.0	*2*	*3*	*3*
20.1+	*16*	*10*	*8*

teria that we used to determine degree of balance. The second column indicates the number of participants who fell into each ratio category at the end of Phase 1 of the interview (prior to any balance edit procedures). Only one participant obtained an income-to-spending ratio good enough—0.91—to obviate the need for a balance edit. Two additional participants obtained ratios of 0.82 and 1.19 in Phase 1, respectively, but the balance edit procedure was triggered because they had annual household incomes over $30,000 (for which "balance" was defined as a deviation of 15 percent or less from a ratio of 1.0). The remaining sixteen participants were considerably unbalanced at the completion of the basic questionnaire, with a mean deviation from unity of 43 percent (i.e., reporting 43 percent more available household income than reported spending).

The third and fourth columns of table 12.1 give one indication of the extent to which the balance edit procedure was effective. Recall that once the balance edit was triggered in this study (using the 15 percent or 20 percent deviation criterion at Phase 1), a household was deemed balanced if it achieved an income-to-spending ratio between 0.90 and 1.10. Examining the first row of the table, we see that two additional participants achieved balance after being given the chance to review a summary of their reports and make revisions (Phase 2, column [3]), and two more participants achieved balance after answering the additional expenditure and income probes (Phase 3, column [4]). Thus, a total of five participants out of nineteen obtained acceptable income-to-spending ratios by the end of the interview; one participant did so without going through the balance edit and four did so only with the help of the balance edit.

The italicized numbers in table 12.1 represent the number of participants who did not reach balance in any phase of the interview by our study criteria. Nevertheless, there is some indication in these cells that the balance edit did have a positive impact on participants' income-to-spending ratios in the aggregate. For example, the number of participants with ratios deviating by more than 20 percent from 1.0 was cut in half by the end of the balance edit process (from 16 to 8 percent). This effect is even more evident when we

Table 12.2 Distribution of the direction of ratio changes in the interview

Type of ratio change	No. of participants
Moved closer to 1.0	13
More income than spending (ratio > 1.0)	8
More spending than income (ratio < 1.0)	3
Below 1.0 to above 1.0	2
Above 1.0 to below 1.0	0
Moved farther from 1.0	1
Flipped	2
No change	2

Note: Includes only participants who were administered the balance edit ($n = 18$).

examine the pattern of changes in income-to-spending ratios for individual participants over the course of their interviews (see table 12.2).

Table 12.2 presents the number of participants whose ratios improved, worsened, stayed the same, or simply switched direction but not magnitude over the course of the interview. The majority of our participants (thirteen of nineteen) moved closer to being balanced as a result of the balance edit process. (However, despite making relative gains in their ratios during the interview, these participants' reports remained notably unbalanced, with an average deviation of 42 percent from unity.) Most of the movement toward balance in this group was the result of participants (eight of thirteen) who initially reported higher income than spending but then subsequently reported additional expenditures (these participants' ratios got smaller but remained above 1.0). Five participants initially reported significantly more spending than income and moved closer to balance throughout the interview by making small reductions to their reported spending and increases to income (e.g., reporting additional wages or forgotten tax refunds). Only one participant did worse over the course of the interview, and four individuals essentially remained unchanged (either by making no adjustments to their reports, or by flipping the sign of their ratio but not decreasing its magnitude).

12.5.2 Levels of Reporting

Table 12.3 gives a more concrete look at participants' reporting throughout the interview. The second column shows the average number of reports and average dollar amounts given by our study participants for the different survey topics in the main questionnaire (Phase 1). Consistent with our earlier findings, prior to the balance edit check our sample in the aggregate reported slightly more income (plus net change in assets and liabilities) than spending. During the initial review and revise component of the balance edit procedure (Phase 2), 61 percent of participants ($n = 11$) made changes to their reported spending (with an average of one change per participant),

Table 12.3 **Reporting incidence and level by topic and interview phase**

	Phase 1 (main questionnaire) (n = 19)		Phase 2 (review and revise) (n = 18)		Phase 3 (additional probes) (n = 16)	
	Average total ($)	Average no. reports	Reporting change (%)	Average change ($)	Reporting change (%)	Average change ($)
Expenditures	4,781	14	61	92	56	188
Income	5,196	2	33	589	50	973
Change in assets	67	1	22	28	n/a	n/a
Change in liabilities	49	0.5	5	–528	n/a	n/a

33 percent (n = 6) made changes to reported income (with an average of two changes per participant), but relatively few revisions were made to reported changes in assets or liabilities (number of changes per participant: median 0, mean 0.2).[5]

The mean dollar change figures shown on the right-hand side of the Phase 2 column reflect averages across the entire eligible sample (i.e., all participants who were administered the Phase 2 balance edit, regardless of whether a change was made or not). Restricting our analyses to only those participants who made revisions in Phase 2, the mean dollar change in expenditures was \$159 ($SD_{\bar{x}}$ = \$429) and the mean change in income was \$1,866 ($SD_{\bar{x}}$ = \$2,621). During Phase 3, in which participants were probed about possible sources of spending and income missed earlier in the questionnaire (but not new changes in assets or liabilities), 56 percent (n = 9) of our sample reported additional expenditures (overall mean, \$188; reporter mean, \$503), and 50 percent (n = 8) reported additional income (overall mean, \$973; reporter mean, \$1,757).

12.5.3 Debriefing

Following the completion of the modified CEQ survey, participants completed a self-administered questionnaire that asked them to rate how comfortable they felt sharing expenditure, income, asset, and liability information during the interview, and how accurate they felt those reports had been. The original questions used a 5-point Likert scale—ranging from "very uncomfortable" to "very comfortable," and "very inaccurate" to "very accurate," with a neutral "neither" middle response option. We collapsed the "very" and "somewhat" categories for each dimension and omit the middle response option data for reporting purposes. Table 12.4 presents those data.

The majority of our study participants reported feeling "very" or "somewhat" comfortable providing the household financial information asked

5. Only one participant in this study revised her reported change in liabilities, recalling that she had paid off a business loan for \$9,500 during the reference month.

Table 12.4 **Participants' ratings of comfort and accuracy (%) by reporting category (n = 19)**

	Comfort		Accuracy	
Category	Very or somewhat uncomfortable (%)	Very or somewhat comfortable (%)	Very or somewhat inaccurate (%)	Very or somewhat accurate (%)
Expenditures	5	65	5	80
Income	30	55	0	85
Change in assets	30	60	5	85
Change in liabilities	15	60	0	95

during the interview, though reporting income was more sensitive for some participants than discussing expenditures or changes in assets and liabilities. In response to a follow-up question, several participants acknowledged that they felt awkward disclosing income to a stranger, but understood the purpose of the question and believed that their responses were important and would be kept confidential. Additionally, several participants said that they were comfortable reporting income information for themselves, but had been reluctant or unable to do so for other household members.

Participants' ratings of the accuracy of their reports were very high (80–95 percent). On the one hand, this may simply reflect the effects of self-presentation management—the desire to represent oneself as a diligent and accurate participant in the data collection effort. On the other hand, a number of participants in this study evinced behavior or made explicit comments that indicated that they had engaged in effortful and thorough recall and reporting. For example, most participants followed along with the information booklet during the interview and said during debriefing that it provided definitional clarity to some of the global expenditure categories/items, and helped them recall expenditures they would otherwise have forgotten to report. Two participants spontaneously brought out their checkbooks or personal calendars during the interview to aid reporting. In addition, the balance edit process itself may have strengthened participants' perceptions of accuracy since most participants made adjustments to their reports (recalling forgotten items, revising earlier reports with greater specificity) and saw visual evidence that those efforts led to improvements in their household cash flow balance.

The self-administered questionnaire also included items to assess participants' perception of survey length and burden. We asked participants to estimate how long the interview lasted under the assumption that those estimates would exceed actual survey length if participants felt burdened (see, e.g., Block 1990). The average estimated interview duration was 32.3 minutes, about five minutes longer than the actual interview duration (mean

27.2), suggesting some degree of respondent burden. However, when asked whether the interview was "too long, too short, or about right," all participants replied that the survey length was "about right." In addition, we asked participants to rate how burdensome they felt the survey was and again the responses were uniformly positive (e.g., "It was not at all burdensome." "It was great." "It was very interesting and easy.").

Factors Affecting Accuracy. In subsequent debriefing, participants identified several factors that they felt affected accuracy. Five participants said that they could have reported more accurately if they had been given advance notice to record their expenses in some form (e.g., using an Excel spreadsheet, using their phone, bringing records to reference, writing it down on paper). The size, frequency, and saliency of expenditures was mentioned by half of the participants as contributing to their reporting accuracy, with smaller everyday expenses like food and transportation reported as more difficult to recall accurately than more stable items like income, mortgage payments, and utilities.

Participants' household composition—its size and the division of financial responsibilities—also played a role. Eleven out of fourteen participants who lived in multiperson households said that they had a "good sense" of what other people in their house bought and how much they spent, but they also identified gaps in that knowledge (e.g., food eaten out by spouses or children, gas expenditures, purchases made on other's credit cards, etc.). To a lesser extent, participants made similar comments about their knowledge of other household members' income, assets, and liabilities. In fact, four participants admitted to "forgetting" to include some or all of their other household members' expenditures and income (e.g., "I was just focused on me!"). Two of those participants were able to remedy these omissions during the balance edit, but two did not realize their mistake until the debriefing discussion.

We also asked participants whether the one-month reference period presented any reporting difficulties and if a three-month reference period would have been easier or more difficult. One self-employed participant said the one-month reference period was difficult because her income varied considerably from month to month, but none of the other participants reported difficulties with the monthly time frame. Participants' views about the efficacy of a longer reference period varied by topic. For expenditures, seventeen participants said that using a three-month reference period would be more difficult than using a one-month reference period because of the additional memory demands and the fact that some expenses are intermittent and therefore more likely to be forgotten. We asked participants how they would come up with their total household expenses for three months. The word "estimate" was used often, and many participants said that they would think of their "typical" monthly expenses and multiply by three. A few participants did say that they would think of the expenses for each

specific month and try to systematically calculate an accurate three-month figure. For income, two-thirds of our study participants said reporting for a three-month period instead of one-month period would be essentially the same because their income was fairly regular and stable. For assets and liabilities, our sample was split: about half preferred the one-month and half preferred the three-month reference period. The responses depended largely on how closely the participant tracked their accounts and how regular or irregular their account activity usually was.

Reactions to the Balance Edit. We were interested in exploring participants' reactions to various features of the balance edit implemented in this study. We began by asking them about the chart they were shown at the end of the basic questionnaire that displayed their income-to-spending ratio. Opinions fell into one of three groups. The first group (n = 3) initially felt confused by the chart; they did not understand its purpose or what it was supposed to represent, and had a hard time comprehending the concept of income-plus-net-assets-and-liabilities. These participants indicated that the accompanying explanation about the chart provided by the interviewer was helpful in deciphering its meaning, or at least in clarifying the overall objective of the balance edit process.

A second group of participants (n = 5) understood the purpose of the chart and balance edit but it elicited some emotional reaction, often somewhat negative. For example, several participants said that it was somewhat surprising and uncomfortable to be confronted with a chart that showed spending in excess (sometimes far in excess) of their income. Others in this group whose reports were unbalanced initially made an inference that they must have done something wrong (e.g., "I felt a little stupid."), or that by being asked to review their previous reports that the interviewer mistrusted them in some way ("I was a little frustrated because I knew I was being truthful."). In contrast, two participants in this group whose reports were reasonably balanced and showed more income than spending expressed satisfaction upon seeing the chart (e.g., "I don't want to be balanced. I want to have more income so I can spend and save more.").

Finally, half of our participants (n = 10) seemed to generally comprehend the chart and the objective of balance edit, and either have no emotional reaction or be genuinely intrigued by the information presented and motivated to resolve reporting discrepancies. The participants in this group were able to clearly articulate the objective of the edit process in their own words and did not appear to have any conceptual difficulties or negative emotional reactions.

Table 12.5 shows the relationship between these subjective participant groupings and the balance edit outcomes. Given the small and disproportionate sizes of the three groups, caution should be used when making inferences based on the results in this table. However, it does appear that there is a relationship between these groupings and the magnitude of change in the

Table 12.5 **Relationship between participants' initial reactions to balance edit and survey outcome**

	Ratio moved closer to 1.0	Average change in ratio (%)	Achieved balance
Group 1—Initially confused (n = 3)	3	–52.3	0
Group 2—Emotional response (n = 5)	3	50.7	0
Group 3—Understood/positive response (n = 10)	7	–2.0	4

Note: Includes only those participants who were administered the balance edit (n = 18).

income-to-spending ratio over the interview (column [3]) and in the final balance status (column [4]).

Participants who were confused by the chart and balance edit objective (Group 1) initially reported greater income than spending. Those who had an emotional reaction initially (Group 2) reported greater spending than income. The magnitude of the average ratio change was very similar for both groups, and no participants in either group achieved balance. Participants who understood the chart and edit objective without any associated negative emotional reaction (Group 3) made smaller (or at least off-setting) changes on average during the balance edit than participants in the other two groups, and all of the balanced households came from this group.

We also queried participants on their reactions to Phase 2 of the interview in which we gave participants the chance to look over a summary page of their reports and make any corrections or additions they felt were needed. Again, reactions tended to be split. Most participants did not express any substantive opinion about this page—they simply engaged in the review process and moved on. Three participants gave only cursory examination of this page and either did not fully understand their task or chose not to exert the effort required to review the information more carefully; they made no changes on this screen. A few participants described the review and revise process as "daunting'" or "chastising" because it forced them to examine (in front of the interviewer) some hard truths about their household finances. And, finally, several other participants said that they really liked the table and task because it prompted them to consider the relative amounts reflected in each category or because the presentation triggered memories of additional items that they then could report.

Some of the most frequent adjustment and additions made on this page stemmed from participants' memories for expenditures related to landmark events (e.g., birthdays, trips, Mother's Day). Another common change that participants made stemmed from financial activities by other household

members. As noted earlier, sometimes this information was neglected altogether in participants' answers to the basic questionnaire, other times it was only partially reported or participants simply provided best guesstimates. This page afforded participants the chance to refocus attention on proxy-related items they may have missed (e.g., spouses contributions to retirement accounts), or revise their estimates for other household member items in order to achieve better balance. There also were a few instances in which we believe that participants were just simply trying to improve the household balance by making seemingly arbitrary adjustments, but this was not common.

In the final phase of the interview, participants who remained unbalanced were asked a brief set of additional probes about possible sources of expenditures and income. Fifteen of the sixteen participants who received these questions said that they were clear and easy to answer. The two most common items that we picked up as a result of these probes were child care expenses and tax refunds/tax payments (the reference month in this study was April).

12.5.4 Discussion

The findings from this small study shed some light on the challenges and opportunities that a balance edit introduces to expenditure surveys. We found that eighteen of nineteen participants in this study provided reports in the initial interview that were sufficiently unbalanced to trigger the balance edit. The balance edit procedure improved income-to-spending ratios for thirteen participants, but only four individuals actually achieved balance (i.e., obtained ratios that were within 10 percent of unity). Despite the relative improvements in the ratios of the remaining nine participants, their final average deviation from unity remained quite high (42 percent). The debriefing session revealed that a sizable minority of individuals in this study either did not understand the purpose in the balance edit or had somewhat negative reactions to the process. In addition, we identified a number of factors that likely contributed to its effectiveness (e.g., participant knowledge of other household members' spending/income; the size, frequency, and salience of the expenditure category).

12.5.5 Limitations

Our study was limited by a number of factors that may have impacted the results. First, we used a small, convenience sample; a larger, probability-based sample would have strengthened our ability to generalize the results and reduced the influence of outliers introduced by individuals with very large (small) reports. Second, some of the participant confusion evidenced in this study may have stemmed from our specific implementation of the balance edit—a hybrid between the approaches taken in the SHS and ALP. In particular, the language we used to introduce the objective of the edit and to

describe the chart reflecting the income-to-spending ratio likely could be improved based on what we learned in this study. In our roles as interviewers, we strove to provide participants additional task clarification as necessary, but further refinement and standardization of this feedback might have improved participant understanding. Third, some of the difficulty in achieving balance may be attributable to the use of a variable reference period for income. Half of the study participants selected reference periods other than the one-month period used for expenditures and change in assets/liabilities (e.g., biweekly or yearly). The calculations used to make these reference periods comparable (e.g., dividing yearly income by twelve) could create problems if the participant's flow of income is not steady over the course of the year. During the interview we did confirm the derived estimates for monthly income with those respondents who initially reported using a different reference period, but the differential use of flexible reference periods may have contributed to the lack of balance for these respondents. Finally, our decision to use global expenditure questions may have affected the incidence of out-of-balance households. Global questions may tend to encourage participants to provide rough estimates for many expenditure categories when other response strategies would be more optimal. Their use also may have affected participants' perceptions of the burden of the survey and the balance edit process; attempting to conduct our balance edit with the full complement of detailed, disaggregated CEQ questions likely would have been far more difficult for participants.

12.6 Priorities for Future Research

Despite evidence that the balance edit procedure used in this study led to some improvement in data quality, this methodology requires substantial additional testing. As suggested elsewhere in this report, the efficacy and appeal of a balance edit will depend largely on the overall survey design (e.g., mode, use of detailed versus global items, length, etc.) and its analytic purposes. Were the CEQ to incorporate significant numbers of global expenditure questions (which we expect are subject to significant reporting errors) and still be conducted primarily in person, then some form of a balance edit may be worth considering. If, on the other hand, the CEQ continues to be a long survey consisting of hundreds of detailed expenditure questions, and/or a significant portion of the interviews each month are administered by telephone, then we believe that this procedure is less attractive given implementation issues and the potential negative impact on respondent and interviewer burden. Additionally, we used a one-month, not a three month, reference period for this study, and we suspect that the longer time frame would reduce the quality of respondents' estimates of change in assets and liabilities, in particular. Depending on design changes under consideration for the CEQ, further research should be done on whether a balance edit is

feasible over the phone, with a larger set of detailed reports, and with a longer reference period. Finally, this study was designed to examine only relative changes in respondent reporting; we had no direct measures of the actual quality of the data collected. Future research should examine measurement error more directly, for example, through the use of record validation, or studies of within-household reporting consistency/reliability (e.g., comparing reports between spouses).

References

Block, R. A. 1990. "Models of Psychological Time." In *Cognitive Models of Psychological Time*, edited by R. A. Block, 1–35. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bosworth, B., G. Burtless, and J. Sabelhaus. 1991. "The Decline in Saving: Some mMicroeconomic Evidence." *Brookings Papers on Economic Activity 1:*183–256.

Brzozowski, M., and T. F. Crossley. 2011. "Measuring the Well-Being of the Poor with Income or Consumption: A Canadian Perspective." *Canadian Journal of Economics* 44:88–106.

Hurd, M., and S. Rohwedder. 2010. "The Effects of the Financial Crisis and the Great Recession on American Households." NBER Working Paper no. 16407, Cambridge, MA.

Jacobs, E., and S. Shipp. 1993. "A History of the US Consumer Expenditure Survey: 1935–36 to 1988–89." *Journal of Economic and Social Measurement* 19 (2): 59–96.

Shields, J., and N. To. 2005. "Learning to Say No: Conditioned Underreporting in an Expenditure Survey." Paper presented at the annual meeting of the American Association for Public Opinion Association, Fontainebleau Resort, Miami Beach, Florida.