

Measurement Issues Study Final Report

Team Composition:

Brett Creech

Branch of Information and Analysis, OPLC

Meaghan Smith

Branch of Production and Control, OPLC

Jeanette Davis, Lucilla Tan, Nhien To

Branch of Research and Program Development, OPLC

Scott Fricker, Jeffrey Gonzalez

Office of Survey Methods Research

Project Lead: Scott Fricker

Analysis Sub-Team: Fricker, Gonzalez, Tan

Training Sub-Team (BLS): Creech, Davis, Fricker, To

November 14, 2011

EXECUTIVE SUMMARY

Overview

The Measurement Issues Study (MIS) was part of a comprehensive and ongoing effort to examine alternative data collection strategies for the Consumer Expenditure Quarterly Interview Survey (CEQ). These strategies seek to improve data quality, maintain or increase response rates, and reduce respondent burden and data collection costs. One component of the 2010 CE Strategic Plan was to address the following three survey design questions: (1) does splitting the questionnaire reduce respondent burden and/or improve data quality; (2) do monthly reference periods provide better quality data than quarterly reference periods; and, (3) do global questions provide data of sufficient quality to replace detailed questions? The MIS was originally designed to offer insights to the first two questions, but the incorporation of global questions in one of study treatment conditions also provided an opportunity to examine the third survey design question raised in the 2010 CE Strategic Plan.

Thus, the MIS had three research objectives: (1) to assess the effects of administering a shorter CEQ instrument on respondent burden, data quality, and nonresponse error; (2) to examine the impact of using a one-month (versus the current three-month) reference period on respondent burden, data quality, and nonresponse error; and (3) to evaluate the quality of data collected from global, as opposed to, detailed questions on expenditures. The findings from this study were intended to help inform future Consumer Expenditure (CE) research activities as well as redesign decisions for the CE Surveys.

The study utilized a basic experimental design in which respondents were randomly assigned to a control group which received no treatment, a test group that was administered a shortened version of the questionnaire, or a test group that was administered a shortened reference period. It used a truncated CEQ interview, a restricted panel design, and was administered by centralized computer-assisted telephone interviews (CATI) using Census Bureau Telephone Center interviewers. The treatment conditions were as follows:

Control Group (C)

In the C condition, sample units completed a bounding interview in wave 1. The bounding interview used a 1-month reference period and consisted of items taken from nine sections of the current CEQ instrument plus a “core” set of questions (e.g., demographic items) that were administered across all study conditions. These same C group sample units were contacted again three and six months later to complete two additional interviews using the same “core + nine sections” questionnaire with a 3-month reference period. The C condition paralleled the existing CEQ survey procedures and served as the basis of comparison for the other experimental conditions.

Shortened Questionnaire (SQ)

In the SQ condition, sample units completed the same full bounding interview in wave 1 as the C condition cases, and then were randomly assigned to one of two subsamples that were administered subcomponents of the full questionnaire in waves 2 and 3 three and six months later. Subsample A (SQ-A) received sections 6, 14, 16, 18, and 20, the “core” questions, and a small number of global expenditure questions from sections 9, 12, 13, and 17. Subsample B (SQ-B) received sections 9, 12, 13, and 17, the “core” questions, and a small number of global expenditure questions from sections 6, 14, 16, 18, and 20.

Reference period (RP)

In the RP condition, sample units received the same “full” bounding interview that was administered to the wave 1 C and SQ respondents. They then received three consecutive monthly interviews using the same “full” questionnaire with a 1-month reference period (rather than the 3-month reference period used in the C and SQ interviews).

Outcome Measures

The key outcome measures for the three dimensions of interest - respondent burden, data quality, and nonresponse error – were defined in the MIS as follows. Respondent burden was assessed by examining respondents’ answers to a series of questions administered at the end of the last interview. These questions addressed respondents’ perceptions of the appropriateness of the survey length and the number of interviews, the effort required, and interest in the survey topic. Data quality was assessed by adopting the premise that “more reporting is better” (in terms of

both number of items and absolute dollar amount). In addition, the MIS Team examined record usage, information book usage, combined expense reporting, and the amount of “don’t know” and “refusal” responses. To assess the potential for nonresponse error in this study, we examined response rates, panel attrition rates, estimates of relative nonresponse bias, and changes in respondent sample composition across the waves of the survey.

Key Findings

Shortened Questionnaire (SQ)

1. Data quality moderately improved under the SQ treatment relative to the control condition. Both SQ subsamples (SQ-A and SQ-B) produced total expenditure estimates that were higher than the control estimates, although only the SQ-B group reached statistical significance. In addition, the SQ-B group reported significantly more expenditure reports than the C group. The SQ treatment did not substantively impact the incidence of negative respondent behaviors (i.e., combined reports, “don’t know/refusals”) or the use of recall aids or records.
2. The effects of the SQ treatment on indicators of nonresponse error were minor, varied, but generally positive. Response rates examined independently by interview wave revealed no treatment effect (i.e., they were comparable for the SQ and C groups at each wave). However, the SQ groups attained significantly lower attrition rates between wave 2 and wave 3 than the C group (0.7% and 0.9% for the SQ groups vs. 2.8% for C). The final wave cumulative response rate (i.e., conditioned on participation in wave 1) also was higher in the SQ groups than the C group. There were no observed differences in sample composition between the SQ and C groups in the final wave. Finally, compared to the C group, the SQ treatment reduced the relative nonresponse bias in total expenditures estimates as well as vehicle operations expenditures estimates, though there was evidence that it also exacerbated the bias existing in a few of the C group expenditure estimates.
3. Respondent burden was significantly lower in the SQ groups than C group. SQ respondents perceived the survey to be less burdensome and of appropriate duration and frequency, compared to the control group respondents. SQ interviews were 6 minutes shorter than C interviews on average.

Reference Period (RP)

1. Evidence on the effect of RP treatment on data quality was mixed. There were some indications that RP improved data quality. For example, respondents in the RP group did report significantly more valid expenditure reports, and the total expenditures estimate in this group was higher than the C estimate (but not significantly so). In addition, RP respondents were more likely than C respondents to use the Information Booklet to prepare for the survey in advance. On the other hand, RP respondents were significantly more likely than the C respondents to engage in undesirable reporting behaviors (e.g., use of combined item reporting and “don’t know” and/or “refused” responses). In particular, the RP group was higher in both of these undesired reporting behaviors for section 9 (clothing), a section that is already problematic in the current instrument using a 3-month recall. The RP group had nearly three times as many “don’t know/refusals” as the C group; represented as a percent of the average total number of reports, the RP group’s rate of DK/REF was 23% compared to 13% for the C group. There was no difference in use of records between the RP and C groups.

2. The RP treatment had a negative impact on survey participation. Response rates examined independently by wave and conditional on wave 1 participation were lower for the RP group than the C group in waves 2 and 3. The attrition rate between wave 1 and 2 also was substantially higher for the RP group (17.2% vs. 13.5%), possibly due to the RP group’s tighter fielding period and/or the saliency of respondents’ prior wave (negative) experience.

3. Overall, it does not appear that RP treatment worsened any potential nonresponse bias that may have existed in the C group. The sample of respondents in RP and C were generally similar in distribution on the selected demographic characteristics. In addition, the RP data showed less relative nonresponse bias in total expenditure estimates and estimates of health insurance spending and regular weekly expenditures compared to the C group. However, the RP group showed worse nonresponse bias for estimates of education and appliances expenditure (which were over-estimated) and non-health insurance expenditures (which was underestimated).

4. There were significant and strong RP treatment effects on a number of respondent burden dimensions. Significantly more RP than C respondents said that the survey was “not very / not at all interesting” and “very / somewhat burdensome,” and that there were “too many” survey

requests. In contrast, more RP respondents than C respondents said that the survey questions were easy.⁶ Moreover, despite the fact that actual interview durations were significantly shorter in the RP group than the C group (by more than 4 minutes in waves 2 and 3), proportionally more respondents in the RP group perceived their final interview to be too long.⁷

Global Questions

Global-based spending estimates were significantly higher than detailed-based estimates in six of the ten expenditure categories examined in this study (clothing, vehicle operations, non-health insurance, health insurance, entertainment, and trips), and significantly lower in only one (books/subscriptions). We present evidence that the use of global questions reduced levels of don't know / refused responses, as well.

Study Limitations

As with any field test, an effort was made to mirror as many of the CEQ survey procedures and conditions in this study as possible (e.g., use of a panel design that incorporated a bounding interview, use of CEQ questions and materials), but there were significant departures that may weaken the inferential value of this study. For example, due to the prohibitive cost of conducting in-person data collection, we relied on centralized computer-assisted telephone interviews (CATI). As a consequence of this mode change, we also eliminated sections of the survey to shorten its overall length. Changes to mode, length, and question context impact the response process and associated errors, so it is likely that some of our results would have been different under a design closer to that of the CEQ. In addition, we were restricted by the project budget to a relatively small sample size. This reduced our power to detect some treatment effects and prevented us from examining effects at lower levels of analysis (below the section-level).

Recommendations

The results of this study suggest that a SQ design may hold promise in a redesigned CEQ. Additional research is needed to determine the optimal length of a shortened survey, composition of questionnaire splits (in terms of their statistical properties and impact on respondent processes/errors), and dataset construction and analysis methods. We are less sanguine about the adoption of a 1-month reference period, given the concomitant need for conducting monthly interviews, and our findings on the negative effects of this design on response rates and

respondent burden. That said, the optimal reference period likely will vary across expenditures, and additional laboratory research is needed in this area. Similarly, we recommend additional research (e.g., cognitive studies, controlled experiments, validation studies) on respondents' use of global questions. Finally, this study underscores the benefits of examining a variety of quality metrics. We recommend that as CE moves forward with its redesign efforts, it considers establishing a panel in the revised design that offers sufficient power to detect statistical differences and track changes, and to include a range of evaluative criteria (e.g., level of reporting, respondent burden index, indirect indicators of data quality, etc.).

1. INTRODUCTION

1.1 Background

The U.S. Consumer Expenditure Survey is an ongoing monthly survey conducted by the U.S. 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 expenses incurred in the 3-month period that just ended.

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, there is a widespread belief that some CEQ data are underreported. Underreporting has been variously attributed to recall error, panel conditioning, respondent fatigue, satisficing, and other causes. The length and perceived burden of the CEQ survey may also have deleterious effects on response rates.

1.2 Study Objectives

This study was the first in a comprehensive and ongoing effort to examine alternative data collection strategies for the CEQ that may improve data quality, maintain or increase response rates, and reduce data collection costs. In particular, this study assessed the effects of administering a shorter CEQ questionnaire on respondent burden, data quality, and nonresponse error. A separate condition in this study examined the extent to which using a 1-month (versus a 3-month) reference period affected underreporting due to recall errors. The study design enabled BLS to perform data quality analyses using both direct measures (e.g., number of expenditure reports, expenditure amounts) and indirect

measures (e.g., response rates, measures of perceived burden, item nonresponse, etc.), and to estimate nonresponse bias by comparing response rates, sample composition, and expenditure estimates across treatment conditions. The results from this study will be used to inform future CEQ research activities and decisions about how to redesign the production survey.

2. PREVIOUS RESEARCH

2.1 Survey Length

Survey organizations routinely limit the length of their surveys under the assumption that longer surveys can negatively impact a number of survey quality outcomes. The empirical literature examining this issue has primarily focused on the effect of length on nonresponse, and the results from these studies are mixed. For example, some studies have found that longer surveys or more frequent survey requests decrease response rates (e.g., Collins et al., 1988; Dillman et al., 1993), increase drop-out rates (e.g., Haraldsen, 2002), and reduce respondents' willingness to respond to future surveys (e.g., Apodaca et al., 1998; Groves et al., 1999). In contrast, other studies have found that longer interviews are associated with higher response rates and panel-survey sample retention (e.g., Champion and Sear, 1969; Branden et al., 1995) or have no association at all (e.g., Sharp and Frankel, 1983; McCarthy, Beckler, and Qualey, 2006). Although data conflict regarding whether survey length increases various forms of nonresponse, the evidence *in toto* suggests that there is at best a weak positive association. One reason for these equivocal findings is that respondent motivation to participate is affected not only by length, but also by a variety of other factors such as topic interest or the survey sponsor.

Motivation may additionally affect data quality more broadly. Individuals' motivation to respond in a thoughtful manner may decrease over the course of a long survey due to respondent fatigue or boredom. Although there has been less empirical attention to the impact of survey length on data quality than nonresponse, several studies provide evidence that respondents in longer surveys have greater likelihood of straight-line responding (Herzog and Backman, 1981), increased rates of item-nonresponse (Galesic, 2006; Peytchev and Tourangeau, 2005), and provide fewer survey reports (Backor, Golde, and Nie, 2007) than those in shorter surveys. Data quality also has been shown to deteriorate over the course of an interview, with increases in item

nonresponse, ‘don’t know’ reports, and response order effects, and less time spent on each question, the longer the duration of the interview (e.g., Krosnick, 1999; Peytchev, 2005; Roberts *et al.*, 2010). Taken together, these findings corroborate the received wisdom that interview length should be kept to a minimum, both to avoid the potential for nonresponse and to reduce satisficing behavior that can jeopardize survey data quality.

2.1.1 Split Questionnaires

The practical reality is that some surveys are excessively long, and often it is not feasible to simply cut items from a questionnaire to achieve reductions in respondent burden. Survey organizations may need to ask a large set of questions to meet stakeholder analytic objectives and to accommodate periodic requests to add new questions to an existing instrument. One method that has been developed to shorten surveys while still achieving the analytic needs of the organization is the use of split questionnaires (also referred to as multiple matrix sampling; see, e.g., Raghunathan and Grizzle, 1995). In one implementation of a split-questionnaire survey design, the original survey is divided into one ‘core’ component containing high-priority questions (e.g., socio-demographic variables) and a number of subcomponents containing approximately equal numbers of the remaining items. The full survey sample is likewise split into distinct subsamples, and each subsample of respondents completes the core component plus a randomly assigned subcomponent. Figure 1 provides an illustration of a split questionnaire survey design with a core component and three subcomponents.

Figure 1. Split Questionnaire Design with Three Components

| Respondent Subsample | Questionnaire Split | | | |
|----------------------|---------------------|----------------|----------------|----------------|
| | Core Component | Subcomponent A | Subcomponent B | Subcomponent C |
| A | | | | |
| B | | | | |
| C | | | | |

Split questionnaire designs reduce the length of the survey while still collecting the necessary information from at least some of the sample members, but they also result in missing data. The goal is to minimize the amount of information lost relative to the complete questionnaire, and appropriate decisions must be made at various phases of the survey process to aid optimal implementation and estimation. Survey designers must determine how to best construct the questionnaire splits (e.g., random allocation of items to subcomponents; grouping logically-

related items within a subcomponent; distributing highly correlated items to different components; use of a core component or not). They must decide which subset of the full sample will receive any given questionnaire component(s) (e.g., through random or predictive assignment). And they must select techniques for analyzing the resultant data (e.g., available case method; single imputation; multiple imputation; adjustments to calibration weighting). Gonzalez and Eltinge (2007a; 2007b; 2008; 2009) provide in-depth treatments of these issues and the foundational discussions of design, implementation, and analysis considerations for a potential application of split questionnaire design to the CEQ.

A relatively small but growing literature suggests that carefully implemented split questionnaire designs can be effective in producing key population and subpopulation estimates, and for reducing respondent burden, compared to full questionnaires. For example, Navarro and Griffin (1993) investigated the possible application of this approach for the 2000 Decennial Census, and found that it achieved adequately reliable small-area population estimates as well as reductions in respondent burden. The seminal paper by Raghunathan and Grizzle (1995) demonstrated that a split questionnaire design, coupled with multiple imputation to produce a complete dataset, could obtain estimates (means and regression coefficients) similar to those derived from the full dataset. And more recently, Wedel and Adiguzel (2008) found that a split questionnaire design yielded parameter estimates that were very close to the complete-data estimates, and that respondents who were administered split questionnaires had more favorable reactions to the survey (e.g., shorter perceived duration, lower ratings of boredom and fatigue, etc.) than those who received the complete questionnaire.

2.1.2 Use of Global Items

For surveys that ask a series of detailed questions about a given topic – as the CEQ does for household purchases across a variety of expenditures categories – another option for shortening the length of the interview is to replace some of the detailed questions with global items. Global questions ask about topics at a more aggregated level. For example, rather than asking separate questions about how much a household spent on shoes, pants, shirts, jackets, etc., a global question might simply ask what was spent on clothing, full stop. Global questions could replace detailed questions for the entire sample or for subsets of the sample as way to collect some

information on expenditures without imposing the burden associated with administering the full set of detailed items. The information obtained from the global reports could then be used to derive estimates at a more detailed level (e.g., in a split-questionnaire design in which respondents' global answers are used as inputs to imputation models)¹.

The gains achieved by the use of global questions (i.e., reductions in survey length and potentially respondent burden) have to be considered against the loss of detailed information and its impact on the needs of the survey stakeholders. In addition, the decision to use global questions needs to be informed by an understanding of their impact on respondent error. For example, because global questions lack the specificity of their more detailed counterparts, they may also fail to provide the definitional clarity and retrieval cues required to elicit full and accurate responses (e.g., Conrad and Schober, 2000; Dashen and Fricker, 2001; Hubble, 1995), which in turn may actually increase respondents' burden. Additionally, some studies suggest that global questions produce overestimates and are less reliable than more detailed questions (e.g., Battistin, 2003; Zimmit, 2004). On the other hand, there is empirical evidence that decomposed (detailed) questions also can lead to increases in measurement error, for example when the granularity of the question does not match the way events are stored in memory, or when respondent fatigue induces satisficing (e.g., Belli *et al*, 2000; Menon, 1997; Shields and To, 2000). As these results suggest, the effectiveness of global questions will vary because information about different topics is encoded and stored in memory in different ways (e.g., depending on its salience, frequency of occurrence and use, and its contextual associations).

2.2 Reference Period

The selection of the length of the survey reference period ideally should be based on a number of factors. First, there are analytic and operational considerations. Survey designers must consider the operational costs associated with different reference period designs, and examine these designs in light of the required levels of precision of the estimates. For example, shorter reference periods may necessitate more frequent interviews, which under a fixed budget would result in smaller sample sizes and less statistical precision. Additionally, if shorter reference

¹ Notwithstanding their treatment in subsequent sections of this report, global items likely would not be used by CE as simply direct substitutes for detailed items. However, an in-depth exploration of imputation models incorporating global reports was beyond the scope of this project.

periods result in more frequent interviews, respondent burden and the likelihood of survey nonresponse may increase (e.g., Bradburn, 1978; Apodaca *et al.*, 1998). Second, designers should understand how information on topics covered in the survey is encoded and structured in respondents' memory. Finally, ideally there needs to be an awareness of the error properties associated with the response processes under different reference period implementations, with the selection of a reference period that minimizes those errors.

The literatures on memory, cognition, and survey response processes indicate that for most surveys no single reference period will be optimal for all items. Memory decays over time and, on the whole, short reference periods may improve recall relative to long reference periods (e.g., Miller and Groves, 1985). But, forgetting occurs at different rates for different events (e.g., Bradburn, Rips, and Shevell, 1987). Respondents tend to forget events that are infrequent, irregular, or not salient, all else being equal (e.g., Menon, 1994), so shorter reference periods should aid recall of these events. Longer reference periods may be more appropriate when asking about salient events or regular events that vary little across time or about which the respondent has abstracted and stored some generalized information (e.g., *I usually spend \$10 at the Laundromat on Fridays.*). In either case, as noted earlier, the granularity of the survey question should match the information stored in respondents' memory. For example, if a question using a short reference period asks respondents to enumerate and provide information about individual events but respondents' memory reflects more aggregated, summary-level information about those events (e.g., as is common for frequently occurring but mundane purchases), reporting errors can occur.

The length of the reference period can also impact another source of recall errors known as *forward telescoping*. In forward telescoping respondents erroneously report events as having occurred during the reference period when in fact they occurred prior to it, a phenomenon that generally leads to overreporting². This effect may be caused by respondents misperceiving the length of the reference period (e.g., respondents given a 3-month reference period actually think about the last 3.5 months) or uncertainty about when a target event occurred (Tourangeau, Rips, and Rasinski, 2000). As with errors of omission, there is not a simple relationship between

² The countervailing effects of backward telescoping – placing in-scope events outside the reference period – are thought to be weaker than those of forward telescoping because memory for older events generally has greater temporal imprecision than memory for more recent events.

reference period length and telescoping errors. Early models suggested that shorter reference periods should reduce recall loss and increase forward telescoping, but a number of studies provide evidence that the occurrence and magnitude of telescoping depends on respondents' ability and motivation, the specificity of the question format, the salience of the event being recalled, and the availability of additional temporal cues such as those provided in bounding interviews (e.g., Groves *et al.*, 2004; Neter and Waksberg, 1964; Prohaska, Brown, and Belli, 1998). In one of the few telescoping studies that used verification data to check the veracity of respondents' reports, Huttenlocher, Hedges, and Prohaska (1988) found no differences in the amount of forward telescoping for 'long' and 'short' reference periods (an academic year and a academic quarter, respectively). Respondents in this study were more accurate when reporting for the shorter reference period, but the authors attributed this finding to steeper forgetting curves for older events and to more effortful memory search under the shorter reference period.

3. OVERVIEW OF THE PRESENT STUDY

3.1 Study Design Issues

The primary purpose of this study was to investigate the effects of shortening the length of the CEQ interview (by implementing a split questionnaire design that incorporated global questions) and the length of the CEQ reference period on survey nonresponse, data quality, and respondent burden. In order to achieve these objectives, staff from the BLS Branch of Research and Program Development (BRPD), Office of Survey Methods Research (OSMR), and Branch of Information and Analysis (BIA) formed the Measurement Issues Study (MIS) Team to plan, implement, and analyze data from a small-scale field test of a modified CEQ.

The study utilized a basic experimental design in which respondents were randomly assigned to a control group which received no treatment, a test group that was administered a shortened version of the questionnaire, or a test group that was administered a shortened reference period. As with any field test, an effort was made to mirror as many of the CEQ survey procedures and conditions as possible (e.g., use of a panel design that incorporated a bounding interview, use of CEQ questions and materials), but there were a number of significant departures. First, the budget for the project prevented

in-person data collection so we relied instead on centralized computer-assisted telephone interviews (CATI). As a consequence of changing from in-person to phone-administered interviewing, the Team decided to shorten the overall length of the survey because of concern that study respondents would not accept a phone survey lasting 60 minutes or more on average. To do this, we eliminated questions about a number of CEQ expenditure categories to develop a basic study instrument with a completion duration target of 30 minutes (less for the treatment groups)³.

Another procedural departure for the MIS was that respondents in the shorter reference period condition were interviewed once a month, not once a quarter as in the current CEQ (and in the MIS control group). This was a necessary consequence of the study objectives, since we wanted to aggregate data from the three monthly interviews that used the 1-month reference period and compare those estimates to estimates derived from the control group's standard 3-month reference period, as well as examine potential differences between the control group and this treatment group in nonresponse and respondent burden. Finally, for the shortened interview treatment group we implemented a basic split questionnaire design. We divided our full study questionnaire (i.e., the 30-minute basic version) into one core component and two subcomponents, split the treatment sample into two random subsamples, and then administered each subsample the core component plus one of the subcomponents. In addition to the detailed expenditure questions in their assigned subcomponent, respondents were asked a smaller number of global expenditure questions to augment the loss of detailed information from the remaining expenditure categories (i.e., those covered in the unassigned subcomponent). This allowed us to derive section-level expenditure estimates for all expenditure categories for both of the shortened interview subsamples, and to examine whether the global items produced data of sufficient quality to replace detailed questions.

3.2 Defining the Key Study Outcome Concepts and Measures

The MIS was designed to shed light on three key concepts – respondent burden, data quality, and nonresponse error – defined as follows.

Respondent Burden – Bradburn (1978) identifies four factors that contribute to respondent burden: (1) length of the interview; (2) effort required by the respondent; (3) amount of perceived

³ Additional details about this and other design issues can be found in the Method section of this report.

stress experienced by the respondent; and, (4) periodicity of the interview. We administered questions covering these four factors and used respondents' answers to determine the effect of each of a shorter questionnaire and a one-month reference period on respondent burden.

Data Quality – The CE Program Office operates under the premise that “more is better,” suggesting that respondents who report more expenditures (in terms of both number of items and absolute dollar amount) have higher data quality than those who report fewer expenditures. In the survey methodological literature, the term “data quality” often is used to refer to multiple error sources (e.g., measurement and sampling) and dimensions (e.g., timeliness and accessibility of data). Therefore, the CE Program Office conceptualization assesses only one component of data quality, namely measurement error, and we adopt this perspective here. Given the design of the MIS, the true value of the household expenditures is unknown, so we assessed data quality by examining six indirect indicators: number of expenditure reports, average expenditures, record usage, information book usage, combined expense reporting, and the amount of “don't know” and “refusal” responses.

Nonresponse Error – Unit nonresponse is the failure to obtain any measurements from a sampled unit. In longitudinal surveys (such as the CEQ and the MIS), it can arise in the form of panel attrition if sample members respond to the first and/or several consecutive interviews, but fail to respond to the remaining interviews (these are often referred to as *dropouts*). Nonresponse error occurs when the values of statistics computed based only on respondent data differ from those based on the entire sample data (Groves, et al. 2004). To assess the potential for nonresponse error in this study, we examined response rates, panel attrition rates, and changes in respondent sample composition across the waves of the survey.

3.3 Analysis Overview

Given the substantial design differences between the CEQ and the MIS and the relatively small sample size of this study, we do not address comparisons between the study data and CE production data in this report. Instead, our primary focus is on statistical comparisons between the study control group and each of the study treatment groups on the dimensions of respondent burden, data quality, and

nonresponse error. Specifically, we investigated the following hypotheses suggested by the literature reviewed in Section 2⁴:

- 1a. A shorter interview achieved by splitting the questionnaire will reduce respondent burden.
- 1b. A shorter interview achieved by splitting the questionnaire will increase data quality.
- 1c. A shorter interview achieved by splitting the questionnaire will reduce nonresponse error.
- 2a. The 1-month reference period treatment will increase respondent burden.
- 2b. The 1-month reference period treatment will improve data quality.
- 2c. The 1-month reference period treatment will increase nonresponse error.
3. Global expenditure questions will increase data quality⁵.

Additional details about the analysis methods (including the variables used) and our evaluative criteria can be found in the MIS Analysis Plan (see Appendix I). In the next section of the report we provide further specification about the study design and method.

4. STUDY METHODS

4.1 General

This MIS investigated the aforementioned issues using a truncated CEQ interview and a restricted panel design. There were three test conditions in this study (see Table 1).

Control Group (C)

In the C condition, sample units completed a bounding interview in wave 1. The bounding interview used a 1-month reference period and consisted of items taken from nine sections of the

⁴ For each of the hypotheses the relevant comparison is the study control group.

⁵ ‘Data quality’ here is defined solely as higher reported average expenditure amounts and lower incidence of ‘_don’t know’ and ‘_refusals;’ this study cannot address whether such responses are valid. Other data quality metrics (e.g., number of reports, combined reporting) were not available for the global questions, by definition. In addition, our study design made it impossible to examine the unique effects of global questions on respondent burden or nonresponse error since the global items were a part of the overall shortened questionnaire treatment.

current CEQ instrument plus a “core” set of questions (e.g., demographic items) that were administered across all study conditions. These same C group sample units were contacted again three and six months later to complete two additional interviews using the same “core + nine sections” questionnaire with a 3-month reference period. The C condition paralleled the existing CEQ survey procedures and served as the basis of comparison for the other experimental conditions.

Shortened Questionnaire (SQ)

In the SQ condition, sample units completed the same full bounding interview in wave 1 as the C condition cases, and then were randomly assigned to one of two subsamples that were administered subcomponents of the full questionnaire in waves 2 and 3. Subsample A (SQ-A) received sections 6, 14, 16, 18, and 20, the “core” questions, and a small number of global expenditure questions from sections 9, 12, 13, and 17. Subsample B (SQ-B) received sections 9, 12, 13, and 17, the “core” questions, and a small number of global expenditure questions from sections 6, 14, 16, 18, and 20. Within each subsample group, respondents were split into two groups. One group received the global expenditure questions prior to the detailed expenditure questions, and the other group received the global questions after the detailed items. This counterbalancing allowed us to control for and examine potential order effects stemming from the placement of the global questions. Both the SQ subsample assignments (SQ-A or SQ-B) and the presentation order for the global questions were fixed for waves 2 and 3.

The process of determining which sections to allocate to SQ-A and SQ-B was determined by examining intra-sectional correlations, average duration per section, incidence rate, and potential data quality concerns (e.g., PCE-CE comparisons, imputation/allocation rates). In looking at intra-sectional correlations, we picked sections that had the highest number of “significant” correlations (i.e., 0.1 or greater) with other sections. Then, for each of those sections we identified the section with which it was most highly correlated, and allocated those two sections to different subsamples of our SQ condition. We also attempted to keep the total interview duration similar in the two subsamples. We then examined the incidence rates and CVs for summary variables in the selected sections, and checked our split against one used in Ghosh and Vogt (2000).

Table 1. MIS Test Conditions

| Condition | Wave 1 | | Wave 2 | Wave 3 |
|-------------------------------------|---|--|--|--|
| Control (C) | Bounding Interview (1-month recall) –FULL” Interview (Core + 9 sections) | → | 2 nd Interview (3-month recall) –FULL” Interview (Core + 9 sections) | 3 rd Interview (3-month recall) –FULL” Interview (Core + 9 sections) |
| Shortened Questionnaire (SQ) | Bounding Interview (1-month recall) –FULL” Interview (Core + 9 sections) | Respondents randomly assigned to one of two sub-samples – (a) or (b) → | 2 nd Interview (3-month recall) | 3 rd Interview (3-month recall) |
| | | | (a) Core + sections 1 – 4 ⁶ | (a) Core + sections 1 - 4 |
| | | | (b) Core + sections 5 - 9 | (b) Core + sections 5 - 9 |
| Reference period (RP) | 4 Consecutive 1-month Interviews | | | |
| | Bounding Interview (1-month recall) –FULL” Interview (Core + 9 sections) | 2 nd Interview (1-month recall) –FULL” Interview (Core + 9 sections) | 3 rd Interview (1-month recall) –FULL” Interview (Core + 9 sections) | 4 th Interview (1-month recall) –FULL” Interview (Core + 9 sections) |

Reference period (RP)

In the RP condition, sample units received the same –full” bounding interview that was administered to the wave 1 C and SQ respondents. They then received three consecutive monthly interviews using the same –full” questionnaire with a 1-month reference period (rather than the 3-month reference period used in the C and SQ interviews). Table 2 shows the six possible interview types.

⁶ The section numbers referenced in the four SQ cells of this table do not correspond to the original section numbers in the CEQ; they are for illustrative purposes only (see Table 2 for corresponding CEQ sections).

Table 2: Interview Content by MI Study Treatment Group for Interviews After the Initial Interview

| | Front | Control Card | Housing | Global (9, 12, 13, 17) | Global (6, 14, 16, 18, 20) | Global Before Detailed | 6 | 9 | 12 | 13 | 14 | 16 | 17 | 18 | 20 | Back |
|-------------------------|-------|--------------|---------|------------------------|----------------------------|------------------------|---|---|----|----|----|----|----|----|----|------|
| Control Group | x | x | x | | | | x | x | x | x | x | x | x | x | x | x |
| SQ-A, Version 1 | x | x | x | x | | | x | | | | x | x | | x | x | x |
| SQ-A, Version 2 | x | x | x | x | | x | x | | | | x | x | | x | x | x |
| SQ-B, Version 1 | x | x | x | | x | | | x | x | x | | | x | | | x |
| SQ-B, Version 2 | x | x | x | | x | x | | x | x | x | | | x | | | x |
| Reference period | x | x | x | | | | x | x | x | x | x | x | x | x | x | x |

4.2 Data Collection

Mode and Fielding Period

All data in this study were collected by computer-assisted telephone interviewing (CATI) conducted by the Census Bureau’s Tucson Telephone Center (TTC) staff. The overall fielding period for this study was nine months, beginning June 1, 2010 and ending in February 18, 2011, but varied across the four treatment conditions. The C and SQ groups consisted of three quarterly interviews, each with a 1-calendar-month fielding period. The RP condition consisted of four consecutive monthly interviews with a three-week fielding period each wave (i.e., the 1st through the 21st of each month).

The sample release was staggered for each treatment group such that one-third of the cases assigned to each group were interviewed in month n , one-third in month $n + 1$, and one-third in month $n + 2$. This approach provided a more manageable case workload for TTC and spread out data collection to minimize potential monthly or seasonality effects. This staggered schedule was carried forward throughout all subsequent interviews, based on when the case was originally released and the appropriate reference period for the condition. These sample segments are henceforth referred to as “panels.” For all treatment conditions, a sample unit’s eligibility for

continued participation in its survey panel was contingent on its completion of the first interview; nonrespondents at wave 1 were dropped from the remainder of the study.

Sampling Frame

Census developed a nationally representative sampling frame for a target of 8,100 completed interviews across all study treatments and interview waves. The Demographic Statistical Methods Division (DSMD) used the CEQ reserve cases from the address-based unit frame and matched them to known telephone numbers using a telematch procedure. The address-to-telephone number match enabled survey advance materials to be sent to sample members prior to CATI contact. DSMD achieved a 31% telematch rate; non-matches were excluded from the study sample. Census provided the sampling frame, conducted the telematch procedure, purged the frame of known nonresidential units and nonworking numbers, and drew the sample. More information about the construction of the sampling frame is available in the MIS Sample Plan (Appendix II).

Advance Materials

Prior to the start of each interview wave, the Census Bureau’s National Processing Center (NPC) mailed advance materials to sample members with an address match (see Table 3 for the list and scheduled mail outs for these materials, and Appendix III and IV for the complete documents).

Table 3. MIS Advance Materials Distribution

| 1st Interview | | 2nd - nth Interview |
|---|--|--|
| 1st Mailing | 2nd Mailing | |
| Advance letter (modified Form CE-303-L1); “Fracking Your Spending Behavior” brochure. | Modified Information Booklet (CE-305(C)) | Advance letter (modified Form CE-303-L2); Modified Information Booklet (CE-305(C)) |

The MIS Team worked with DSMD staff to modify the existing CEQ advance letters (Form CE-202-L1 – L5) and the CEQ Information Booklet (CE-305(C)). The two biggest changes to the advance letter were that respondents were asked to participate in the “Consumer Expenditure Telephone Survey” and told that “the average interview takes about 25 minutes”⁷. The revised MIS Information Booklet

⁷ The MIS Team estimated that interviews would take 25 minutes to complete *averaging across treatment groups and interview waves*. Pre-tests indicated that the wave 1 “full” interview took an average of 30 minutes to complete in each condition; waves 2 and 3 C interviews took 27 minutes each; waves 2 – 4 RP

eliminated sections that were not administered in the study and added examples for the global expenditure category questions, but otherwise was identical to the production CEQ Information Booklet.

Within-Household Respondent Selection

As in the CEQ, any adult member of the sampled household age 16 or older could serve as a MIS respondent, but an attempt was made to collect household spending information from the most knowledgeable adult household member (e.g., the owner/renter or their spouse). Changes in respondents between survey waves were allowed and tracked by the instrument when they occurred.

Survey Instrument

Census modified the existing CEQ interview Blaise source code to develop and implement the MIS survey instrument. Table 4 outlines the sections that were taken directly from the CEQ instrument, the set of new questions that the MIS Team provided Census for integration into the Blaise instrument, and the distribution of section/item assignments for the subsamples in the SQ condition. In each interview, respondents were asked about their household purchases in each of the expenditure categories over the

Table 4. MIS Questions: Subject, Origin, and SQ Allocations

| CEQ Section | Subject | Question Origin | SQ Section Allocation |
|--------------------|----------------------------------|------------------------|------------------------------|
| FRONT | Case Management | Existing CEQ | Core |
| CONTROL | Demographics/Roster | Existing CEQ | Core |
| BACK | Contact Information/CHI | Existing CEQ | Core |
| n/a | Rent/Mortgage | New – MIS-provided | Core |
| n/a | Income | New – MIS-provided | Core |
| 6 | Appliances | Existing CEQ | SQ-B |
| 9 | Clothing | Existing CEQ | SQ-A |
| 12 | Vehicle Operating Expenses | Existing CEQ | SQ-A |
| 13 | Insurance (non-health) | Existing CEQ | SQ-A |
| 14 | Hospital/Health Insurance | Existing CEQ | SQ-B |
| 16 | Education Expenses | Existing CEQ | SQ-B |
| 17 | Subscriptions/Entertainment | Existing CEQ | SQ-A |
| 18 | Trips | Existing CEQ | SQ-B |
| 20 | Expense Patterns/Food | Existing CEQ | SQ-B |
| n/a | Global Expenditure Questions | New – MIS-provided | SQ-A & SQ-B |
| n/a | Post-Survey Assessment Questions | New – MIS-provided | Core |

interviews took 25 minutes each; and wave 2 and 3 SQ interviews took an average of 18 minutes to complete.

reference period. In addition, in their final interview (wave 3 for the C and SQ groups, wave 4 for the RP group), respondents were asked a set of post-survey assessment questions (PSAQs) that measured how burdensome they found the survey experience to be, their interest in the survey content, the perceived difficulty of responding to the survey questions, perceived appropriateness of interview length and frequency of survey requests, their estimate of the interview length, and their use of MIS recall aids. Appendices V – VIII provide the full instrument specifications and question wordings for the set of questions that MIS added to the existing CEQ instrument (full instrument specifications are available upon request). Formal systems and verification tests of the instrument were carried out by the MIS Team and Census prior to the start of data collection to ensure that instrument navigation, flow, edits, and database capture and output met study specifications⁸.

Interviewer Staffing, Training, and Monitoring

Approximately 30 TTC CATI interviewers and supervisors worked the MIS data collection over the course of the study fielding period. Census was responsible for the staffing assignments and produced monthly reports for the MIS Team on survey operations. The MIS Team and Census jointly developed an extensive set of interviewer training materials for this study based on the existing CEQ training documentation. However, because the existing materials were designed for CEQ field representatives (not centralized-CATI interviewers) and also did not cover topics and procedures unique to the MIS study, considerable revisions were necessary. Training was developed in two formats: MIS self-study and MIS classroom training. In addition, we provided interviewers with a MIS-tailored Interviewer Manual. The 2-day classroom training was conducted at the TTC facilities during the week of May 10, 2010. The sessions were lead by TTC supervisors and attended by members of the MIS Team who answered study-related questions as required. Throughout the data collection period, TTC supervisors

⁸ Two instrumentation problems were identified after the start of data collection. From September – December 2010, the instrument would not accept values larger than 4 digits in global expenditure variable fields. Some interviewers entered notes about the correct global values in the CATI NOTES field. The authors examined instances that had the maximum allowable 4-digit value, and reviewed the NOTES table to identify global values that needed editing. This occurred in fewer than 10 cases, primarily for education and trip items. In addition, the original MIS study design specified that the order of PSAQ response options would be reversed in half the sample to reduce possible order effects. Census discovered a programming error in the fall of 2009 indicating that this counterbalancing was only done for the Split panel group; the RP and C group respondents only received one response option ordering. We accounted for this problem and coding difference during analyses.

randomly monitored interviewers to ensure that they asked questions as worded, probed effectively, and recorded respondents' answers accurately. In addition, MIS Team members routinely monitored interviews from the remote observation facility at Census, and provided corrective feedback to TTC interviewers when appropriate. Finally, the MIS Team conducted an interviewer survey in September, 2010 to identify any potential problems in survey administration or interviewer understanding of the MIS concepts or procedures. Neither the regular monitoring nor the debriefing survey revealed significant issues that would have negatively impacted survey administration or the quality of the MIS study data⁹.

Data Collection Procedures

MIS cases were assigned to individual interviewers using a WebCATI control system that Census has employed on other CATI surveys (e.g., ACS). At initial contact, the MIS interviewer verified that they had contacted the correct address and attempted to complete the interview. If the respondent agreed to participate, the interviewer proceeded to collect household roster and demographic information (in wave 1; this information was simply verified and updated in subsequent waves) and to administer the expenditure questions appropriate to the MIS treatment group. The control system's set of integrated checks helped to minimize errors (e.g., out-of-range responses, inappropriate skips). In the event of a refusal, the case was reassigned to a supervisory interviewer or refusal conversion specialist; after two refusals WebCATI removed the case from the interview queue and coded it as a noninterview. At the end of each fielding period (three weeks or one month, depending on treatment group), all cases were assigned one of three CATI outcome code types: interview, noninterview, or ineligible for CATI (e.g., incorrect address or telephone numbers). If a sample unit was assigned noninterview or ineligible at the end of the wave 1 fielding period, no further interviews were attempted with this unit.

⁹ There was an unexpected dip in response rates in July 2010 due to insufficient TTC staffing. This did not appear to have differential effects on the study treatment outcomes, and staffing shortages were remedied for the remainder of the data collection period.

5. FINDINGS

5.1 Overall Response Rates and Sample Sizes

For this study, we calculated the response rate for each treatment group and interview wave using the AAPOR RR#4 formula (with 0.33 as the estimated proportion of “eligibility unknown” cases assumed to be eligible):

$$\text{Response Rate} = \frac{\text{Interviews}}{\text{Interviews} + (\text{Refusals} + \text{Others} + \text{Non-contacts}) + e(\text{Unknown Eligibles})}$$

Table 5 shows the final response rates for the treatment conditions, averaging across wave and panel.

Table 5. Final Response Rates by Treatment, Averaging Across Wave and Panel

| Treatment Group | Response Rate | Sample Size (n) |
|------------------------------|---------------|-----------------|
| Control Group (C) | 51.7% | 3,951 |
| Shortened Questionnaire (SQ) | 52.9% | 8,092 |
| SQ-A | 51.2% | 3,906 |
| SQ-B | 54.3% | 4,186 |
| Reference period (RP) | 49.1% | 6,525 |

Table 6 outlines the number of completed interviews by MIS wave.

Table 6. Number of Completed Interviews by MIS Condition and Wave

| Condition | Wave 1 | | | Wave 2 | Wave 3 |
|------------------------------------|---------------|---------------|---------------|---------------|-------------|
| Control (C) | 805 | | | 533 | 477 |
| Shortened Questionnaire (SQ) | 1,686 | | Total | (n = 1,067) | (n = 1,036) |
| | | | SQ-A | 487 | 474 |
| | | | SQ-B | 580 | 562 |
| | Wave 1 | Wave 2 | Wave 3 | Wave 4 | |
| Reference period (RP) | 1,102 | 607 | 606 | 598 | |

5.2 Verifying Random Assignment to Treatment Groups

Wave 1 Sample Composition

We next examined data from wave 1 completed interviews to compare the frequency distributions of CU size, respondent's age, gender, race, education attainment, and housing tenure across the four treatment groups (see Table 7). This served as a manipulation check of our random assignment to the study treatments (i.e., under random assignment of sample units, the group attributes of the different treatment groups should be roughly equivalent). There were no statistical differences between treatment groups in CU size, respondent's age, gender, race, or educational attainment. There was evidence of association between treatment group and housing tenure, $\chi^2(3, n = 3,580) = 15.5, p < .01$. The RP group contained more owners in wave 1 than the C and SQ groups.

| Table 7. Demographic Comparisons by Treatment Condition | | | | |
|--|-----------------------------|---------------|------------|------------|
| (1st Interview) | | | | |
| | TREATMENT GROUP | | | |
| | CONTROL | RECALL | SQA | SQB |
| N | 805 | 1,102 | 774 | 912 |
| | Percent Distribution | | | |
| Number of CU Members | | | | |
| 1 | 25.3% | 25.2% | 26.4% | 21.3% |
| 2 | 38.8 | 37.4 | 39.0 | 41.8 |
| 3+ | 35.9 | 37.4 | 34.6 | 37.0 |
| Age Group | | | | |
| <25 | 1.2 | 1.5 | 1.6 | 1.8 |
| 25-34 | 4.1 | 5.3 | 5.4 | 4.6 |
| 35-64 | 55.8 | 54.3 | 55.9 | 56.4 |
| 65+ | 37.4 | 37.7 | 36.3 | 35.9 |
| Female | 61.1 | 58.6 | 59.7 | 58.1 |
| Race | | | | |
| White | 83.6 | 86.0 | 85.1 | 84.4 |
| Black | 7.7 | 6.4 | 6.3 | 6.8 |
| Asian | 4.2 | 4.1 | 4.1 | 4.1 |
| Marital Status | | | | |
| Married | 60.5 | 61.7 | 61.6 | 64.7 |
| Widowed | 14.9 | 15.2 | 12.0 | 11.6 |
| Divorced | 11.8 | 12.3 | 11.8 | 10.5 |
| Separated | 1.4 | 1.6 | 1.8 | 1.0 |
| Never married | 9.7 | 7.6 | 11.8 | 10.4 |
| Education | | | | |
| < HS | 8.7 | 7.6 | 8.9 | 6.6 |
| HSgrad | 22.1 | 24.6 | 24.7 | 23.7 |
| Some college | 27.7 | 26.2 | 27.1 | 26.0 |
| Undergrad | 22.2 | 21.6 | 20.8 | 24.1 |
| Postgrad | 17.5 | 18.8 | 17.4 | 18.1 |
| Own/Rent* | | | | |
| Own | 87.6 | 89.7 | 86.2 | 91.8 |
| Rent | 10.3 | 8.8 | 12.0 | 6.9 |
| No rent or mort | 2.1 | 1.5 | 1.8 | 1.3 |
| Received Info Book? | | | | |
| No | 33.0 | 29.6 | 32.3 | 32.9 |

* p < .01

Wave 1 Outcome Measures

Since MIS wave 1 interviews were identical for all treatment groups, we can also compare key data quality outcome measures across the groups. Again, given random assignment and the lack of compositional differences between groups observed in Table 7, we would also expect there to be no differences in these measures, and that is what we found (see Table 8). The four treatment groups obtained very similar average expenditures, number of expenditure reports, and incidence of combined expense and ‘don’t know’ reporting.

Table 8. Wave 1 Outcome Measures Verifying Random Group Assignment

| Variable | SQ-A | SQ-B | RP | C |
|----------------------------------|----------|----------|----------|----------|
| \bar{X} total expenditure (\$) | 1,154.50 | 1,299.60 | 1,233.30 | 1,351.00 |
| \bar{X} # of reports | 7.22 | 7.49 | 7.43 | 7.14 |
| \bar{X} # of combined reports | 0.18 | 0.18 | 0.21 | 0.18 |
| \bar{X} # of DK reports | 0.54 | 0.71 | 0.60 | 0.64 |

The availability of MIS frame data additionally allowed us to compare wave 1 respondents and nonrespondents in each of the four treatment groups on characteristics of area poverty, urbanicity, and Census Region. To the extent that these frame variables are correlated with one or more key data items collected in the MIS, differences between nonrespondents and respondents may indicate the potential for nonresponse bias. The bolded cells in Table 9 show the values that reached statistical difference between respondent and nonrespondent on these variables within each condition. The most consistent finding is that households living in high poverty areas appear to be underrepresented in the MIS respondent pool (by 2.3% to 5.2%, depending on the treatment group). In addition, an examination of the relative magnitudes of the difference estimates across the treatments suggests that the C group is at greatest risk for nonresponse bias (i.e., its difference scores are generally larger than those in the other groups). These results, however, do not address the issue of treatment effects on nonresponse bias since the different treatment manipulations were not implemented until wave 2. We explore these analyses in subsequent sections of the report.

Table 9. Nonrespondent – Respondent Differences on Frame Variable in Wave 1

| | SQ-A | | SQ-A DIFF | SQ-B | | SQ-B DIFF | RP | | RP DIFF | C | | C DIFF |
|-------------------------|-----------|------|--------------|-----------|------|--------------|-----------|------|-------------|-----------|------|-------------|
| | Type A | Resp | | Type A | Resp | | Type A | Resp | | Type A | Resp | |
| 20% + in poverty | 13.1% | 9.7% | 3.4% | 12% | 8.8% | 3.2% | 11.5% | 9.2% | 2.3% | 14.5% | 9.3% | 5.2% |
| Urban area | 87.9 | 85.7 | 2.2 | 83.9 | 84.2 | -0.3 | 86.4 | 84.6 | 1.8 | 86.8 | 83.4 | 3.4 |
| Census Region | | | | | | | | | | | | |
| NE | 23.2 | 26.2 | -3.0 | 24.6 | 23.8 | 0.8 | 24.6 | 25.8 | -1.2 | 25.3 | 23.5 | 1.8 |
| MW | 24.0 | 25.6 | -1.6 | 25.6 | 26.3 | -0.7 | 24.8 | 28.3 | -3.5 | 23.9 | 27.7 | -3.8 |
| S | 31.6 | 29.3 | 2.3 | 34.2 | 32.9 | 1.3 | 32.5 | 31.3 | 1.2 | 35.9 | 29.9 | 6.0 |
| W | 21.2 | 18.9 | 2.3 | 15.5 | 17.0 | -1.5 | 18.0 | 14.6 | 3.4 | 14.8 | 18.9 | -4.1 |

5.3 Effect of a Shortened Questionnaire

We next examined the effect of our SQ group on data quality, respondent burden, and nonresponse error. The questionnaire for the two SQ treatment groups was a shortened form of the C questionnaire through the use a split questionnaire design with global questions for a subset of expenditure categories. The reader will recall that in SQ-A the global questions were asked in place of detailed questions in sections 9 (clothing), 12 (vehicle operations), 13 (health insurance), and 17 (subscriptions), and in SQ-B global questions were asked in sections 6 (appliances), 14 (non-health insurance), 16 (education), 18 (trips), and 20A (regular weekly expenditures – i.e., grocery shopping). Since the effects of the global questions may differ by expenditure category, and there may be distinct effects resulting from the unique composition of the detail-global item combinations in each SQ subgroup, we conducted and present separate analyses comparing SQ-A to the Control group and SQ-B to the Control group for each of our analytic dimensions.

Data preparation

Since the source variables required for the C and SQ group comparisons differed depending on whether expenditures were collected from detailed or global questions, we created new analysis variables based on the appropriate source in order to analyze group differences. The mapping between source variables and analysis variables for these group comparisons are documented in Appendix IX (see “Mapping the Variables for the SQ-C Comparisons” and Tables SC1 and SC2).

In addition, before creating the analysis variables, we zero-filled the source variables since the sample mean (i.e., the average expense incurred per category across all sample units) is the statistic of interest.

5.3.1 SQ – Data Quality

Recall that the hypothesis was that the SQ treatment would result in better data quality than the C group (as defined by higher average expenditure amounts overall, a greater number of detailed expenditure reports, and fewer instances of combined reporting and use of DK and REF among the detailed items). Table 10 presents the results of our data quality analysis comparing SQ-A and C groups. We found no significant difference between the SQ-A and C groups in total expenditures in either wave 2 or 3, although in both interviews the SQ-A group produced estimates that were approximately 20% higher than those obtained under in the C (i.e., in the hypothesized direction). Restricting our analysis to estimates derived only from the detail questions, we found that SQ-A and C performed essentially the same in terms of the number of valid reports, combined (or aggregated) reports, and “don’t know or refused” responses (see Table 10). In addition, when we drilled down further to examine section-level comparisons between these two groups for their common detailed sections (6, 14, 18, and 20), we found that they were similar in dollar expenditures, number of valid reports, number of combined (or aggregated) reports, and number of “don’t know/refuse” in both waves 2 and 3 (data not shown; see Appendix X for the section-level expenditure tables for all treatment conditions). Stated differently, the total average expenditure amounts were higher in the SQ-A group (though not significantly so) in both interviews solely because respondents reported higher expenditures in response to the global expenditure questions than they did to the detailed questions. In one sense, the lack of effect of the SQ-A treatment on respondents’ detailed reports – no reduction in the number or amount of reporting, and no increase in combined reporting or “don’t know/refusals” – coupled with the higher overall dollar spending estimates, offers some support to the hypothesis that the SQ treatment should produce better data quality.

Table 10. Comparison of Aggregate Data Quality Measures for SQ-A and CONTROL

| Variable | SQ-A (A) | Control (C) | Difference (A-C) | | | | |
|---|-------------|----------------|---------------------|------------|------------|--------|-----------------------|
| | | | Mean | 95% LCL | 95% UCL | SE | p-value for t-test |
| Wave 2 (quarterly recall) | | | | | | | |
| \bar{X} Total expenditures (\$) | 3,318.00 | 2,752.20 | 565.80 | -58.24 | 1,189.80 | 318.02 | 0.0769 |
| \bar{X} # of valid reports (detailed) | 3.26 | 3.30 | -0.04 | -0.43 | 0.36 | 0.20 | 0.8527 |
| \bar{X} # of combined reports (detailed) | 0.03 | 0.02 | 0.01 | -0.01 | 0.04 | 0.01 | 0.2215 |
| \bar{X} # of DK/REF responses (detailed) | 0.18 | 0.23 | -0.05 | -0.12 | 0.02 | 0.04 | 0.1771 |
| Wave 3 (quarterly recall) | | | | | | | |
| \bar{X} Total expenditures (\$) | 2,968.20 | 2,516.70 | 451.49 | -107.10 | 1,010.00 | 284.61 | 0.1131 |
| \bar{X} # of valid reports (detailed) | 3.41 | 3.57 | -0.16 | -0.55 | 0.23 | 0.20 | 0.4100 |
| \bar{X} # of combined reports (detailed) | 0.03 | 0.01 | 0.01 | -0.01 | 0.03 | 0.01 | 0.2418 |
| \bar{X} # of DK/REF responses (detailed) | 0.18 | 0.24 | -0.06 | -0.14 | 0.01 | 0.04 | 0.0903 |

The trends that existed in the SQ-A to C comparisons were even stronger in the SQ-B to C comparisons (See Table 11). The average overall total expenditure amount was significantly higher in the SQ-B group than the C group in both waves 2 and 3, and SQ-B had more expenditure reports overall than the C group in the detailed question sections for both waves, as well. Moreover, when we examined the indicators of poor data quality (i.e., use of combined reports and “don’t know/refused” reports) – we found no difference between the SQ-B and C groups, and note that in both groups incidence of these behaviors is exceedingly low.

As before, we also examined our data quality metrics at the section level for the detailed sections common to both SQ-B and C (9, 12, 13, and 17). There were no section-level differences in dollar expenditure amounts between SQ-B and C in wave 2, but the SQ-B group did have significantly more expenditure reports for vehicle operating expenses (section 12) (1.52 vs. 1.31, $p < .05$) and entertainment (section 17B) (2.12 vs. 1.83, $p < .01$), as well as significantly more

–don’t know/refused” reports to questions about non-health insurance policies (0.14 vs. 0.08, $p < .05$). Similarly, in wave 3, SQ-B respondents reported significantly higher dollar expenditures for vehicle operations than C respondents (\$263.2 vs. \$177.3, $p < .01$), as well as significantly more reports in this category (1.39 vs. 1.13, $p < .01$). There were no wave 3 differences between SQ-B and C groups in combined reports or –don’t know/refused” reporting. So, here again we see the impact of global items inflating the total expenditure amounts, but there is also evidence that the shortened interview in the SQ-B group had some independent, additive effect.

Table 11. Comparison of Aggregate Data Quality Measures for SQ-B and CONTROL

| Variable | SQ-B (B) | Control (C) | Difference (B-C) | | | | |
|---|-------------|----------------|---------------------|------------|------------|--------|-----------------------|
| | | | Mean | 95% LCL | 95% UCL | SE | p-value for t-test |
| Wave 2 (quarterly recall) | | | | | | | |
| \bar{X} Total expenditures (\$) | 3,955.60 | 2,674.20 | 1,281.40 | 645.17 | 1,917.60 | 324.25 | <.0001 |
| \bar{X} # of valid reports (detailed) | 9.37 | 8.41 | 0.96 | 0.22 | 1.71 | 0.38 | 0.0107 |
| \bar{X} # of combined reports (detailed) | 0.35 | 0.33 | 0.02 | -0.06 | 0.10 | 0.04 | 0.6367 |
| \bar{X} # of DK/REF responses (detailed) | 0.28 | 0.28 | 0.00 | -0.10 | 0.10 | 0.05 | 0.9717 |
| Wave 3 (quarterly recall) | | | | | | | |
| \bar{X} Total expenditures (\$) | 4045.40 | 2,442.90 | 1,602.60 | 909.03 | 2,296.10 | 353.43 | <.0001 |
| \bar{X} # of valid reports (detailed) | 9.99 | 8.98 | 1.01 | 0.16 | 1.86 | 0.43 | 0.0197 |
| \bar{X} # of combined reports (detailed) | 0.27 | 0.31 | -0.04 | -0.12 | 0.03 | 0.04 | 0.2667 |
| \bar{X} # of DK/REF responses (detailed) | 0.18 | 0.20 | -0.02 | -0.10 | 0.05 | 0.04 | 0.5921 |

In addition to comparing average expenditures between the SQ and C treatment groups, we also examined the distributions of expenditure shares between the treatment groups. Expenditure shares are a common way of representing how total expenditures are allocated to the different components of spending. Changes in relative shares can impact the CPI cost weights, so we

wanted to explore potential treatment effects on this measure¹⁰. To test for differences in shares between the SQ and C groups, we used the Chi-square test of homogeneity in proportions (where the null hypothesis is that different treatment groups have the same proportion of consumer units (CUs) in the expenditure categories) and the adjusted Rao-Scott chi-square test statistic which accounts for the complex sample design. The analysis was implemented with Proc SurveyFreq in SAS v 9.1.3, with the CU as the unit of observation (cluster), and the CU's expenditures in each category as the weights. Non-positive expenditures such as those for reimbursements were dropped from the analyses, and for the SQ groups both the detailed and global items served as source variables.

Expenditure shares were calculated as follows:

Aggregate expenditure on category j for group g , X_{jg} , is the sum of expenditures on category j by households i in group g : $X_{jg} = \sum_i x_{ijg}$

Total expenditures for group g : $TX_g = \sum_j X_{jg}$

The relative share of category j for group g : $S_{jg} = \frac{X_{jg}}{TX_g}$

Table 12 shows the expenditures shares for the SQ and C groups for waves 2 and 3. As the distributional differences in Table 11 and the associated chi-square results indicate, there was a large and significant treatment effect for both the SQ-A and SQ-B groups relative to the control group. Although the differences between the SQ and C groups' expenditure shares were relatively small for some expenditure categories (e.g., health insurance and trips in SQA – C waves 2 and 3), they were quite large for others (e.g., health insurance and trips for SQ-B – C in waves 2 and 3). This variability may be due in part to the influence of the global questions (e.g., health insurance and trips were measured by detailed questions in SQ-A, but by global questions in SQ-B).

¹⁰ Our method of calculating expenditure shares differs from the current BLS methods computing expenditure shares. The MIS did not account for various weighting steps used by CPI, and our expenditures base is different because we excluded a number of CEQ sections.

Table 12. Relative Expenditure Shares for SQ and Control Groups for Waves 2 and 3

| | Wave 2 (column %) | | | Wave 3 (column %) | | |
|---|----------------------|-------|---------|----------------------|-------|---------|
| | SQ-A | SQ-B | CONTROL | SQ-A | SQ-B | CONTROL |
| Appliances | 9.8 | 6.9 | 13.1 | 11.6 | 9.1 | 14.8 |
| Clothing | 8.3 | 6.4 | 8.5 | 12.0 | 7.5 | 9.8 |
| Vehicle operations | 9.2 | 6.8 | 8.1 | 10.7 | 6.5 | 7.0 |
| Non-Health insurance | 25.2 | 13.8 | 17.7 | 27.9 | 14.7 | 20.6 |
| Health insurance | 5.0 | 15.5 | 4.6 | 5.9 | 17.9 | 6.1 |
| Education | 26.8 | 21.2 | 24.4 | 16.5 | 23.9 | 22.1 |
| Subscriptions – Entertainment | 6.2 | 6.7 | 9.0 | 7.8 | 6.6 | 9.5 |
| Trips | 3.9 | 19.7 | 7.8 | 1.2 | 10.6 | 2.6 |
| Weekly groceries | 5.6 | 3.0 | 7.0 | 6.4 | 3.1 | 7.4 |
| Total | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| Test of Homogeneity of SQ expenditure relative shares against CONTROL <ul style="list-style-type: none"> • SQA Wave 2: Rao-Scott Chi-Square Test Statistic = 21.23(df=8, p=0.0066) • SQB Wave 2: Rao-Scott Chi-Square Test Statistic =74.60 (df=8, p<0.0001) • SQA Wave 3: Rao-Scott Chi-Square Test Statistic =23.44 (df=8, p=0.0028) • SQB Wave 3: Rao-Scott Chi-Square Test Statistic =70.32(df=8, p<0.0001) | | | | | | |

We also examined respondents' use of recall aids (records and the MIS Information Booklet) in their final interview (questions on recall aid usage were only administered in the SQ and C conditions in wave 3)¹¹. As can be seen in Table 13, the prevalence of information booklet use (before or during the interview) was quite similar across the SQ-A, SQ-B, and C treatments, with slightly more respondents in the two SQ groups than the C group using the Information Booklet to prepare prior to the interview, and slightly fewer in the SQ groups using the Booklet during the interview.¹² Record use trended higher for the C group than either SQ group, but this effect was not significant.

¹¹ We did not have information about respondents' actual use of recall aids, only their self-reported usage.

¹² However, the high missing rate for the later variable – over 50% in all three treatment groups – suggests that there may have been administration problems with this question.

Table 13. Use of Recall Aids for SQ and C Respondents, Wave 3 (column percent shown)

| | SQ-A N=474 | SQ-B N=562 | Control N=477 | Chi-sq p- value |
|---|---------------|---------------|------------------|--------------------|
| Information book use to prepare before interview | | | | 0.0774 |
| missing | 8.9 | 8.0 | 13.0 | |
| Yes | 44.7 | 44.0 | 43.4 | |
| No | 46.4 | 48.0 | 43.6 | |
| Information book use during interview | | | | 0.0437 |
| missing | 55.3 | 56.1 | 56.4 | |
| Yes | 35.0 | 33.3 | 36.3 | |
| No | 9.7 | 10.7 | 7.3 | |
| Record use | | | | 0.4124 |
| missing | 0.21 | 0.36 | 0.21 | |
| Yes | 31.22 | 33.99 | 37.11 | |
| No | 68.57 | 65.66 | 62.68 | |

Finally, we also collected information at the end of the final interview about respondents' use of computer-based financial applications (online and software-based tools for conducting and tracking financial transactions) and asked them whether or not they would be hypothetically willing to share those types of records in lieu of reporting that information in a recall-based expenditure survey if it would significantly shorten the interview. Although respondents' answers to these questions had no bearing on the MIS data, we report this information for all treatment groups in Appendix XI because gaining access to respondents' electronic records is one data-collection strategy being considered in the CE Redesign.

5.3.2 SQ – Nonresponse Properties

To assess the potential for nonresponse error, we began by comparing SQ and C response rates by interview wave and selected characteristics (see Table 14). When each MIS wave is treated as independent, both SQ groups achieved higher response rates in the final wave than the C, though this result failed to reach statistical significance ($p=0.8843$). There also was no indication of a treatment effect in the distribution of response rates by geographic characteristics (Census region, percent of poverty in the area, and urban area).

Table 14. Response Rates for SQ and C by Selected Characteristics

| Characteristics | C | | SQA | | SQB | | Chi-sq p-value |
|---|---------------------------|-----------------|---------------------------|-----------------|---------------------------|-----------------|----------------|
| | No. Eligible ¹ | Response rate % | No. Eligible ¹ | Response rate % | No. Eligible ¹ | Response rate % | |
| Interview wave | | | | | | | 0.8843 |
| 1 | 2,019.75 | 39.9 | 1,973.43 | 39.2 | 2,087.33 | 43.7 | |
| 2 | 756.11 | 70.5 | 714.75 | 68.1 | 856.08 | 67.8 | |
| 3 | 735.05 | 64.9 | 699.37 | 67.8 | 839.03 | 67.0 | |
| Percent of population in poverty in the area | | | | | | | 0.5235 |
| 20% or more | 396.39 | 42.4 | 372.64 | 41.1 | 374.03 | 45.2 | |
| Less than 20% | 3,114.52 | 52.9 | 3,014.91 | 52.5 | 3,408.41 | 55.3 | |
| Census region | | | | | | | 0.0505 |
| North-East | 833.90 | 51.9 | 850.18 | 52.1 | 915.66 | 53.1 | |
| Mid- West | 926.88 | 55.5 | 842.74 | 54.0 | 980.53 | 56.7 | |
| South | 1,127.87 | 47.2 | 1,028.77 | 48.3 | 1,263.60 | 52.8 | |
| West | 622.26 | 54.0 | 665.86 | 51.1 | 622.65 | 55.4 | |
| Urban area | | | | | | | 0.4258 |
| Rural | 546.03 | 54.8 | 453.35 | 57.1 | 599.73 | 55.0 | |
| Urban | 2,964.88 | 51.1 | 2,934.20 | 50.3 | 3,182.71 | 54.2 | |

¹ The proportion of eligibility among cases with “unknown” final disposition was assumed to be e=0.33

We then calculated the cumulative response rates for the C and SQ groups, where the response rate at each wave is conditional on eligibility in wave 1 (see Table 15). This provides a cleaner picture of the potential impact of longitudinal burden on response rates and controls for the initial take rate in each treatment group. The cumulative response rate at wave *t* was computed as:

$$\frac{\text{Interviews at wave } t}{[\text{Interviews} + (\text{Refusals} + \text{Others} + \text{Non-contacts}) + e(\text{Unknown Eligibles})] \text{ at wave } 1}$$

Table 15. SQ and C Group Cumulative Response Rates by Wave

| Wave | Response Rate Conditional on Eligibility at Wave 1 (%) | | |
|------|--|------|------|
| | C | SQ-A | SQ-B |
| 1 | 39.9 | 39.2 | 43.7 |
| 2 | 26.4 | 24.7 | 27.8 |
| 3 | 23.6 | 24.0 | 26.9 |

Recall that all wave 1 interviews were identical and respondents were administered the SQ treatment for the first time in wave 2. Thus, changes in the SQ cumulative response rates

between wave 1 and 2 are unlikely to be the result of a treatment effect. Treatment effects are more likely to occur in wave 3 given respondents' experience with the full wave 2 interview. If hypothesis 1c is correct, we would expect the SQ groups to have lower attrition rates than the C group. That is what we found: the attrition rates between wave 2 and wave 3 for SQ-A (-0.7%) and SQ-B (-0.9%) were substantially lower than the one observed for the C group (-2.8%).

Another way of assessing potential effects of the SQ treatment on nonresponse error is to compare this group's sample composition in the final interview to that of the C group. As discussed in section 5.2, there were essentially no differences between these groups in wave 1 (the only significant difference was a higher proportion of homeowners in SQ-B than in the C and SQ-A groups). Here we found no evidence of differential changes in sample composition between the SQ and C groups over the life of the panel, suggesting that the magnitude of potential nonresponse bias was at least no greater in the SQ conditions than in the C group. Although the SQ-B group continued to have more homeowners than the other two groups in wave 3, the association was not significant at the final wave ($p=0.1142$); the distribution of other characteristics also were similar between the groups.

Table 16 presents the estimated relative nonresponse bias for each expenditure category (and associated 95 percent confidence interval) for nonrespondents at the final interview wave. The following formula was used to compute this estimate for wave 3 nonrespondents:

$$\text{Relative Bias} = \frac{\hat{Z}_{R,j} - \hat{Z}_{T,j}}{\hat{Z}_{T,j}}$$

where:

- $\hat{Z}_{R,j}$ = mean expenditure estimate for expenditure category j in wave 3 from the total sample. Where there was a nonresponse in wave 3 on category j, the expenditure value was substituted from wave 2 (if reported); if it was not reported in wave 2, then the wave 1 value used for wave 3.
- $\hat{Z}_{T,j}$ = mean expenditure estimate for expenditure category j from respondents in wave 3.

Variance estimates of the relative nonresponse bias for each expenditure category were computed using the random groups method (Wolter, 1985), and the data were weighted using the base weights provided by DSMD (see Appendix XII for MIS weighting documentation).

Table 16. Estimated Relative Nonresponse Bias for SQ and C in the Final Wave (Using Base Weights)

| | SQ-A | | | SQ-B | | | CONTROL | | |
|--|-------------------|-------|-------|-------------------|-------|-------|-------------------|-------|------|
| | Relative bias (%) | 95%CI | | Relative bias (%) | 95%CI | | Relative bias (%) | 95%CI | |
| Appliances | 24.2 | 5.4 | 43.0 | 26.0 | 3.3 | 48.7 | 18.8 | 2.4 | 35.2 |
| Clothing | 6.5 | -8.5 | 21.5 | -6.4 | -18.6 | 5.8 | 6.4 | -9.5 | 22.3 |
| Education | 14.0 | -23.7 | 51.6 | 0.6 | -19.8 | 21.1 | 1.4 | -32.4 | 35.2 |
| Health insurance | -3.3 | -33.6 | 27.1 | -52.4 | -58.8 | -46.0 | -22.7 | -44.4 | -1.0 |
| Non-health insurance | -24.6 | -31.0 | -18.3 | 7.3 | -1.8 | 16.4 | -5.3 | -11.7 | 1.0 |
| Weekly groceries | 65.5 | 61.1 | 69.9 | 55.4 | 50.3 | 60.6 | 53.9 | 50.3 | 57.5 |
| Subscriptions & entertainment | -42.8 | -54.6 | -31.0 | -36.1 | -48.3 | -23.8 | -12.7 | -25.7 | 0.3 |
| Trips | 24.8 | -17.2 | 66.8 | -20.0 | -40.2 | 0.1 | 34.4 | -15.3 | 84.1 |
| Vehicle operations | 10.3 | -6.2 | 26.7 | 18.1 | 5.9 | 30.4 | 21.0 | 8.6 | 33.3 |
| Total expenditures | 6.1 | -4.8 | 17.0 | -3.6 | -9.9 | 2.6 | 9.7 | 1.7 | 17.7 |

A negative value for the relative nonresponse bias indicates that by using only data collected from final wave respondents we would underestimate the expenditure (assuming no other sources of error); conversely, a positive value for the relative nonresponse bias suggests that on average, final wave respondents report higher expenditures than nonrespondents. If zero is included in the 95 percent confidence interval of the estimated relative nonresponse bias, it indicates that nonresponse bias is not affecting the estimated expenditure for that item.

The evidence presented in Table 16 suggests that, due to nonresponse, we may be over-estimating final wave SQ-A expenditures for weekly grocery shopping by 66 percent (95CI: 61.1% to 69.9%) and appliances by 24 percent (95CI: 5.4% to 43.0%), but under-estimating non-health insurance by 25 percent (95CI: -31.0% to -18.3%) and subscriptions and entertainment by 43 percent (95CI: -54.6% to -31.0%). However, nonresponse bias does not appear to affect the SQ-A total quarterly expenditures estimate significantly (estimated relative bias of 6.1%, 95CI: -4.8% to 17.0%).

There is some indication of nonresponse bias at the expenditure section level for the SQ-B group, as well. We appear to be over-estimating final wave expenditures on vehicle operations by 18 percent (95CI: 5.9 %to 30.4%), groceries by 55 percent (95CI: 50.3% to 60.6%), and appliances by 26 percent (95CI: 3.3% to 48.7%), but under-estimating health insurance by 52 percent (95CI: -58.8% to -46.0%), and subscriptions and entertainment by 36 percent (95CI: -48.3% to -23.8%). However, again, nonresponse bias does not appear to affect the total expenditure estimate (estimated relative bias of -3.6%, 95CI: -9.9% to 2.6%).

We see a similar pattern of section-level nonresponse bias in the C group, as well. Final wave expenditure estimates for vehicle operating costs, appliances, and groceries appear to be significantly over-estimated in our respondent pool, whereas health insurance estimates appear to be significantly under-estimated. More troubling, the total quarterly expenditure estimate for the C group appears to be positively biased by 9.7 percent (95CI: 1.7% to 17.7%).

We can also examine Table 16 to compare the relative bias measures between the C and SQ groups to assess the effects of our treatment. That is, where there is evidence of bias in the C group, we can look to see if the SQ treatment alleviated, eliminated, or added to the bias. Conversely, we can identify instances where the SQ group may introduce nonresponse bias not present in the C group. For example, nonresponse in both SQ groups appears to exacerbate the bias existing in the C group in the expenditure estimates for appliances, subscriptions and entertainment, and weekly groceries, and to reduce the nonresponse bias in estimates of vehicle operations expenditures and total expenditures.

5.3.3 SQ – Respondent Perceptions of Survey Burden

Table 17 displays the distribution of SQ-A, SQ-B, and C respondent answers to the Post-Survey Assessment Questions (PSAQs) which were designed to capture different dimensions of survey burden. We found a strong association between treatment group and perceived burden, with significantly fewer SQ respondents (27.4% SQ-A and 30.6% SQ-B) saying that they found the survey to be “very burdensome” or “somewhat burdensome” than the C respondents (36.5%). Similarly, SQ respondents were more likely to say that the number of pre-interview calls/contact attempts was “reasonable” (76.0% SQ-A, 73.4% SQ-B) compared to C respondents (68.6%).

And, SQ respondents also were less likely to perceive the final interview to be “too long” (10.1% SQ-A, 8.2% SQ-B) compared to C respondents (17.8%). Finally, we examined the actual length of interview in waves 2 and 3 as another proxy measure of perceived burden. As can be seen in Table 18, there were no differences between treatment groups in wave 1, but the SQA and SQB interviews were significantly shorter than the C interviews in waves 2 and 3 (by more than 6 minutes). Together these results lend strong support to hypothesis 1a that the SQ treatment would reduce respondent burden.

Table 17. Distribution of PSAQ Responses for SQ-A, SQ-B, and C

| | SQA | SQB | Control | Chi-sq p-value |
|---|------|------|---------|----------------|
| Sample size | 474 | 562 | 477 | |
| Interest in survey | | | | 0.4556 |
| missing | 0.8 | 0.9 | 1.1 | |
| Very | 21.7 | 19.2 | 21.0 | |
| Somewhat | 53.8 | 49.3 | 51.4 | |
| Not very | 13.5 | 16.7 | 16.1 | |
| Not at all | 10.1 | 13.9 | 10.5 | |
| Ease in answering survey questions | | | | 0.6503 |
| missing | 0.4 | 0.4 | 0.8 | |
| Easy | 47.9 | 47.2 | 44.2 | |
| Some easy | 33.5 | 37.2 | 38.0 | |
| Some difficult | 16.0 | 14.2 | 15.5 | |
| Very difficult | 2.1 | 1.1 | 1.5 | |
| Survey was burdensome | | | | 0.0083 |
| missing | 3.6 | 2.7 | 0.8 | |
| Very | 1.9 | 3.2 | 4.8 | |
| Somewhat | 25.5 | 27.4 | 31.7 | |
| Not very | 30.6 | 32.4 | 27.7 | |
| Not at all | 38.4 | 34.3 | 35.0 | |
| Number of survey requests for survey panel | | | | 0.9103 |
| missing | 1.5 | 1.4 | 1.5 | |
| Too many | 28.9 | 30.4 | 31.9 | |
| Reasonable | 69.6 | 68.2 | 66.7 | |
| Number of pre-interview calls (contact attempts) | | | | 0.0482 |
| missing | 1.5 | 2.0 | 3.6 | |
| Too many | 22.6 | 24.6 | 27.9 | |
| Reasonable | 76.0 | 73.5 | 68.6 | |
| Perceived length of final survey | | | | <0.0001 |
| missing | 1.1 | 1.1 | 0.2 | |
| Too long | 10.1 | 8.2 | 17.8 | |
| Too short | 0.4 | 0.2 | 0.2 | |
| About right | 88.4 | 90.6 | 81.8 | |

Table 18. Actual Survey Length for the C and SQ Groups by Wave (Minutes)

| | Wave 1 | | | Wave 2 | | | Wave 3 | | |
|-------------------------------------|--------|------|------------------|--------|------|------------------|--------|------|------------------|
| | Mean | SE | p-value for diff | Mean | SE | p-value for diff | Mean | SE | p-value for diff |
| SQ-A | 28.87 | 0.42 | | 21.77 | 0.48 | | 20.58 | 0.4 | |
| SQ-B | 29.51 | 0.37 | | 20.53 | 0.38 | | 19.59 | 0.35 | |
| C | 28.57 | 0.42 | | 28.85 | 0.56 | | 26.69 | 0.6 | |
| Estimated difference* SQ-A – CON | 0.30 | 0.58 | 0.5994 | -7.09 | 0.69 | <0.0001 | -6.11 | 0.68 | <0.0001 |
| Estimated difference* SQ-B – CON | 0.94 | 0.55 | 0.0896 | -8.32 | 0.66 | <0.0001 | -7.10 | 0.65 | <0.0001 |

* from ANOVA

5.4 Effect of a Shortened Reference Period

In this section we report the results of our examination of the effects of a shortened reference period on data quality, nonresponse error, and respondent burden. Specifically, we compared the findings from the C group to those from the RP group, which employed a 1-month reference period and four consecutive monthly interviews. To do so, we first constructed quarterly estimates from the RP group by aggregating across completed interviews in waves 2 through 4 to compare to the quarterly estimates directly obtained from the C group’s wave 2 interview¹³; the reference period months common to both conditions were June 2010 through October 2010. In addition, for the RP group, the “usual weekly expense” variable in section 20 (which includes groceries, alcoholic beverages, and meals away from home) was divided by three after aggregating across the three interviews to account for the “usual weekly” reference period. As before, all analysis variables were zero-filled to compute the sample mean.

5.4.1 RP – Data Quality

Table 19 presents the comparisons between the C and RP groups on each of the key data quality metrics at the aggregate survey (not section) level. The derived (aggregated) RP estimate for overall average expenditure amount (\$) exceeded that of the C group (\$3107.1 vs. \$2752.2, respectively), but the difference was not statistically significant ($p=0.2225$). The RP group did produce significantly more expenditure reports than the C group (difference of 9.9 reports, se

¹³ For the RP-C analyses we dropped RP CUs that had not completed all four interview waves.

0.69, $p < 0.0001$), but it also evinced a greater number of combined reports and “don’t know/refused” responses.

Table 19. Comparison of RP and C Aggregate Data Quality Measures

| Variable | Recall (R) | Control (C) | Difference (R-C) | | | | |
|-------------------------------------|------------|-------------|------------------|---------|---------|--------|--------------------|
| | | | Mean | 95% LCL | 95% UCL | SE | p-value for t-test |
| \bar{X} Expenditures (\$) | 3,107.10 | 2,752.20 | 354.90 | -255.60 | 935.45 | 295.81 | 0.2225 |
| \bar{X} # of valid reports | 21.59 | 11.71 | 9.88 | 8.53 | 11.24 | 0.69 | <.0001 |
| \bar{X} # of combined reports | 0.67 | 0.35 | 0.32 | 0.20 | 0.45 | 0.06 | <.0001 |
| \bar{X} # of DK/refused responses | 4.99 | 1.67 | 3.32 | 3.00 | 3.64 | 0.16 | <.0001 |

We also examined the same measures at the section level (data not shown) and found that the RP group obtained more valid expenditure reports ($p < 0.05$) than the C group for all sections except 18B (trips). However, the RP results for expenditures amounts were mixed. They were significantly higher than the C group for sections 6A (major appliances; \$120.6 vs. \$61.6), 12 (vehicle operation; \$437.7 vs. \$223.7), and 14 (health insurance; \$386.2 vs. \$117.9), but lower for sections 17 (subscriptions/entertainment; \$192.9 vs. \$247.4), 18 (trips; \$41.6 vs. \$215.6), and 20A (regular expenditure patterns; \$173.8 vs. \$192.4). These differences are also reflected in the expenditure shares changes between the RP and C groups (see Table 20).

Table 20. Relative Expenditure Shares for the RP Group

| Expenditure Category | Aggregates (\$) | | Relative Share (% distribution) | |
|--|-----------------|-----------|---------------------------------|---------|
| | Recall | Control | Recall | Control |
| Appliances | 192,291 | 199,976 | 15.3 | 13.1 |
| Clothing | 124,457 | 96,791 | 7.4 | 8.5 |
| Vehicle operations | 119,203 | 183,413 | 14.0 | 8.1 |
| Non-Health insurance | 260,286 | 203,932 | 15.6 | 17.7 |
| Health insurance | 67,282 | 166,461 | 12.7 | 4.6 |
| Education | 358,506 | 284,866 | 21.8 | 24.4 |
| Subscriptions & entertainment | 131,877 | 80,808 | 6.2 | 9.0 |
| Trips | 114,927 | 17,435 | 1.3 | 7.8 |
| Regular weekly expenditures | 102,536 | 72,835 | 5.6 | 7.0 |
| Total expenditures | 1,471,365 | 1,306,517 | 100.0 | 100.0 |

RC: Test of Homogeneity of expenditure relative shares;
Rao-Scott Chi-Square Test Statistic = 57.42(df=8, p < 0.0001)

The RP group also evidence poorer data quality by producing a greater number of combined reports than the C group in sections 9A (clothing) and 13B (non-health insurance), and more “don’t know/refused” responses in sections 13B, 14B, 18A, and 20A, though the incidence of both behaviors was very low overall.

Finally, RP respondents were more likely than CG respondents to use the Information Booklet to prepare for the interview (49.7% vs. 43.4 %; p < 0.05). They were also slightly more likely use it to follow along during the interview (38.8% vs. 36.3%; p < 0.05), but again this variable may be suspect given its high missingness rate (over 50%). There was no difference between the RP and C groups in their prevalence of records use (36.1% vs. 37.1%, respectively; p =.9321).

5.4.2 RP – Nonresponse Properties

Table 21 displays RP response rates by selected characteristics. Treating each wave as independent, the response rate was higher for the C group than RP group in waves 1 – 3 (p=0.0503). There was a significant effect for Census region, as well – the RP group achieved a

Table 21. Response Rates for RP and C Groups by Selected Characteristics

| Characteristic | CONTROL | | RECALL | | Chi-sq p-value |
|---|---------------------------|-----------------|---------------------------|-----------------|----------------|
| | No. Eligible ¹ | Response rate % | No. Eligible ¹ | Response rate % | |
| Interview wave | | | | | 0.0503 |
| 1 | 2,019.8 | 39.9 | 2,873.9 | 38.3 | |
| 2 | 756.1 | 70.5 | 1,044.4 | 58.1 | |
| 3 | 735.1 | 64.9 | 1,007.6 | 60.1 | |
| 4 | n/a | n/a | 1,006.0 | 59.4 | |
| Interview panel | | | | | 0.1385 |
| 1 | 1,189.0 | 51.2 | 2,108.4 | 49.0 | |
| 2 | 1,084.4 | 46.5 | 1,798.2 | 46.5 | |
| 3 | 1,237.6 | 56.7 | 2,025.3 | 51.5 | |
| Percent of population in poverty in the area | | | | | 0.4517 |
| 20% or more | 396.4 | 42.4 | 577.7 | 43.5 | |
| Less than 20% | 3,114.5 | 52.9 | 5,354.2 | 49.7 | |
| Census region | | | | | 0.0067 |
| North-East | 833.9 | 51.9 | 1,503.8 | 48.9 | |
| Mid-West | 926.9 | 55.5 | 1,628.4 | 50.8 | |
| South | 1,127.9 | 47.2 | 1,880.9 | 48.8 | |
| West | 622.3 | 54.0 | 918.9 | 46.9 | |
| Urban area | | | | | 0.5377 |
| Rural | 546.0 | 54.8 | 892.0 | 56.1 | |
| Urban | 2,964.9 | 51.1 | 5039.9 | 47.9 | |

¹ The proportion of eligibility among cases with “unknown” final disposition was assumed to be 0.33

lower response rate in the West and mid-West than the C group (p=0.0067) – but the RP – C groups obtained similar response rates within the other geographic groups. The cumulative response rates (conditional on wave 1 participation) presented in Table 22 reveal that the rate of attrition in the RP group was highest between waves 1 and 2 (-17.3% vs. -13.5% for C) but remained essentially unchanged after that, whereas respondents in the C group continued to attrite between waves 2 and 3.

Table 22. Cumulative Response Rates for RP and C Groups by Wave

| Wave | Response Rate Conditional on Eligibility at Wave 1 (%) | |
|----------|--|------|
| | C | RP |
| 1 | 39.9 | 38.3 |
| 2 | 26.4 | 21.1 |
| 3 | 23.6 | 21.1 |
| 4 | na | 20.8 |

Analysis of the RP and C groups' sample compositions in waves 1 and 4 revealed no significant differences between treatments in either wave, and no significant changes over the life of the panel, suggesting that the RP group's lower wave response rates and steeper attrition rate at wave 2 may not have increased nonresponse error relative to the C group (assuming that the sample composition variables examined are correlated with expenditure reporting).

Table 23 displays the estimated relative nonresponse bias for each expenditure category in the RP and C final interviews. Thus, we are examining relative nonresponse bias for wave 4 monthly expenditure estimates in the RP group, and nonresponse bias for wave 3 quarterly estimates in the C group. The procedure for computing the total sample expenditure estimate at the final wave was the same as before (i.e., final wave nonrespondent estimates are based on the reported value for that item given in most recent available interview and we used sample base weights in the analysis). Again, negative values suggest that we are underestimating the expenditure due to nonresponse, and positive values suggest we are overestimating it. If zero is included in the 95 percent confidence interval of the estimated relative nonresponse bias, it suggests nonresponse bias is not affecting the estimated expenditure for that item.

Table 23. Estimated Relative Nonresponse Bias in the Final Wave for RP and C (base-weighted)

| Expenditure category | RP (monthly expenditure estimates) | | | C (quarterly expenditure estimates) | | |
|--|---------------------------------------|-------|-------|--|-------|------|
| | Relative bias (%) | 95%CI | | Relative bias (%) | 95%CI | |
| Appliances | 17.2 | 0.7 | 33.6 | 18.8 | 2.4 | 35.2 |
| Clothing | -15.2 | -31.8 | 1.3 | 6.4 | -9.5 | 22.3 |
| Education | 46.1 | 30.9 | 61.2 | 1.4 | -32.4 | 35.2 |
| Health insurance | 17.3 | -3.5 | 38.1 | -22.7 | -44.4 | -1.0 |
| Non-health insurance | -25.5 | -37.1 | -14.0 | -5.3 | -11.7 | 1.0 |
| Regular weekly expenditures | -1.7 | -6.0 | 2.6 | 53.9 | 50.3 | 57.5 |
| Subscriptions & entertainment | 5.1 | -9.4 | 19.5 | -12.7 | -25.7 | 0.3 |
| Trips | -20.6 | -59.2 | 18.0 | 34.4 | -15.3 | 84.1 |
| Vehicle operations | 15.8 | 0.1 | 31.5 | 21.0 | 8.6 | 33.3 |
| Total expenditures | 5.6 | -3.4 | 14.6 | 9.7 | 1.7 | 17.7 |

The data indicate that attrition in the RP group may be causing us to overestimate final wave RP expenditures for vehicle operations by 16 percent (95CI: 0.1% to 31.5%), education by 46 percent (95CI: 30.9% to 61.2%), and appliances by 17.2 percent (95CI: 0.7% to 33.6%), but under-

estimate non-health insurance by 26 percent (95CI: -37.1% to -14.0). However, nonresponse bias does not appear to be significantly affecting the RP total monthly expenditure estimate (5.6% estimated relative bias; 95CI: -3.4% to 14.6%). Comparing these results to those obtained for the C group, we find that the RP treatment reduces potential nonresponse bias in the total expenditure estimate as well as for a number of expenditure categories (appliances, health insurance, regular weekly expenses, and vehicle operations), but worsens it for two others (education and non-health insurance).

5.4.3 RP – Respondent Burden

Table 24 displays the distribution of RP and C respondent answers to the Post-Survey Assessment Questions (PSAQs), and there is evidence of significant and strong treatment effects on a number of burden dimensions. For example, 34.4 percent of RP respondents said that the survey was not very / not at all interesting compared to 26.6 percent in the C group. Significantly more RP respondents than C respondents said that the survey questions were easy (52.2% vs. 44.2%). RP respondents also were more likely than C respondents to rate the survey as very burdensome or somewhat burdensome (45.3% vs. 36.5%), and to say that there were too many MIS survey requests (42.1% vs. 31.9%). Moreover, despite the fact that interviews were significantly shorter in the RP group than the C group (more than 4 minutes shorter in waves 2 and 3; see Table 25), proportionally more respondents in the RP group perceived their final interview to be too long though this difference did not reach statistical significance ($p=0.1207$).

Table 24. Distribution of PSAQ Responses for RP and C Groups

| | RP (wave 4) | C (wave 3) | Chi-sq p-value |
|---|-------------------------|-----------------------|---------------------------|
| Sample size | 598 | 477 | |
| | <i>Column percent %</i> | | |
| Interest in survey | | | 0.0478 |
| missing | 0.7 | 1.1 | |
| Very | 16.2 | 21.0 | |
| Somewhat | 48.8 | 51.4 | |
| Not very | 19.4 | 16.1 | |
| Not at all | 14.9 | 10.5 | |
| Ease in answering survey questions | | | 0.0098 |
| missing | 0.0 | 0.8 | |
| Easy | 52.2 | 44.2 | |
| Some easy | 35.5 | 38.0 | |
| Some difficult | 10.9 | 15.5 | |
| Very difficult | 1.5 | 1.5 | |
| Survey was burdensome | | | 0.0157 |
| missing | 0.7 | 0.8 | |
| Very | 7.5 | 4.8 | |
| Somewhat | 37.8 | 31.7 | |
| Not very | 27.4 | 27.7 | |
| Not at all | 26.6 | 35.0 | |
| Number of survey requests for survey panel | | | 0.0015 |
| missing | 2.0 | 1.5 | |
| Too many | 42.1 | 31.9 | |
| Reasonable | 55.9 | 66.7 | |
| Number of pre-interview calls (contact attempts) | | | 0.0968 |
| missing | 1.7 | 3.6 | |
| Too many | 30.9 | 27.9 | |
| Reasonable | 67.4 | 68.6 | |
| Perceived length of final survey | | | 0.1207 |
| missing | 0.5 | 0.2 | |
| Too long | 20.9 | 17.8 | |
| Too short | 1.2 | 0.2 | |
| About right | 77.4 | 81.8 | |

Table 25. Actual Survey Length (minutes) for RP and C Groups by Interviewer Wave

| | Wave 1 | | | Wave 2 | | | Wave 3 | | | Wave 4 | | |
|---------------------------------------|--------|------|------------------|--------|-----|------------------|--------|-----|------------------|--------|-----|------------------|
| | Mean | SE | p-value for diff | Mean | SE | p-value for diff | Mean | SE | p-value for diff | Mean | SE | p-value for diff |
| RP | 28.9 | 0.3 | | 22.4 | 0.4 | | 22.3 | 0.5 | | 21.6 | 0.4 | |
| C | 28.6 | 0.4 | | 28.9 | 0.6 | | 26.7 | 0.6 | | | | |
| Estimated difference* (RP – C) | 0.3 | 0.53 | 0.6015 | -6.5 | 0.7 | <0.0001 | -4.4 | 0.6 | <0.0001 | | | |

* from ANOVA

5.4 Effect of a Global Expenditure Questions on Data Quality

The third research objective of this study was to examine whether global questions can solicit data of sufficient quality to replace detailed questions. The use of global questions in the SQ treatment condition provided an opportunity to address this question through an analysis of expenditure amounts and reporting rates of “don’t know/refused” responses collected through global questions and detailed questions. In this section, we begin by revisiting in more detail the differences between the detailed-based expenditure estimates obtained in the C group and the global-based estimates observed in the SQ-A and SQ-B groups. We then examine comparisons of estimates from a single SQ group (formed by simply concatenating SQ-A and SQ-B records) and the C group both at the section level and overall. Finally, as an ancillary analysis, we investigate whether the order of the block of global questions and the block of detail questions had an effect on expenditure estimates within each SQ group.

In order to compare global and detail estimates, we created analysis variables that were sourced either from the detailed items in the C and SQ groups, or from the global variables in the two SQ groups. The mapping between source variables and analysis variables for these group comparisons are documented in Appendix IX (see “Mapping the Variables for Global vs. Detailed Question Comparisons” and Tables SC3 and SC4). As before, we zero-filled the source variables.

Tables 26 and 27 display expenditure estimates derived from the detailed questions in the C group and those derived from the global questions about the same category in the SQ-A and SQ-B groups, respectively. Global questions elicited significantly higher expenditure reports in five of the ten expenditure categories in wave 2 (six in wave 3), and a significantly lower expenditure estimate of subscriptions in both waves (see bolded cells). There was no significant effect in either wave of question form (detailed vs. global) on expenditure estimates of appliances, education, and weekly groceries.

Table 26. Detailed-Based (C) and Global-Based (SQ-A) Expenditure Estimates for Waves 2 and 3

| | Global (SQ-A) | Detailed (C) | Difference | 95LCI diff | 95UCI diff | SE diff | p-value for t-test |
|-----------------------------|----------------------|---------------------|-------------------|-------------------|-------------------|----------------|---------------------------|
| Wave 2 | | | | | | | |
| Clothing | 274.9 | 233.5 | 41.4 | -12.6 | 95.4 | 27.5 | 0.1378 |
| Vehicle operations | 304.6 | 223.7 | 80.9 | 19.4 | 142.5 | 31.4 | 0.0112 |
| Non-health insurance | 839.1 | 488.3 | 350.8 | 232.0 | 469.5 | 60.5 | <.0001 |
| Subscriptions | 37.5 | 121.8 | -84.3 | -120.0 | -48.6 | 18.2 | <.0001 |
| Entertainment | 168.7 | 125.6 | 43.1 | 2.6 | 83.7 | 20.7 | 0.0406 |
| | | | | | | | |
| Wave 3 | | | | | | | |
| Clothing | 357.8 | 249.2 | 108.6 | 8.5 | 208.7 | 51.0 | 0.034 |
| Vehicle operations | 319.1 | 177.3 | 141.8 | 59.6 | 224.1 | 42.0 | 0.0008 |
| Non-health insurance | 832.0 | 520.8 | 311.2 | 183.6 | 438.8 | 65.0 | <.0001 |
| Subscriptions | 45.3 | 122.6 | -77.3 | -107.3 | -47.2 | 15.3 | <.0001 |
| Entertainment | 188.1 | 118.4 | 69.7 | 13.6 | 125.6 | 28.5 | 0.0152 |

Table 27. Detail-Based (C) and Global-Based (SQ-B) Expenditure Estimates for Waves 2 and 3

| | Global (SQ-B) | Detailed (C) | Difference | 95LCI diff | 95UCI diff | SE diff | p-value for t-test |
|-------------------------|----------------------|---------------------|-------------------|-------------------|-------------------|----------------|---------------------------|
| Wave 2 | | | | | | | |
| Appliances | 271.3 | 360.8 | -89.5 | -186.2 | 7.3 | 49.3 | 0.0748 |
| Health insurance | 613.4 | 117.9 | 495.5 | 422.0 | 568.9 | 37.4 | <.0001 |
| Education | 839.5 | 672.6 | 166.9 | -289.7 | 623.3 | 232.6 | 0.4704 |
| Trips | 779.0 | 215.6 | 563.4 | 372.0 | 754.8 | 97.6 | <.0001 |
| Weekly groceries | 120.0 | 114.4 | 5.6 | -4.2 | 15.3 | 5.0 | 0.2649 |
| | | | | | | | |
| Wave 3 | | | | | | | |
| Appliances | 368.8 | 375.4 | -6.6 | -113.7 | 100.5 | 54.6 | 0.9040 |
| Health insurance | 725.7 | 140.4 | 585.3 | 447.4 | 723.2 | 70.3 | <.0001 |
| Education | 966.5 | 559.8 | 406.7 | -139.1 | 952.6 | 278.2 | 0.1289 |
| Trips | 430.2 | 65.0 | 365.2 | 256.3 | 474.1 | 55.5 | <.0001 |
| Weekly groceries | 124.7 | 114.0 | 10.7 | -0.6 | 22.0 | 5.8 | 0.0642 |

Another way to look at the effects of global questions is to examine their impact on the final survey estimates. In a typical SQ design, data from the various questionnaire subcomponents are combined in some way to produce a single dataset for analysis. Although systematic examination and evaluation of the methods for combining SQ files was beyond the scope of this project, we were interested to see how estimates from the C group would compare to those from a single, combined SQ dataset. To create this data files, we simply concatenated the SQ-A and SQ-B files and summed across responses from the detailed and global questions in each expenditure category. We then calculated expenditure estimates and counts of ‘don’t know/refuse’ for this combined SQ file, and compared them to those obtained in the C group (see Tables 28 and 29).

Table 28: Comparison of Combined-SQ and C Group Estimates of Quarterly Expenditures (\$)

| Expenditure category | Wave 2 | | | | | Wave 3 | | | | |
|-------------------------------|---------------|---------------|--------------|---------|----------------|---------------|---------------|---------------|---------|----------------|
| | SQ | C | Diff (SQ-C) | SE diff | p-value T-test | SQ | C | Diff (SQ-C) | SE diff | p-value T-test |
| Appliances | 296.2 | 360.8 | -64.6 | 43.1 | 0.1765 | 358.7 | 375.4 | -16.7 | 45.1 | 0.7225 |
| Clothing | 263.1 | 233.5 | 29.6 | 22.2 | 0.1601 | 328.7 | 249.2 | 79.5 | 38.0 | 0.0138 |
| Vehicle operations | 285.7 | 223.7 | 62.1 | 26.3 | 0.0108 | 288.8 | 177.3 | 111.5 | 33.4 | <.0001 |
| Non-health insurance | 679.0 | 488.3 | 190.7 | 49.2 | <.0001 | 704.2 | 520.8 | 183.4 | 53.5 | <.0001 |
| Health insurance | 405.7 | 117.9 | 287.8 | 32.5 | <.0001 | 468.0 | 140.4 | 327.6 | 55.4 | <.0001 |
| Education | 862.9 | 672.6 | 190.3 | 201.6 | 0.3281 | 749.1 | 559.8 | 189.3 | 218.2 | 0.3363 |
| Subscriptions & entertainment | 238.1 | 247.4 | -9.3 | 25.2 | 0.7233 | 251.0 | 241.0 | 10.0 | 28.3 | 0.6898 |
| Trips | 483.1 | 215.6 | 267.5 | 77.2 | 0.0008 | 249.7 | 65.0 | 184.7 | 42.4 | <.0001 |
| Groceries | 50.5 | 114.4 | -63.9 | 4.4 | <.0001 | 51.9 | 114.0 | -62.1 | 4.5 | <.0001 |
| Total expenditures | 3564.3 | 2674.2 | 890.1 | 284.9 | 0.0001 | 3449.9 | 2442.9 | 1007.1 | 297.0 | 0.0001 |

Significant differences between the combined-SQ and C group estimates are bolded in both tables. Total reported expenditures were higher for the combined-SQ group than the C group in both waves 2 and 3 (by \$890 and by \$1007, respectively). At the section level, wave 2 and wave 3 estimates from the single SQ data file were significantly higher than those from the C group in the categories of vehicle operations, health insurance, non-health insurance, and trips. In addition, although there was no difference between the combined-SQ and C groups in wave 2 estimates of clothing expenses, the SQ estimate was significantly higher in wave 3.

Each of these results is consistent with what we found earlier when we looked separately at the SQ-A and SQ-B comparisons to the C group. One departure from those findings here is that estimates of weekly grocery spending were significantly lower in the combined-SQ group than the C group. The significant differences we noted in our earlier SQA-C and SQB-C comparisons of higher entertainment estimates and lower subscriptions estimates in the SQ group were not evident when using the combined-SQ data.

We next examined the incidence of “don’t know/refused” responses as an indicator of data quality. However, since the number of questions asked about each expenditure category varied by question format (i.e., there were more opportunities for DK/REF response with detailed questions than global items), we created a section-level flag to indicate their presence or absence. As shown in Table 29, the proportion of “don’t know/refused” responses trended lower in the combined-SQ group than the C group, with significant differences observed for non-health insurance and weekly grocery spending in waves 2 and 3, and subscriptions and memberships in wave 2.

Table 29. Comparison of Combined-SQ and C Group: Proportion of “Don’t Know/Refused”*

| Expenditure category | Wave 2 | | | | | Wave 3 | | | | |
|-------------------------------|---------------|---------------|----------------|---------|----------------|---------------|---------------|----------------|---------|----------------|
| | SQ | C | Diff (SQ-C) | SE diff | P-value T-test | SQ | C | Diff (SQ-C) | SE diff | P-value T-test |
| Appliances | 0.0112 | 0.0206 | -0.0090 | 0.0063 | 0.1775 | 0.0068 | 0.0168 | -0.0100 | 0.0055 | 0.1189 |
| Clothing | 0.0375 | 0.0525 | -0.0150 | 0.0226 | 0.5529 | 0.0125 | 0.0273 | -0.0150 | 0.0082 | 0.1338 |
| Vehicle operations | 0.0169 | 0.0356 | -0.0190 | 0.0101 | 0.1288 | 0.0087 | 0.0231 | -0.0140 | 0.0069 | 0.0936 |
| Non-health insurance | 0.1078 | 0.0844 | 0.0234 | 0.0202 | 0.2156 | 0.0840 | 0.0901 | -0.0060 | 0.0184 | 0.7372 |
| Health insurance | 0.1097 | 0.1895 | -0.0800 | 0.0227 | 0.0011 | 0.1014 | 0.2055 | -0.1040 | 0.0234 | <.0001 |
| Education | 0.0056 | 0.0094 | -0.0040 | 0.0048 | 0.4477 | 0.0039 | 0.0042 | -0.0003 | 0.0040 | 0.9302 |
| Subscriptions & entertainment | 0.0337 | 0.0675 | -0.0340 | 0.0110 | 0.0057 | 0.0222 | 0.0252 | -0.0030 | 0.0083 | 0.7225 |
| Trips | 0.0028 | 0.0019 | 0.0009 | 0.0027 | 0.7061 | 0.0029 | 0.0042 | -0.0010 | 0.0032 | 0.7029 |
| Groceries | 0.5511 | 1.1614 | -0.6100 | 0.0516 | <.0001 | 0.4884 | 1.109 | -0.6210 | 0.0525 | <.0001 |

* A “dk/refused” is flagged for a section if there is one or more “dk/refused” response in any subsection.

Finally, as noted in Section 4.1, the MI study design allowed us to examine potential order effects stemming from the placement of the block of global items within each SQ subsample. Half of

the respondents in each SQ group received the global expenditure questions prior to the detailed expenditure questions, and half received the global questions after the detailed items (with the order of presentation fixed across interview waves). We investigated the impact of the global-detail ordering on respondents' reported expenditure amounts (for both global and detailed items), number of expenditure reports, DK/REF reports, and use of combined reports for detailed items. Tables 30 and 31 present the results of the global-detail order analyses for the SQ-A and SQ-B groups, respectively.

Table 30. Effect of Global-Detail Ordering on Key Outcome Measures – SQ-A

| | Wave 2 | | | Wave 3 | | |
|---|-----------------|-----------------|--------------|-----------------|-----------------|--------------|
| | Global Q Order | | Diff | Global Q Order | | Diff |
| | 1 st | 2 nd | | 1 st | 2 nd | |
| \bar{X} Expenditures on Global Items (\$) | 1,906.8 | 1,382.7 | 524.1 | 2,210.6 | 1,333.2 | 877.4 |
| \bar{X} Expenditures on Detailed Items (\$) | 1,985.9 | 1,147.4 | 838.5 | 1,359.9 | 804.4 | 555.5 |
| \bar{X} # of valid reports | 3.59 | 2.99 | 0.60 | 3.58 | 3.26 | 0.32 |
| \bar{X} # of DK/refused responses | 0.20 | 0.17 | 0.03 | 0.19 | 0.16 | 0.03 |
| \bar{X} # of combined reports | 0.05 | 0.01 | 0.04 | 0.03 | 0.02 | 0.01 |

Table 31. Effect of Global-Detail Ordering on Key Outcome Measures – SQ-B

| | Wave 2 | | | Wave 3 | | |
|---|-----------------|-----------------|--------------|-----------------|-----------------|--------|
| | Global Q Order | | Diff | Global Q Order | | Diff |
| | 1 st | 2 nd | | 1 st | 2 nd | |
| \bar{X} Expenditures on Global Items (\$) | 2,908.6 | 2,341.4 | 567.2 | 3,008.1 | 2,212.3 | 795.8 |
| \bar{X} Expenditures on Detailed Items (\$) | 1,381.7 | 1,284.1 | 97.61 | 1,542.1 | 1,313.8 | 228.3 |
| \bar{X} # of valid reports | 9.23 | 9.51 | -0.28 | 9.89 | 10.10 | - 0.21 |
| \bar{X} # of DK/refused responses | 0.21 | 0.34 | -0.13 | 0.17 | 0.19 | -0.02 |
| \bar{X} # of combined reports | 0.40 | 0.31 | 0.09 | 0.28 | 0.26 | 0.02 |

The bolded cells in these tables indicate significant differences ($p < .05$) between the outcome measures based on the order of the global item administration. As shown in Table 30, when the block of global questions came before the block of detailed questions, SQ-A respondents reported higher expenditure estimates for both the global and detailed items than when the block of detailed questions were administered first, and this result was obtained in both waves 2 and 3. In addition, in wave 2, SQ-A respondents who were asked the global items first reported significantly more detailed expenditure reports than those who were asked the block of detail items first. Table 31 reveals a similar though non-significant trend for the expenditure amounts reported by SQ-B respondents in waves 2 and 3 (i.e., asking global items first was associated with higher expenditures; rows 1 and 2). In addition, SQ-B respondents who were asked global questions first provided significantly fewer DK/REF responses than those who were asked the detailed items first.

6. DISCUSSION

6.1 Summary

The objectives of this study were to: (1) assess the effects of administering a shorter CEQ instrument on respondent burden, data quality, and nonresponse error; (2) examine the impact of using a one-month (versus the current three-month) reference period on respondent burden, data quality, and nonresponse error; and (3) evaluate the quality of data collected from global, as opposed to, detailed questions on expenditures. To achieve these objectives, the MIS study implemented an experimental design in which respondents were randomly assigned to a control group (C) which received no treatment, a test group that received a shortened questionnaire design (SQ), or a test group that was administered a shortened reference period (RP).

The following is a summary of the study findings as they pertain to the set of hypotheses laid out in Section 3.3 of this report.

Did a shorter interview achieved by splitting the questionnaire reduce respondent burden?

Yes. Respondent burden was significantly lower in the SQ groups than C group. SQ respondents perceived the survey to be less burdensome and of appropriate duration and frequency, compared to the control group respondents. SQ interviews were 6 minutes shorter than C interviews on average.

Did a shorter interview achieved by splitting the questionnaire increase data quality?

Somewhat. Data quality moderately improved under the SQ treatment relative to the control condition. Both SQ subsamples (SQ-A and SQ-B) produced total expenditure estimates that were higher than the control estimates, although only the SQ-B group reached statistical significance. In addition, the SQ-B group reported significantly more expenditure reports than the C group. The SQ treatment did not substantively impact the incidence of negative respondent behaviors (i.e., combined reports, “don’t know/refusals”) or the use of recall aids or records.

Did a shorter interview achieved by splitting the questionnaire reduce nonresponse error?

The effects of the SQ treatment on indicators of nonresponse error were minor, varied, but generally positive. Response rates examined independently by interview wave revealed no treatment effect (i.e., they were comparable for the SQ and C groups at each wave). However, the SQ groups attained significantly lower attrition rates between wave 2 and wave 3 than the C group (0.7% and 0.9% for the SQ groups vs. 2.8% for C). The final wave cumulative response rate (i.e., conditioned on participation in wave 1) also was higher in the SQ groups than the C group. There were no observed differences in sample composition between the SQ and C groups in the final wave. Finally, compared to the C group, the SQ treatment reduced the relative nonresponse bias in total expenditures estimates as well as vehicle operations expenditures estimates, though there was evidence that it also exacerbated the bias existing in a few of the C group expenditure estimates.

Did the 1-month reference period treatment increase respondent burden?

Yes. There were significant and strong RP treatment effects on a number of respondent burden dimensions. Significantly more RP than C respondents said that the survey was *‘not very / not at all interesting’* and *‘very / somewhat burdensome,’* and that there were *‘too many’* survey requests. In contrast, more RP respondents than C respondents said that the survey questions were *‘easy.’* Moreover, despite the fact that actual interview durations were significantly shorter in the RP group than the C group (by more than 4 minutes in waves 2 and 3), proportionally more respondents in the RP group perceived their final interview to be *‘too long.’*¹⁴

Did the 1-month reference period treatment improve data quality?

Evidence on the effect of RP treatment on data quality was mixed. There were some indications that RP improved data quality. For example, respondents in the RP group did report significantly more valid expenditure reports, and the total expenditures estimate in this group was higher than the C estimate (but not significantly so). In addition, RP respondents were more likely than C respondents to use the Information Booklet to prepare for the survey in advance. On the other hand, RP respondents were significantly more likely than the C respondents to engage in

¹⁴ The burden in the RP condition stems from both longitudinal burden (the burden associated with being interviewed multiple times in relative quick succession) and the cognitive burden associated with the reference period. The MI study design did not permit an examination of the effects of these two sources separately.

undesirable reporting behaviors (e.g., use of combined item reporting and “don’t know” and/or “refused” responses). In particular, the RP group was higher in both of these undesired reporting behaviors for section 9 (clothing), a section that is already problematic in the current instrument using a 3-month recall. The RP group had nearly three times as many “don’t know/refusals” as the C group; represented as a percent of the average total number of reports, the RP group’s rate of DK/REF was 23% compared to 13% for the C group. There was no difference in use of records between the RP and C groups.

Did the 1-month reference period treatment increase nonresponse error?

The RP treatment had a negative impact on survey participation. Response rates examined independently by wave and conditional on wave 1 participation were lower for the RP group than the C group in waves 2 and 3. The attrition rate between wave 1 and 2 also was substantially higher for the RP group (17.2% vs. 13.5%), possibly due to the RP group’s tighter fielding period and/or the saliency of respondents’ prior wave (negative) experience.

Overall, it does not appear that RP treatment worsened any potential nonresponse bias that may have existed in the C group. The sample of respondents in RP and C were generally similar in distribution on the selected demographic characteristics. In addition, the RP data showed less relative nonresponse bias in total expenditure estimates and estimates of health insurance spending and regular weekly expenditures compared to the C group. However, the RP group showed worse nonresponse bias for estimates of education and appliances expenditure (which were over-estimated) and non-health insurance expenditures (which was underestimated).

Did global expenditure questions increase data quality?

Global-based spending estimates were significantly higher than detailed-based estimates in six of the ten expenditure categories examined in this study (clothing, vehicle operations, non-health insurance, health insurance, entertainment, and trips), and significantly lower in only one (books/subscriptions). We present evidence that the use of global questions reduced levels of “don’t know / refused” responses, as well.

6.2 Limitations

As noted earlier in this report, the prohibitive cost of conducting in-person data collection impelled us to rely on centralized computer-assisted telephone interviews (CATI). As a consequence of this mode change, we also eliminated sections of the survey to shorten its overall length. Changes to mode, length, and question context impact the response process and associated errors, so it is likely that some of our results would have been different under a design closer to that of the CEQ. We also were restricted by the project budget to a relatively small sample size. This reduced our power to detect some treatment effects and prevent us from examining effects at lower levels of analysis (below the section-level).

In addition, as noted elsewhere in this report, there were potential limitations with some of our analytic techniques and outcome measures. For example, we had no direct way to assess the ‘more is better’ hypothesis of data quality because we did not have true values on expenditures. This limitation is not unique to the MI study, but it deserves underscoring as CE embarks on redesign efforts that will look to measures of data quality improvements. In addition, the nonresponse bias analyses we conducted should only be viewed as suggestive. The method involved carrying forward the last available observation, and this may be tenuous for expenditures that are unlikely to be recurring in two subsequent interview periods. Moreover, our nonresponse bias estimates may provide a worse-case scenario since the data in this study were not subject to the same rigorous nonresponse adjustment procedures as utilized in the CE production environment.

6.3 Recommendations

The results of this study suggest that a SQ design may hold promise in a redesigned CEQ. Additional research is needed to determine the optimal length of a shortened survey, composition of questionnaire splits (in terms of their statistical properties and impact on respondent processes/errors), and dataset construction and analysis methods. We are less sanguine about the adoption of a 1-month reference period, given the concomitant need for conducting monthly interviews, and our findings on the negative effects of this design on response rates and respondent burden. That said, the optimal reference period likely will vary across expenditures, and additional laboratory research is needed in this area. Similarly, we recommend additional research (e.g., cognitive studies, controlled experiments, validation studies) on respondents’ use of global questions. Finally, this study underscores the benefits of examining a variety of quality

metrics. We recommend that as CE moves forward with its redesign efforts, it considers a range of evaluative criteria (e.g., level of reporting, respondent burden index, indirect indicators of data quality, etc.), and establishes a panel in the revised design that offers sufficient power to detect statistical differences and to track changes in these metrics.

References

- Apodaca, R. , Lea, S. and Edwards, B. (1998). "The Effect of Longitudinal Burden on Survey Participation." Proceedings of the 1998 American Statistical Association, Survey Research Methods Section.
- Backor, K., Golde, S. and Norman, N. (2007). "Estimating Survey Fatigue in Time Use Study." Paper presented at the 2007 International Association for Time Use Research Conference, Washington, DC.
- Battistin, E. (2003) Errors in Survey Reports of Consumption Expenditures. Working Paper W03/07, Institute for Fiscal Studies, London.
- Belli, R., Schwarz, N., Singer, E. and Talarico, J. (2000). "Decomposition Can Harm the Accuracy of Behavioural Frequency Reports." *Applied Cognitive Psychology*, 14, 295-308.
- Bradburn, N. M. (1978). "Respondent Burden." Proceedings of the Section on Survey Research Methods, American Statistical Association, 35—40.
- Bradburn, N.M., Rips, L.J., & Shevell, S.K. (1987). "Answering autobiographical questions: The impact of memory and inference on surveys." *Science* , 236.
- Branden, L., Gritz, R. M., and Pergamit, M. R. (1995). "The Effect of Interview Length on Nonresponse in the National Longitudinal Survey of Youth," Proceeding of the 1995 Census Bureau Annual Research Conference, U.S. Bureau of the Census, pp. 129-154.
- Champion, D. and Sear, A. (1969). "Questionnaire Response Rate: A Methodological Analysis," *Social Forces*, 47, 335-339.
- Collins, M., Sykes, W., Wilson, P., and Blackshaw, N. (1988). "Nonresponse: The UK experience." In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II and J. Waksberg (Eds.), *Telephone Survey Methodology* (pp. 213-231). New York: Wiley.
- Conrad, F.G., & Schober, M.F. (2000). "Clarifying question meaning in a household telephone survey." *Public Opinion Quarterly*, 64, 1-28.

- Dashen, M. and Fricker, S. (2001). "Understanding the cognitive processes of open-ended categorical questions and their effect on data quality." *Journal of Official Statistics*, 17, 457–477.
- Dillman, D., Sinclair, M. and Clark, J. (1993). "Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys," *Public Opinion Quarterly*, 57, 289-304.
- Galesic, M. (2006). "Dropouts on the Web: Effects of Interest and Burden Experienced During an Online Survey." *Journal of Official Statistics*, 22, 313-328.
- Ghosh, D. and Vogt, A. (2000). "Determining an optimal split for a lengthy questionnaire." Paper presented at the American Statistical Association (ASA) meetings, Indianapolis.
- Gonzalez, J. M. and Eltinge, J. L. (2007a). "Multiple Matrix Sampling: A Review." Proceedings of the Section on Survey Research Methods, American Statistical Association, 3069–75.
- Gonzalez, J. M. and Eltinge, J. L. (2007b). "Properties of Alternative Sample Design and Estimation Methods for the Consumer Expenditure Surveys." Paper presented at the 2007 Research Conference of the Federal Committee on Statistical Methodology, Arlington, VA, November, 2007.
- Gonzalez, J. M. and Eltinge, J. L. (2008). "Adaptive Matrix Sampling for the Consumer Expenditure Quarterly Interview Survey." Proceedings of the Section on Survey Research Methods, American Statistical Association, 2081–8.
- Gonzalez, J. M. and Eltinge, J. L. (2009). "Imputation Methods for Adaptive Matrix Sampling." A poster presented at the 2009 Joint Statistical Meetings, Washington, DC, August.
- Groves, R., Singer, E., and Corning, A. (1999). "A Laboratory Approach to Measuring the Effects on Survey Participation of Interview Length, Incentives, Differential Incentives, and Refusal Conversion," *Journal of Official Statistics*, 15(2), 251-268.
- Groves, R., Fowler, F.J., Couper, M., Lepkowski, J., Singer, E., and Tourangeau, R. (2004). *Survey Methodology*. New York: Wiley.

- Haraldsen, G. (2002). "Identifying and Reducing the Response Burden in Internet Business Surveys." Paper presented at the International Conference on Questionnaire Development, Evaluation and Testing Methods (QDET). Charleston, South Carolina, November 14-17.
- Herzog, A.R. and Bachman, J.G. (1981). "Effects of Questionnaire Length on Response Quality." *Public Opinion Quarterly*, 45, 549-559.
- Hubble, D. L. (1995). "The National Crime Victimization Survey Redesign: New Questionnaire and Procedures Development and Phase-in Methodology." Paper presented at the Joint Statistical Meetings, Orlando, FL.
- Huttenlocher, J., Hedges, L.V., and Prohaska, V. (1988). "Hierarchical organization in ordered domains: Estimating the dates of events." *Psychological Review*, 95, 471 – 484.
- Krosnick, J. (1999). "Survey research." *Annual Review of Psychology*, 50, 537-567.
- Lee, L. and Carr, J. (2009). "Evaluation of a 12-Month Reference Period in the National Crime Victimization Survey (NCVS)," presented at the Federal Committee on Statistical Methodology Conference, Washington, DC. November.
- Menon, G. (1994). "Judgments of Behavioral Frequencies: Memory Search and Retrieval Strategies," in *Autobiographical Memory and the Validity of Retrospective Reports*, Norbert Schwarz and Seymour Sudman, eds., Springer-Verlag, 161-172.
- Menon, G. (1997). "Are the Parts Better than the Whole? The Effects of Decompositional Questions on Judgments of Frequency Behaviors." *Journal of Marketing Research*, 34, 335-346.
- McCarthy, J. S., Beckler, D., and Qualey, S. (2006). "An Analysis of the Relationship Between Survey Burden and Nonresponse: If We Bother Them More, Are They Less Cooperative?" *Journal of Official Statistics*, 22(1), 97-112.
- Miller, P. V. & R. M. Groves. (1985). "Matching survey responses to official records: An exploration of validity in victimization reporting." *Public Opinion Quarterly*, 49, 366-380.

- Navarro, A. and Griffin, R.A. (1993). –Matrix Sampling Designs for the Year 2000 Census.” *American Statistical Association, Proceedings of the Section on Survey Research Methods*, 480-5.
- Neter, J., & Waksberg, J. (1964). –A study of response errors in expenditures data from household interviews.” *Journal of American Statistical Association*, 59, 18 -5 5.
- Peytchev, A. and Tourangeau, R. (2005). –Causes of Context Effects: How Questionnaire Layout Induces Measurement Error.” Presented at the American Association for Public Opinion Research (AAPOR) 60th Annual Conference.
- Prohaska, V., Brown, N. R., & Belli, R. (1998). [Forward telescoping: The question matters.](#) *Memory*, 6, 455-465.
- Raghunathan, T.E. and Grizzle, J.E. (1995). –A Split Questionnaire Survey Design.” *Journal of the American Statistical Association*, 90, 54-63.
- Roberts, C. Eva, G., Allum, N., and Lynn, P. (2010). "Data quality in telephone surveys and the effect of questionnaire length: a cross- national experiment," ISER Working Paper Series 2010-36, Institute for Social and Economic Research.
- Sharp, L.M. and Frankel, J. (1983). –Respondent Burden: A Test of Some Common Assumptions.” *Public Opinion Quarterly*, 47(1), 36-53.
- Shields, J. and N. To (2005). –Learning to Say No: Conditioned Underreporting in an Expenditure Survey.” American Association for Public Opinion Research Annual Conference, Miami Beach, American Statistical Association.
- Tourangeau, R., Rips, L.J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Wedel, M. and Adiguzel, F. (2008), –Split Questionnaire Design for Massive Surveys,” *Journal of Marketing Research*, 25(5), 608-617.
- Wolter, K.M. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.

Zimmet, P. (2004). –Comparison of Fruit and Vegetable Frequency Data from Two Australian National Surveys.” *Nutrition & Dietetics: The Journal of the Dietitians*, 61, 88-97.

Appendix I

CEQ Measurement Issues Study Analysis Plan

FINAL: September 30, 2011

Background Information

This study is part of a comprehensive and ongoing effort to examine alternative data collection strategies for the Consumer Expenditure Quarterly Interview Survey (CEQ). These strategies seek to improve data quality, maintain or increase response rates, and reduce respondent burden and data collection costs. One component of the 2010 CE Strategic Plan is to address the following three survey design questions: (1) does splitting the questionnaire reduce respondent burden and/or improves data quality; (2) do monthly recall periods provide better quality data than quarterly recall periods; and, (3) do global questions provide data of sufficient quality to replace detailed questions? As originally designed, the CEQ-Measurement Issues Study (MIS) would offer insights to the first two questions. However, the incorporation of global questions in the CEQ-MIS also provided the opportunity to examine the third survey design question raised in the 2010 CE Strategic Plan.

The CEQ-MIS has three research objectives. First, this study will assess the effects of administering a shorter CEQ instrument on respondent burden, data quality, and nonresponse error. Second, a separate condition will examine the extent to which using a one-month (versus the current three-month) reference period affects respondent burden, data quality, and nonresponse error. A final objective will be to evaluate the quality of data collected from global, as opposed to, detailed questions on expenditures. The findings from this study will help inform future Consumer Expenditure (CE) research activities as well as redesign decisions for the CE Surveys.

Purpose of Analysis Plan

The purpose of this analysis plan is to outline the primary analytic details for each research objective. It has two components: (1) a narrative describing major steps in the analysis and, (2) a table (in the Appendix) relating the analytic objectives (ranked by order of importance) to variables of interest and successful outcomes. This plan should be viewed as an evolving document that will be updated as new issues arise. It should also be used to guide the analysis of the data collected from the CEQ-Measurement Issues Study (MIS).

General Definitions

The two primary research objectives that the CEQ-MIS is designed to shed light on involve three key concepts. They are respondent burden, data quality, and nonresponse error. Below we define these three concepts in the context of the CEQ-MIS.

1. **Respondent Burden** – Bradburn (1978) identifies four factors that contribute to respondent burden. They are (1) length of the interview; (2) effort required by the respondent; (3) amount of perceived stress experienced by the respondent; and, (4) periodicity of the interview. We will use respondent answers to questions covering

these four factors to determine the effect of each of a shorter questionnaire and a one-month recall period on respondent burden.

2. **Data Quality** – The CE Program Office operates under the premise that “more is better,” suggesting that respondents who report more expenditures, in terms of both number of items and absolute dollar amount, are deemed to have higher data quality than those respondents who report fewer expenditures. In the survey methodological literature, the term “data quality” is often used to refer to multiple error sources (e.g., measurement and sampling) and dimensions (e.g., timeliness and accessibility of data). Therefore, the premise assumed by the CE Program Office only assesses one component of data quality, namely measurement error. More formally, measurement errors refer to deviations from answers given to a survey question and the underlying attribute being measured (Groves, et al. 2004). For this analysis plan, we will adopt the CE Program Office’s current perspective on data quality. Given the design of the CEQ-MIS, it is impossible to know exactly the true value of the attribute being measured; thus, the assessment of data quality for both conditions will be based on indicators (e.g., number of items reported) of measurement errors.
3. **Nonresponse Error** – Nonresponse is the failure to obtain measurements on the sampled units. Sometimes the failure is complete, meaning that no measurements are obtained from the sample unit. We use the term “unit nonresponse” to describe this situation. Sometimes the failure is partial, meaning that it affects one or more survey items, but not all of them. We refer to this situation as “item nonresponse.” For the Split Questionnaire (SQ) condition of the CEQ-MIS, there will be item nonresponse since not all sample members will be asked all survey items. This type of item nonresponse will be referred to as “planned-item nonresponse.” In contrast, “unplanned-item nonresponse” is beyond the survey designer’s control and is a consequence of a sample unit’s unwillingness or inability to provide a response to survey item(s). Since the CEQ is a panel survey and the CEQ-MIS incorporates this feature into its design, panel attrition is relevant to this study. Panel attrition arises when sample members are unit respondents to the first and/or several consecutive interviews, but fail to respond to the remaining interviews. Nonresponse error arises when the values of statistics computed based only on respondent data differ from those based on the entire sample data (Groves, et al. 2004). The evaluation of nonresponse error will be primarily based on indicators of nonresponse error such as response rates, respondent sample composition changes across the waves of the survey.

Analysis Details

As indicated in the CEQ-MIS Statement of Work (SOW), this study has four treatment conditions – a control group (CG), two shortened questionnaire groups (denoted as SQA and SQB and interchangeably referred to as the split questionnaire groups), and a one-month recall period group (RP). The full details of these four treatment groups are documented in the SOW and will not be extensively detailed here, rather only the details necessary for describing the anticipated analyses will be highlighted here. Statistical comparisons among these groups will serve as the basis by which we will address our two primary research objectives: (1) assessing the effect of administering a shorter CEQ instrument on respondent burden, data quality, and nonresponse

error, (2) assessing the effect of using a one-month recall as opposed to a three-month recall period on respondent burden, data quality, and nonresponse error, and (3) assessing the quality of global questions relative to detailed questions. The primary constructs and indicators that will be used to address these questions are briefly described in Table 1 below.

As a first step in the analysis, we describe the sampling procedures used to select the study sample as well as how the sample units were allocated to each of the four test conditions. Special attention should be given to the differences between the procedures typically implemented by the CE to select the sample and the procedures utilized in CEQ-MIS. Another key component of the pre-analysis details is a narrative on the data collection procedures. This should include, but should not be limited to, the telephone calling strategy and soft refusal follow-up procedures. Clearly documenting the sampling and data collection procedures will provide the analysts with valuable context for interpreting and understanding the collected data. It is also important because a possible secondary analysis may be to compare the study data (i.e., the data collected in the CEQ-MIS to CE production data) so it is essential to understand the differences between the two data sources.

Once these procedures are adequately documented, the analysis should systematically address each primary objective as outlined in the SOW and the three requirements of the CEQ-MIS as documented in the 2010 Strategic Plan. The description that follows lays out the steps of this systematic assessment. Additionally, we offer possible hypotheses/explanations of the findings. It should be noted, however, that although we offer several hypotheses about why we would expect to observe certain associations, this is by no means an exhaustive, complete, and/or correct listing of these hypotheses. Furthermore, it is unlikely that we will have the data necessary to directly study these hypotheses. At the very least, documenting them, however, might help guide future research activities for the CE program. As a final cautionary note, we stress that we can only assess whether we have statistical associations among various items and treatment conditions that we are comparing and we cannot draw conclusions about the causal mechanisms underlying these associations.

The First Objective

The first objective is to study the effect of administering a shorter CEQ questionnaire on respondent burden, data quality, and nonresponse error. Based on the given definition of respondent burden, we included in our survey instrument explicit questions pertaining to length of interview, perceived stress experienced by the respondent, effort required by the respondent, and periodicity of the interview. These questions will be asked of every respondent at the end of their scheduled final interview. Thus, we can compare the distribution of the responses on these questions as estimated from the CG and each of the split questionnaire groups. This will help us determine whether splitting the questionnaire, and perhaps equivalently, a shorter questionnaire is associated with a reduction in respondent burden. An indication that respondent burden may be reduced by administering a shorter questionnaire would be a lower perception of burden associated with the SQA and SQB groups relative to the CG. This means the proportion of respondents in the two split questionnaire groups who perceive that their interviews are shorter, feel less stress, require less effort in providing

responses, and think that the periodicity of interviews is low should be higher than the proportion of respondents in the CG group with those same characteristics.

There may be many reasons why a shorter questionnaire, or splitting the questionnaire, may be associated with a lower perception of burden. One possible explanation is that the amount of time it takes to complete the interview task is lower, so it may be viewed as less of an infringement on the time of the respondent. Of course, length of interview is only one component of the definition of respondent burden. Also, the difficulty and stress associated with recalling and correctly reporting purchases made over the previous three months are components of burden. If there are fewer items that the respondent is prompted to recall, then the perception of burden may be lower. In a shorter questionnaire, there would naturally be fewer items being asked about so this is why we may observe a lower perceived burden in the split questionnaire groups.

Once we have studied the effect of administering a shorter questionnaire on respondent burden, the next step in the analysis would be to compare the split questionnaire groups and the CG on components thought to be associated with data quality. As indicated by the definition of data quality, an improvement in data quality among the respondents in the SQA and SQB groups would be evidenced by (among other factors) reporting more items and having higher mean expenditure estimates. This is consistent with the notion of “more is better.” To assess the “more is better” hypothesis, we will compute, initially at the section-level, the average number of items reported for each of the SQA and SQB groups and the CG. We will then compare the corresponding estimate from the split questionnaire groups to the CG. The desirable outcome is that the average number of reports in the split questionnaire group is higher than (and statistically different from) the average number of reports for that same section as calculated from the CG. A similar calculation would be completed for the mean dollar amount for expenditures per section. Again, if a higher mean expenditure is evidenced in the particular split questionnaire group when compared to the CG, then the notion of “more is better” is supported. Therefore, we have evidence suggesting that reducing the length of the CEQ instrument may be associated with an improvement in data quality. These types of comparisons will also be made at the total expenditure level and similar conclusions could be drawn if the split questionnaire groups have higher averages, on both metrics, than the CG.

A shorter questionnaire might lead to more reports among the items actually being asked about in that shorter questionnaire for many reasons. A respondent may not feel rushed to complete the interview because he/she anticipates in advance the length of the interview, thus, the respondent may give more thoughtful responses. By putting more thought into the response process, respondents should be able to recall more of the expenditures and details of those expenditures that they did, in fact purchase over the past quarter year.

Other factors that are thought to be associated with higher data quality should also be investigated as a potential indication of whether a shortened version of the questionnaire results in higher data quality. Some of these factors are related to “good” reporting practices: the use of records (e.g., bank statements, check registers), information book usage, fewer

combined expense reporting, i.e., expenditures for a combination of items reported as an overall expenditure (e.g., a shopping trip during which pants, shirts, and socks were purchased, but the respondent cannot recall the prices for the individual items), and fewer “don’t know” or “refusal” responses. It is plausible that these good reporting behaviors are likely to lead to more accurate reporting. We should also note that these four factors do not comprise of an exhaustive list of factors that are thought to be associated with data quality. Other factors may be investigated at the discretion of the analyst as long as the necessary data are collected and hypotheses justifying why those factors should be investigated are provided.

For the first two characteristics, use of records and information book usage, we will estimate the proportion of respondents within each group that used each recall aid during their completion of the survey¹. An indication that respondents in either split questionnaire group may have higher quality data than respondents in the CG would be a higher proportion of members in each of SQA and SQB using either or both recall aids. Understanding why the respondents in the split questionnaire groups would be more likely to use these recall aids is somewhat challenging to explain or justify. One possible explanation is that a shorter interview may be associated with lower perceived burden (we will actually test this hypothesis earlier in the analysis), in turn making the reporting task easier on the respondents. A lower perceived burden and easier reporting task may make the respondent more motivated to complete the task. If the use of these particular recall aids were made salient to the motivated respondent as a tool useful for completing the task well, then these respondents may be more likely to use the recall aids. Finally, since the use of recall aids are thought to lead to higher quality data, then respondents in the split questionnaire groups should have higher quality data relative to respondents in the CG.

For the third and fourth characteristics, combined expense reporting and the amount of “don’t know” and “refusal” responses, we will compute the proportions of respondents in each of the split questionnaire groups and the CG who have each characteristic. Initially, this will be done at the interview level, meaning that a respondent will be flagged as having the characteristic if at any point in the interview they provide a combined expense or give a “don’t know” or “refusal” response. The rationale for initially doing this at the interview level is that, in general, these characteristics may occur with relatively low incidence. Our study may already be hampered by a low sample size, so in order to obtain reliable estimates of these traits; we will conduct the analysis at the interview level. Of course, the analyst may modify this plan and investigate these characteristics at the section-level.

If a shorter questionnaire is associated with higher quality data, then we would expect a smaller proportion of respondents in SQA and SQB would report combined expenses or give “don’t know” or “refusal” responses relative to the CG. A plausible explanation for this may be that a shorter questionnaire may allow the respondent to focus their attention and recall efforts on a fewer number of broad expenditure categories. A more focused attention would lead to better

¹ We need to verify whether these data are only collected during the final interview as this may impact the interpretation and explanation of the results.

reporting, perhaps in the form of more complete responses (i.e., fewer combined expenses and “don’t know” and “refusal” responses).

Finally, to assess the effects of administering a shorter questionnaire on nonresponse error, we can compare some of the nonresponse properties of the two split questionnaire groups to those of the CG. As indicated in the definition of nonresponse error, we cannot directly assess nonresponse bias as that would require the expenditure reports from the nonrespondents. Since no effort will be made to obtain these measurements from the nonrespondents, our assessment of nonresponse error will consist of analyzing factors that may be associated with nonresponse error. We will compute the response rates, the proportion of eligible sample members completing the interview, for each of the two split questionnaire groups and the CG. If the response rates are low² in any of the groups there exists the potential for nonresponse to adversely affect the results of the study. We would also hope to observe that the response rates in either of the split questionnaire groups would be higher than the response rate in the CG. A possible hypothesis is that respondents may regard length of interview as an important factor when making the decision to participate in a survey. If the survey completion time is too long, then this may make sample members less inclined to complete the interview; thus, offering a shorter questionnaire should guard against the possibility of a lengthy questionnaire negatively impacting the participation decision.

As part of analyzing the nonresponse properties of our split questionnaire condition, we will also compare the demographic characteristics of the respondents in each of the split questionnaire groups to the demographic characteristics of the CG. This amounts to estimating the proportion of each study group that falls into each demographic category. If the distribution of a particular demographic characteristic differs among the groups and those demographic characteristics are also associated with the primary outcome of interest (e.g., expenditures on certain items), then there is the potential for nonresponse bias. It is often the case, that demographics are associated with spending patterns; thus, it is plausible that if the groups have different demographic characteristics then there is the potential for the nonresponse error properties to differ among the groups.

The Second Objective

Similar to the manner in which we analyzed the split questionnaire groups and compared them to the CG, we will use the same methods for comparing the RP to the CG. Since the methods will be almost identical, we will only describe possible hypotheses as to why we may observe various associations. The first component of the second objective is to assess the effect of using a one-month recall period on respondent burden versus a three-month recall period. We hypothesize that by asking a respondent to focus their recall efforts on the past month as opposed to the past three months, respondents in the RP should have a lower perception of burden than respondents in the CG. Furthermore, three months of recall may result in a more difficult reporting task and may make the interview more stressful; thus, a shorter recall period,

² Low is a relative term; thus, it needs to be established as to what we mean by “low.”

e.g., a one-month recall period, may be easier and less stressful and subsequently viewed as less burdensome.

The next step in the analysis is to study the effect of a one-month recall period on data quality. Prior to computing estimates from this treatment condition, we will need to aggregate, or sum, across the second through fourth interviews for each responding unit. The second through fourth interviews correspond to three months of expenditure information, or a quarter's worth of data. Since the CG uses a three-month, or quarterly, recall period, aggregating the data in this manner is the only way we can obtain comparable estimates between the two groups. As a reminder, the six main metrics for data quality that we will be using are: number of expenditure reports, average expenditures, record usage, information book usage, combined expense reporting, and the amount of "don't know" and "refusal" responses.

We will compare the means or proportions of these six characteristics between the two groups. An indication that data quality is higher in the RP than in the CG would be some combination of the following findings: higher average number of expenditure reports and/or average expenditures, a greater proportion of the respondents in the RP consulting records and/or the information book during their interview, a smaller proportion of the respondents in the RP resorting to combined expense reporting and/or giving "don't know" or "refusal" responses.

There are several reasons why a one-month recall period would lead to improved data quality over a three-month recall period. Research on memory suggests that events that happened long ago are harder to recall, so the natural implication for survey design is to shorten the recall period. By shortening the recall period respondents in the RP should be less likely to forget about and/or omit purchases that occurred longer ago. If these respondents forget fewer purchases, then when aggregated over three months, the estimated average total number of reports as well as the estimated average of total expenditures should be higher in the RP than in the CG.

Reducing the length of the recall period may also lessen the likelihood of a retrieval failure. A retrieval failure occurs when there is a failure to bring to mind information that is store in the long-term memory. This would be applicable to entire "purchase events" or individual details about the "purchase events." If the respondent is able to bring to mind the individual details about the purchase event, then there should be less combined expense reporting and "don't know" and "refusal" responses.

A one-month recall period may also make the use of records more likely. It is reasonable to believe that if there is a smaller time lag between interviews, then respondents may be less likely to lose or misplace records; thus, those respondents would have them available to consult during the interview. Furthermore, if interviewers remind respondents that a record system should be maintained and since the interviews are closer in time, these reminders might seem more salient, or significant, to the sample unit. Therefore, the sample unit may be more likely to heed the advice of the interviewer and keep records for use during the interview.

The final task of this research objective is to assess the nonresponse properties of a survey instrument utilizing a one-month recall period. Using the same metrics as in the comparison of the split questionnaire groups to the CG, we will study how the CEQ may be potentially impacted by administering a one-month recall period as opposed to the standard three-month recall period. This will be interesting to study because it is possible that a shorter recall period and the subsequent quicker turnaround time between interviews may actually have an adverse affect on the nonresponse properties of the data collected from the RP.

The first step in analyzing the nonresponse properties would be to compute the wave response rates for the RP and the CG. Response rates are defined as the proportion of eligible units completing the interview. We will then compare the response rates obtained at each pair of interviews between the two groups. In other words, we would compare the interview 1 response rate from the CG to the interview 1 response rate from the RP. We would do this comparison for both interviews 2 and 3. One of the challenges with this analysis is that the RP will be interviewed a total of four times whereas the CG will only be interviewed three times.

It seems likely that the response rate would decrease at a faster rate for the RP rather than the CG. One reason for this may be the shorter time period used in the RP to obtain a completed interview. The data collection period for one interview in the CG is almost twice as long as the data collection period for the RP. Thus, in the CG there is more time to finalize cases, so more completes are likely to be obtained.

The next step in this analysis would be to compare the respondent distributions on various demographic characteristics. Any differences between the two groups would be an indication of nonresponse having a potential impact on the estimates “over what is already present.” Remember that we cannot study nonresponse bias directly (since we are not making an attempt to obtain measurements on nonrespondents) and that we are using the CG as a baseline for comparison. This last part is crucial to our understanding on what our assessment of the nonresponse properties will tell us. By using the CG as the comparison group, we assume that there is some amount of nonresponse that we are “stuck with.” So severe deviations from what we observe in the CG may be an indication that estimates derived from the data obtained from the RP group are severely impacted by nonresponse.

It is plausible that a one-month recall period would affect the representativeness of the respondent population more than a three-month recall period because of the higher frequency of data collection. The discrepancy between groups that are already “harder-to-get” may become larger. So, our data collection efforts might make these groups under-represented in our respondent sample. If these demographic groups also have a different spending pattern than the other demographic groups, then this increases the potential for nonresponse bias affecting the estimates derived only from the RP. As part of this analysis, it may be important to fully understand which demographic groups are historically “harder-to-get,” in terms of a shorter data collection period, and hope that the representation of these groups in the RP is not statistically different from the representation of these groups in the CG.

The Third Objective

The third research objective will be an analysis of the global questions to determine whether they elicit data of sufficient quality to replace detailed questions. The analysis of the global questions will only involve data from the two split questionnaire groups since the RP sample members will not be administered these questions. We should note that not all data quality metrics from Table 1 will be useful in this analysis. In particular, only mean expenditure amounts and the amount of “don’t know” and “refusal” responses will be useful in determining if the global questions are of sufficient quality to replace the detailed questions.

An essential first step in this analysis is to link the global questions to their corresponding detailed questions. We should also document the key differences, if any, between the global questions and detailed questions. We will then estimate the average amount spent on an expenditure category from both sets of questions. We will also estimate the proportion that responses to the global questions were either “don’t know” or “refused” and the proportion that a detailed question resulted in a “don’t know” or “refusal” response. The latter calculation will be done at the section-level.

An indication that the global questions *may* be of sufficient data quality to replace the detailed questions would be if the estimates obtained from each set are statistically identical. One potential problem with this criterion is that it assumes we are satisfied with the quality of the *existing* data. One reason to believe that global questions might be of sufficient data quality to replace detailed questions is that there would be fewer questions asked. With fewer questions being asked, respondents may be less likely to get fatigued during the interview, subsequently they may be less likely to use shortcuts (e.g., satisficing) when responding to questions.

Table 1: Constructs Related to Components of Research Questions³

| Respondent Burden | Data Quality | Nonresponse Error |
|---|---|--|
| <ul style="list-style-type: none"> • Interest • Difficulty • Burdensome • Frequency of participation • Number of phone calls prior to each interview • Perception of length of interview • Estimation of length of interview | <ul style="list-style-type: none"> • Record usage • Information book usage • Mean expenditure amounts • Global expenditure questions • Amount of unplanned item nonresponse (e.g., “don’t know” or “refusal” responses) • Number of expenditure reports • Amount of combined expense reporting | <ul style="list-style-type: none"> • Wave response rates • Cumulative response rates • Sample composition |

Success Determination

We describe what is to be regarded as a successful outcome in favor of an intervention when comparing treatment groups on the indicators and measures related to the analytic objectives in the Appendix. We also document any notes/issues regarding the specific analysis.

Secondary Research Objectives

After the three research objectives have been completed, other research questions and objectives that CEQ-MIS data can address involve:

1. An assessment of the effect of administering a shorter questionnaire on the interviewer (i.e., are subsequent interviews easier for the interviewer to complete because they are shorter?)
 - a. Interviewer expectations could affect the outcome
2. A comparison of collected data, specifically the SQ and R groups, to production data
3. An assessment of whether the collected data, specifically the SQ and R groups, meets the needs of the following primary data users:
 - a. Consumer Price Index (CPI)
 - b. Published tables
 - c. Microdata users
4. An assessment of data collection costs.
5. An assessment of data quality for globals asked before the detailed questions versus after the detailed questions
 - a. Depending on response rates in each of the split questionnaire groups (respondent) sample sizes may make this comparison tenuous.
 - b. Make use of the variable: SQ_SPLIT
6. An assessment of PSAQ_8, PSAQ_9, and PSAQ_10.
7. Imputation model development
 - a. One of the interesting features of the CEQ-MIS is that there will be “planned-item nonresponse” with the both split questionnaire groups.

³ A more detailed version of this table can be found in the Appendix

- b. Although any estimate, specifically mean expenditure estimates, may be computed using only the available cases, i.e., those sample cases that are directly asked the specific expenditure question, it is possible to compute estimates based on the available cases in conjunction with imputed data.
- c. There are four issues relevant to imputation model development:
 - i. Why do we want to impute the missing information?
 - ii. At what level should the imputation be done?
 - iii. How do we build the imputation model(s)?
 - iv. How do we appropriately reflect the imputation model uncertainty?

Because these are secondary research objectives, the analyses to address these questions are not fully documented here. However, a possible analysis addressing (1) above would be group comparisons between the CG and SQ groups on perceived concerns, number of contact attempts, soft refusals, strategies used by the interviewer to solicit participation, etc. Before this analysis is completed, it is necessary to double check whether perceived concerns and strategies used by the interviewer to solicit participation are contained in the CATI CHAI data files (these types of variables are currently captured in the CAPI CHI).

Timeline

We are committed to delivering a final report, summarizing the results of the CEQ-MIS and identifying key findings, to CE Management six months after the conclusion of data collection.

References

- Bradburn, N. M. (1978). Respondent Burden. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 35—40.
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., and Tourangeau, R. (2004). *Survey Methodology*. New Jersey: Wiley.
- Raghunathan, T. E., and Grizzle, J. E. (1995). A Split Questionnaire Survey Design. *Journal of the American Statistical Association*, **90**, 54—63.

Appendix

| CEQ-MIS Success Determination⁴ | | | |
|---|---|--|--|
| Analysis Category | Variable/Construct of Interest | Successful Outcome | Analysis Notes |
| <p>Respondent Burden (RANK: 3)</p> <ul style="list-style-type: none"> • A shorter CEQ, in terms of number of sections asked to the respondent, should be less burdensome to complete because it will take less time. • It should be easier to recall expenses incurred in the last month than over the past three months; however, this gain in recall ease likely will be outweighed by the additional burden of being contacted/interviewed every month in the 1-month recall period condition. Thus, the burden associated with the R group should be more than the CG. | <ul style="list-style-type: none"> • Interest • Difficulty • Burdensome • Frequency of participation • Number of phone calls prior to each interview • Perception of length of interview • Estimation of length of interview | <ul style="list-style-type: none"> • Interest level in the survey should be the <u>same or higher</u> in the SQ group, and <u>lower</u> in the R group, when compared to the C group. • Perceived level of difficulty completing the survey should be <u>lower</u> in the SQ group and <u>higher</u> in the R group when compared to the C group. • Perceived level of burden should be <u>lower</u> in the SQ group and higher in the R group when compared to the C group. • Perception of frequency of participation should be <u>the same or lower</u> in the SQ group and higher in the R | <ul style="list-style-type: none"> • Determine how close, on average, the respondent's estimate of the length of the interview was to the actual length. • What factors will potentially modify the dimensions of respondent burden? <p>Rankings of burden dimensions for SQ</p> <ol style="list-style-type: none"> 1. Burden 2. Perception of length 3. Difficulty 4. Frequency of participation 5. Estimate of length 6. Number of phone calls 7. Interest <p>Rankings of burden dimensions for R</p> |

⁴ This Table only details the Success Determination for the original objectives of the CEQ-MIS. Recall that the original objectives of the CEQ-MIS were to assess the effects of administering a shorter CEQ instrument on data quality, nonresponse error, and respondent burden and to assess the effects of using a one-month recall period as opposed to a three-month recall period on those same constructs.

| | | | |
|---|---|---|--|
| | | <p>group when compared to the C group.</p> <ul style="list-style-type: none"> • Perception of the number of phone calls prior to each interview should be <u>the same or lower</u> in the SQ group, higher in the R group when compared to the C group. • Perception of length of the interview should be <u>shorter</u> in the SQ and R groups when compared to the C group. • Average estimates of length of interview based on PSAQ_5 should be <u>lower</u> from the SQ groups when compared to the C group. | <ol style="list-style-type: none"> 1. Difficulty 2. Frequency of participation 3. Burden 4. Number of phone calls 5. Perception of length 6. Estimate of length 7. Interest |
| <p>Data Quality (RANK: 1)</p> <ul style="list-style-type: none"> • A shorter CEQ, in terms of number of sections asked to the respondent, should yield higher quality data because respondents may not get fatigued as much as in a longer CEQ. | <ul style="list-style-type: none"> • Record usage • Information book usage • Number of expenditure reports per section • Mean expenditure amounts | <ul style="list-style-type: none"> • Record usage rates in the SQ and R groups should be <u>the same or higher</u> when compared to the C group. • Information book usage rates in the SQ and R groups should be <u>the same</u> | <ul style="list-style-type: none"> • Appropriate adjustments should be made in order to accurately compare the R group estimates of number of expenditure reports, amounts, and global expenditure amounts to the C group. |

| | | | |
|---|---|--|---|
| <ul style="list-style-type: none"> • A one-month recall period should result in higher quality data than a three-month recall period because purchases made more recently should be easier to recall and report. | <p>per section</p> <ul style="list-style-type: none"> • Global expenditure amounts • Amount of unplanned item nonresponse • Amount of combined expense reporting | <p><u>or higher</u> when compared to the C group.</p> <ul style="list-style-type: none"> • The average number of expenditure reports per section should be <u>the same or higher</u> in the SQ and R groups when compared to the C group. • The average amount of expenses incurred per section should be <u>the same or higher</u> in the SQ and R groups when compared to the C group. • Average global expenditure estimates from the SQ groups should be <u>the same or higher</u> than estimates computed from the corresponding detailed sections in the C group. • Rates of unplanned item nonresponse should be <u>lower</u> in the SQ and R groups than in the C group. • Rates of combined expense reporting should be <u>lower</u> | <ul style="list-style-type: none"> • How well do the global expenditure questions and amounts coincide with the corresponding detailed item questions? • In order to compare global questions to detailed item questions, we must link the specific global question to the set of appropriate detailed questions. • In determining how to calculate an estimate of “unplanned missing data” we should attempt to mimic what is done in production as closely as possible. It is also possible use appropriate codes (e.g., XX8, XX9) to identify “Don’t know” and “Refusal” options for any of the expenditure items. • To compute the number of expenditure reports, one could sum an indicator for non-zero dollar amounts for any of the expenditure variables of interest. <p>Rankings of DQ dimensions</p> |
|---|---|--|---|

| | | | |
|---|--|---|--|
| | | in the SQ and R groups than in the C group. | <ol style="list-style-type: none"> 1. Number of expenditure reports 2. Mean expenditure amounts 3. Global expenditure amounts 4. Combined expense reporting 5. Unplanned missing data 6. Information book usage 7. Record usage |
| <p>Nonresponse Error (RANK: 2)</p> <ul style="list-style-type: none"> • A shorter CEQ, in terms of number of sections asked to the respondent, should result in higher unit response rates than a longer questionnaire because a shorter questionnaire would be less of an infringement on the respondent's time than a longer one and thus influence the participation decision in a positive manner. • A one-month recall period may, in fact, have a deleterious effect on nonresponse error due to | <ul style="list-style-type: none"> • Wave response rates • Cumulative response rates • Sample composition | <ul style="list-style-type: none"> • Wave and cumulative response rates from the SQ group should be <u>the same or higher</u> when compared to the C group. • Wave and cumulative response rates from the R group should <u>be maintained</u> when compared to the C group. • There should <u>be no</u> sample composition differences between the four treatment groups and across all waves. | <ul style="list-style-type: none"> • Compute response rates by various subgroups <p>Rankings of NE dimensions</p> <ol style="list-style-type: none"> 1. Wave response rates 2. Cumulative response rates 3. Sample composition |

| | | | |
|---|--|--|--|
| the short turnaround time in contacting and interviewing respondents. | | | |
|---|--|--|--|

Table 1: Expanded list of Variables to be used in Analyses

| Analysis Category | Construct | CATI Variables |
|--------------------------|---|---|
| Respondent Burden | Interest | PSAQ_1 |
| | Difficulty | PSAQ_2A |
| | Burden | PSAQ_3 |
| | Perception of the frequency of participation requests | PSAQ_3A |
| | Perception of the number of phone calls prior to each interview | PSAQ_3B |
| | Perception of the length of the interview | PSAQ_4 |
| | Estimation of the length of the interview | PSAQ_5 |
| | Timing variables | SHLTR SECT_06 SECT_09 SECT_12 SECT_13 SECT_14 SECT_16 SECT_17 SECT_18 SECT_20 GLBL INCOME PSAQ BACK FRONT CONTROL TOT_TIME |
| Data Quality | Record usage | PSAQ_7 PSAQ_7B |
| | Information book usage | PSAQ_11 PSAQ_11A PSAQ_11B PSAQ_11C PSAQ_11D |
| | Number of expenditure reports | Section 6 – Major HH Appliances <ul style="list-style-type: none"> • MAJINSTX • MAJPURX • MAJRENTX • INSTLLEX • MINPURX • MINRENTX Section 9 – Clothing |

| | | |
|--|--|--|
| | | <ul style="list-style-type: none"> • CLOTHXA Section 12 – Vehicle Operating • VOEXPX • VOPRMBXA Section 13 – Non-health Insurance • INSEX3AX • INSEXXA • INSEXPBX • INSNEXXB Section 14 – Health Insurance • HHIPDAMT • HHICMXXA • HHIRPMXB • HHIIRGXB • HHICMXXB Section 16 – Education • EDEXOXA • EDREIMBX Section 17 – Subscriptions • S17PURXA • S17CMEXX • SPRTFEEX • SPORTADX • RECADMX • ENTADMX • OTHBKRFX • NEWSMAGX • RECORDX • FILMX • FLMPRCSX • VIDOPURX • VIDORNTX • SPFEECMX • SPRTADXC • RECADMXC • ENTADXCM • OTHBKRCM • NEWMAGCM • RECORDXM • FILMXCM • FLMPRXCM • VDPURXCM • VDRNTXCM Section 18 – Trips • TOTYUPDX • PKGTRIPX |
|--|--|--|

| | | |
|--|-----------------------------------|--|
| | | <ul style="list-style-type: none"> • CMLOCALX • CMPLANEX • CMTRAINX • CMBUSX • CMSHIPX • RTCARX • RTTRUCKX • RTMOPEDX • RTPLANEX • RTBOATX • RTCAMPX • RTOTHERX • GASOILX • TRPTOLLX • PARKINGX • LDGCOSTX • TRPFOODX • TRPALCHX • TRPGROCX • TRPALCGX • TRPSPRTX • TRSPORTX • TRPETRTX • TRMISCX • LOCDEALX • LOCLODGX • LOCMEALX • ALCMEALX • LOGGROCX • ALCGROCX • LOCADMSX <p>Section 20 – Expense Patterns</p> <ul style="list-style-type: none"> • GROCWEKX • OTHSTUFX • OSTORWKX • DINE_WKX • ALC_HOMX • ALC_OUTX |
| | <p>Mean expenditure amounts</p> | <p>NOTE: Use variables listed under “Number of expenditure reports”</p> |
| | <p>Global expenditure amounts</p> | <p>G9a G12a G13a G13b G17a G17b</p> |

| | | |
|--------------------------|--------------------------------------|--|
| | | G17c G17d G6a G6b G6c G14a G16a G18a G20 |
| | Amount of unplanned nonresponse | NOTE: Use variables listed under “Number of expenditure reports” |
| | Amount of combined expense reporting | MJAPPLn (1-9) MNAPPLn (1-9) CLOTHAn (1-9) VPSRVYn (1-9) EDUC_An (1-9) POLICYYB |
| Nonresponse Error | Wave response rates | OUTCOME CONVREF SINGL_INT INTERI |
| | Cumulative response rates | OUTCOME CONVREF SINGL_INT INTERI |
| | Sample composition | QTYPE SQ_SPLIT PANEL PSU SEGMENT SEX AGE HORIGIN EDUCA RC_WHITE RC_BLACK RC_ASIAN RC_NATAM RC_PACIL RC_OTHER H1 |

Appendix II
MIS Sampling Plan

Draft – May 17, 2010

MEMORANDUM FOR Scott Fricker
 Project Lead, Measurement Issues Study
 Office of Survey Methods Research
 Bureau of Labor Statistics

From: Stephen Ash
 Chief, Victimization and Expenditures Branch
 Demographic Statistical Methods Division
 U.S. Census Bureau

Subject: Sample for 2010 Consumer Expenditures Quarterly Measurement
 Issues Study (CEQ-MIS)

I. Introduction

This memorandum describes the steps used to produce the sample for the CEQ-MIS.

The Bureau of Labor Statistics (BLS) is conducting the CEQ-MIS, which is the first in a comprehensive ongoing effort to examine alternative data collection strategies for the CEQ Survey that would improve data quality, maintain or increase response rates, and reduce data collection costs. This study focuses on the effects of a shorter questionnaire on respondent burden, data quality, and nonresponse error. In addition, a separate condition in this study assesses the effects of a one-month versus three-month reference period on underreporting due to recall errors. The results of this study are being used to inform future CEQ research activities and decisions about how to redesign the production survey.

The Census Bureau is collecting data for this study by computer-assisted telephone interviewing (CATI). Interviews begin June 1, 2010 and end February 28, 2011. There are three waves of data collection for the conditions, control group (CG) and split questionnaire (SQ), and four waves for the recall period (RP) condition. Within each condition, each sample unit is contacted once per wave. The Demographic Statistical Methods Division (DSMD) selected the sample of addresses.

II. Design for the CEQ-MIS Sample

The goal of the 2010 CEQ-MIS sample design was to have 8,100 completed interviews over all three conditions (or treatment groups). The three conditions were equivalent to three separate samples. To account for cases that would not yield completed interviews for various reasons, we determined that we needed at least 17,374 reserve CEQ, unit-

frame cases from the DSMD sample database. Some examples of incomplete interviews that we accounted for were unavailable telephone number for the chosen address; ineligible case due to a nonworking or nonresidential telephone number; interview refusal; or respondent could not be reached.

DSMD used the CEQ reserve cases from the unit frame for this study. Reserve cases are separate from the regular CEQ sample. However, both contain the same types of cases and have the same national distribution. At the beginning of the 2000 sample design, the Census Bureau set aside the reserve cases to be used for research and/or sample expansions as needed during the decade. (Census 2004.) The unit frame makes up about 80% of the CEQ reserve sample (as well as the regular CEQ sample). Unit-frame cases refer to CEQ cases obtained from the 2000 census that typically have complete addresses that include a house number and street name.

For each of the conditions, the sample release is staggered. One-third of the cases assigned to the SQ condition have their first interview in June, one-third have their first in July, and one-third have their first interview in August. These thirds are referred to here as panels. This staggered schedule will be carried forward throughout all subsequent interviews, based on when the case was originally released and the appropriate reference period for the condition.

For all three conditions, only completed and sufficient partial cases, final codes of 01 and 02, are carried forward to wave two for their second interview. Between waves two and three and waves three and four, only out-of-scope cases and congressional-refusal cases are dropped (final codes of 20, 21 and subtype 003, and 176.) (A congressional-refusal case is removed upon request of a U. S. Congressman that we stop interviewing one of their constituents. See Attachment A for definitions of final codes.)

The telephone center will attempt to conduct three interviews for each case in the CG and SQ conditions, and will attempt four interviews for the RP condition, except for the incomplete interviews in wave 1 as mentioned above. Also for the second and third waves for the SQ condition, cases were split into two subsamples to receive a shorter version of the CEQ-MIS questionnaire. Table 1 illustrates the timing, conditions, and number of completed interviews expected by condition and wave.

Table 1. Expected CEQ-MIS Interviews by Condition & Wave

| Condition | Completed Interviews by Condition & Wave | | | | |
|---|--|--------------|--------------|--------------|--------------|
| | Wave 1 (Jun) | | | Wave 2 (Sep) | Wave 3 (Dec) |
| Control Group (CG) | 750 | | | 525 | 450 |
| Split Questionnaire (SQ) | 1,515 | | Total | (n = 1,060) | (n = 900) |
| | | | Subsample A | 530 | 450 |
| | | | Subsample B | 530 | 450 |
| | Wave 1 (Jun) | Wave 2 (Jul) | Wave 3 (Aug) | Wave 4 (Sep) | |
| Recall Period (RP) | 1,036 | 725 | 617 | 525 | |
| Total Completed Interviews / Sample Units | 3,301 | | | | |

We selected 17,374 reserve CEQ addresses with an inflation rate of 5.26 given the following calculation:

$$17,374 \text{ addresses} = 3,301 \text{ completed interviews} \times \frac{1 \text{ address}}{50\%} \times \frac{1 \text{ address}}{40\%} \times \frac{1 \text{ address}}{95\%}$$

The following assumptions define the denominators in the preceding calculation.

Assumptions

A. Address to Telephone-Number Match (Telematch) Rate

We expected about 50% of the addresses we initially selected to match to a telephone number in the telematch procedure (address to telephone-number matching operation.) The non-matches were excluded from the CEQ-MIS sample, and will not be available for other samples that may be drawn from the reserve sample in the future.

From the start of the telematch procedure to the start of interviewing is three months, March 1 to June 1. This time lag should not be long enough to significantly affect the quality of the telephone numbers obtained in the matching operation. Also, the time lag between the telematch operation and interviewing

will be consistent for each condition, since wave 1 interviewing starts in the same month within each condition.

However, the time lag for the different panels within a condition may result in differing ineligibility rates across panels. (Sample for each condition was divided into panels, which followed the same interviewing schedule; panel one for all three conditions starts in June, panel two starts in July, and panel three starts in August.)

After the telematch operation, depending on the number of complete telephone numbers obtained, DSD and DSMD may determine that more should be done to maintain eligibility rates in the later panels. Two options available for the later panels are sending telematch cases without numbers to the telephone center for further research for the later panels or repeating the telematch operation later to keep the time lag similar for each panel. However, our decision depends on whether repeating the telematch operation would be feasible to complete before the start of production and on how smoothly the March telematch operation would run.

B. Response Rates

For planning the sample size, we estimated the response rate for the first wave of interviews would be 40%. This was a conservative estimate compared to Telephone Point of Purchase Survey's (TPOPS') first interviews of about 52%. This rate was based on the TPOPS sample that received an advance letter, as planned for the measurement study as well.

The 40% response-rate estimate was also based on the AAPOR (American Association for Public Opinion Research) Response Rate #3, which estimates what proportion of cases of unknown eligibility is actually eligible. (*Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, by the American Association for Public Opinion Research.)

C. Ineligibility Rate

Using recent historical ineligibility rates for the ongoing CEQ survey, we expected 3 to 5% of sample to become ineligible after telephone numbers were obtained. This rate is sometimes called the type C rate also. Since the addresses come from the same source as regular CEQ, we assumed they should have the same type C rate. We will use the more conservative 5% estimate to account for these cases. The calculation above uses the percentage *eligible* – 95%.

D. Bounding Interviews

For the CEQ Survey, the first interviews are called bounding interviews, and data from these cases are excluded from survey estimates. In this study, data from first interviews *is* used, so we did not inflate the sample for these cases.

III. Design for the Supplemental CEQ-MIS Sample Universe

The telematch rate did not come in as high as our assumption of 50%, and as a result, we decided to obtain more cases for the study by creating a supplemental sample universe and sending it through a second telematch operation. For the rest of this document, we refer to the set of cases we initially selected for telematch as the original universe and the second set of cases as the supplemental universe.

For the supplemental universe, we selected 14,226 reserve CEQ addresses with an inflation rate of 10.08 given the following calculation and revised assumptions:

$$14,226 \text{ addresses} = 924 \text{ completed interviews} \times \frac{1 \text{ address}}{31\%} \times \frac{1 \text{ address}}{40\%} \times \frac{1 \text{ address}}{80\%}$$

Assumptions Applied to the Supplemental Universe

A. Telematch Rate

We selected the supplemental universe based on the rate achieved by the initial operation -- 31% -- thereby inflating the supplemental universe significantly more than the original.

B. Response Rates

We used the same response rate for the supplemental as for the original universe -- 40%.

C. Ineligibility Rate

Instead of using Census's historical CEQ ineligibility rate, we used a more conservative ineligibility rate of 20%, similar to that used in the 2009 study, "An Evaluation of Bias in the 2007 National Households Education Surveys Program: Results From a Special Data Collection Effort." (Dept. of Education, 2009.)

D. Bounding Interviews

Bounding-interview data is used from the supplemental cases as well as the original cases.

IV. Description of the CEQ-MIS Sample Production

Table 2 shows how many CEQ reserve sample units were available and needed for the CEQ-MIS at different stages of the sample production.

Table 2. Available CEQ Reserve Units and CEQ-MIS Sample Counts

| Conditions: | CG | SQ | | RP | | Total |
|---|--------|-------|-------|--------|-------|--------|
| Sample Designations: | X19 | X11 | X17 | X18 | X20 | |
| QTYPE: | 1 | 2 | 3 | 4 | 4 | |
| Original Universe Based on Original Assumptions | | | | | | |
| 1. CEQ Reserve Units | 9,137 | | 9,317 | 9,209 | | 27,663 |
| 2. Sample to Send to Telematch | 4,614 | | 9,317 | 6,139 | | 20,070 |
| 3. Minimum Sample Required for Telematch | 3,947 | | 7,974 | 5,453 | | 17,374 |
| 4. Sample to Send for Interviewing | 1,974 | | 3,987 | 2,726 | | 8,687 |
| 5. Completed Interviews Desired – (Wave 1) | 750 | | 1,515 | 1,036 | | 3,301 |
| Supplemental Universe Based on Revised Assumptions | | | | | | |
| 6. CEQ Reserve Units | 4,529 | 9,135 | | 3,053 | 9,135 | 25,852 |
| 7. Sample to Send to Telematch | 2,221 | 8,857 | | 1,574 | 1,574 | 14,226 |
| 8. Minimum Sample Required for Telematch | 2,117 | 4,276 | | 1,462 | 1,462 | 9,317 |
| 9. Sample to Send for Interviewing | 656 | 1,326 | | 453 | 453 | 2,888 |
| 10. Completed Interviews Desired – (Wave 1) | 210 | 424 | | 145 | 145 | 924 |
| Original & Supplemental Counts | | | | | | |
| 11. CEQ Reserve Units | 13,666 | 9,135 | 9,317 | 12,262 | 9,135 | 53,515 |
| 12. Sample <i>Sent</i> to Telematch | 6,829 | 8,857 | 9,317 | 7,656 | 1,574 | 34,233 |
| 13. Sample <i>Received</i> from Telematch | 2,411 | 3,785 | 2,903 | 2,537 | 698 | 12,334 |
| 14. Sample <i>to Send</i> for Interviewing | 2,344 | 2,367 | 2,367 | 2,537 | 698 | 10,313 |

As illustrated by the table, each condition was selected from different CEQ reserve sample designations. The CEQ reserve cases are a sample of cases separate from the regular sample, but both were selected from the same universe and contain the same types of cases with the same national distribution. The reserve sample is about seven-tenths the size of the regular sample. Both samples are also structured and coded the same way with cases evenly grouped by sample designations. There are 10 sample designations in the reserve sample and 14 sample designations in the regular sample: X11-X20 and Q33-Q46, respectively. Each CEQ sample designation is further divided into 101 reduction groups. Each of these groups is representative of the entire sample designation. The sample was divided this way so we can reduce sample quickly by dropping reduction groups without introducing bias. For example, cases in each reduction group are spread out geographically across the sample designation so that eliminating one reduction group does not eliminate any geographical area, but reduces across all areas evenly.

Starting in September 2009, we expanded the ongoing CEQ sample to maintain the appropriate workload size for the 2000 sample design. We used two reduction groups, 001 and 084, from each of the CEQ reserve sample designations, except X17. We did not use X17 in the CEQ expansion because the other nine sample designations provided just the amount of sample that was needed through the end of the 2000 design. (Census 2009a.) Since reduction groups 001 and 084 were used for the expansion, we have excluded those groups from the CEQ-MIS. However, since all 101 reduction groups were available in the X17 sample designation, we used them all for the largest condition, the split questionnaire. We chose two other sample designations, each having 99 reduction groups, to use for the smaller CG and RP conditions. We decided to use X18 and X19 since they were the first introduced in the CEQ expansion and had already cycled through the sampling system. Line 1 of Table 2 demonstrates these numbers by sample designation and condition. In addition, several CE areas were reduced in 2006, so cases from those areas are also excluded from the sample units available for selection.

A. Defining the CEQ-MIS Original Universe

Line 2 of Table 2 gives the number of cases from each sample designation that we selected for telematch. Below we describe the method we used to derive the number of cases for each condition:

- Condition CG; Sample Designation X19:

We selected all odd reduction groups, excluding 001, which had been used for the 2009 expansion. The groups selected were 003, 005, 007, ..., 097, 099, 101. These groups yielded 4,608 cases for telematch.

- Condition SQ; Sample Designation X17:

We selected all reduction groups: 001, 002, 003, ..., 099, 100, 101. These groups yielded 9,317 cases for telematch.

- Condition RP; Sample Designation X18:

We selected two-thirds of the reduction groups, excluding 001 and 084, which had been used for the 2009 expansion: The groups selected were 002, 003, 005, 006, 008, 009, 011, 012, 014, ..., 095, 096, 098, 099, 101. Following this pattern, we excluded 34 reduction groups (001, 004, 007, ..., 094, 097, 100). However, reduction group 084 did not fall in this set, so we also excluded it for a total of 35 reduction groups excluded and 66 reduction groups included. The included groups yielded 6,082 cases for telematch.

B. Defining the CEQ-MIS Supplemental Universe

Line 7 of Table 2 gives the number of cases from each sample designation that we selected for telematch. Below we describe the method we used to derive the number of cases for each condition:

- Condition CG; Sample Designation X19:

We selected every fourth reduction group, excluding 084. The groups selected were 004, 008, 012, ..., 092, 096, 100. These groups yielded 2,221 cases for telematch.

- Condition SQ; Sample Designation X11

We selected all reduction groups, excluding 001, 031, 061, 091, and 084. These groups yielded 8,857 cases for telematch.

- Condition RP; Sample Designation X18 and X20

We selected half of the remaining reduction groups in sample designation X18 by selecting every sixth group, starting with 004. The groups selected were: 004, 010, 016, 022, 028, 034, 040, 046, 052, 058, 064, 070, 076, 082, 088, 094, 100. These groups yielded 1,574 cases for telematch. (We excluded reduction groups 001 and 084.)

We selected 17 reduction groups (the same number of groups selected above) from sample designation X20 to complete the supplemental sample needed for the RP condition. We selected every sixth group starting with 004 as in the previous step. These groups yielded 1,574 cases for telematch. (We excluded reduction groups 001 and 084.)

Note: There were 3,053 cases available for selection in sample designation X18, which was enough to provide the required number of completed interviews. However, we only took one-half of these cases and obtained the other half from sample designation X20 to allow for the possibility of future sample expansions. If future expansions were necessary, we would need to use the same design as before, which calls for case selection from across all the reserve sample designations, including X18.

C. Creating the Original Sample Universe

Within DSMD, the Victimization and Expenditures Branch (VEB) specified for the CPS, Health, and Income Surveys Programming Branch (CHISPB) how to select the address-based cases for the sample universe. Per specifications from the Demographic Surveys Division (DSD), CHISPB created a universe file of these sample cases, consisting of case control numbers.

This stage of sampling corresponds to lines 2 and 3 of Table 2. For the CG condition, to obtain the estimated 3,947 cases needed at this point, CHISPB selected all odd reduction groups as described in Section, IV.A. This selection yielded more than the desired number of cases, approximately 4,614, but we sent them all to be matched to telephone numbers. With this approach, we were more certain of getting enough usable cases back from the telematch operation. The next step would be to systematically subsample the units with a telephone number to obtain the desired sample size.

For the SQ and RP conditions, we followed the above procedure except we selected all reduction groups from sample designation X17 for the SQ condition, and two-thirds of the reduction groups from sample designation X18 for the RP condition.

D. Creating the Supplemental Sample Universe

DSMD repeated the same process for the supplemental universe as for the original universe described in the previous section, except we used numbers of cases desired based on the revised rates of inflation. Table 2, lines 7 and 8 demonstrate the number of cases desired for telematch by condition.

E. Selecting the Final Sample

The cases with phone numbers from both the original and supplemental universes were combined before creation of the final sample. Cases on the final sample file are not identified as original or supplemental, although we have the ability to identify which universe each unit came from. The origin of the cases is not

relevant while the cases are in the field, and if that information were passed on with the sample, it could bias efforts toward obtaining interviews.

Even though we revised assumptions from the original to the supplemental universe, after combining the telematched cases, we confirmed that each sample designation had enough sample to say the number was inflated by the more-conservative, supplemental rates. With our second telematch-rate assumption, only 31% of the 14,226 cases sent to telematch would have telephone numbers. But since the second match had a 43% match rate, we had enough cases for inflation at the higher, 20% ineligibility rate. Row 14 of Table 2 shows the counts we determined would meet the inflated rate of 20% for ineligible cases and 40% for nonresponse, by condition and sample designation. We created the final sample using those counts; the total sample size was 10,314.

The following steps describe how we created the final sample. For the CG condition, the number of cases to send out for wave one was 2,344 cases. See Table 2, line 14. To meet this goal (thus avoiding unnecessary interviewing costs), VEB subsampled the telephone-number cases by removing cases in excess of 2,344 using the steps below.

1. Determined the number of cases to remove from sample, r = Number of units with telephone numbers (after the telematch procedure) – desired number of cases for CG wave 1, 2,344.
2. Determined the take every, te = Number of telephone cases from the matching operation, 2,411 / number of cases to remove from sample, r .
3. Ordered the telephone cases by QTYPE, Sample Designation, Stratification Primary Sampling Unit (SPSU), Basic PSU Component (BPC), Reduction Group, and Final Hit Number (FHN).
4. Starting at the beginning of the ordered set of telephone-number cases, used the take every to identify the cases to remove from the CG sample.
5. Placed the remaining cases on the final sample file for CHISPB.

Next VEB subsampled the SQ and RP conditions by the same process described above. VEB then systematically assigned condition (using QTYPE 1 to 4) and panel to each sample unit on the final sample file for CHISPB. VEB made the assignments separately for each condition, using the same sort described in the subsampling steps. The 2,344 cases from sample designation X19 were assigned to condition CG, QTYPE 1; the 2,367 cases from X11 and X17 were assigned to condition SQ, QTYPE 2 and 3, respectively; and the 2,537 cases from sample designation X18 and the 698 cases from X20 to condition RP, QTYPE 4.

VEB then provided the final sample file to CHISPB who loaded it onto the 2000 sample database. CHISPB created a wave-1 sample control input file (SCIF) containing the CG and SQ samples and a separate wave-1 SCIF with the RP sample, according to the instructions in the “Specification for the 2010 Consumer Expenditure Quarterly Survey – Measurement Issues Study Instrument SCIF” (Census, 2009b). DSMD then verified and released the SCIFs to the Technologies Management Office (TMO).

Note: For each sample designation, only the reduction groups excluded from this study sample will be available for potential selection of CEQ reserve sample in future studies or surveys.

V. Timing:

1. DSMD-CHISPB delivers the universe file to DSD by 3/1/10.
2. DSMD-VEB/CHISPB receives a list of control numbers for cases that received a telephone number from the telematch operation, by 3/19/10.
3. DSMD-CHISPB delivers the supplemental universe file to DSD by 4/5/10.
4. DSMD-VEB/CHISPB receives a list of control numbers from the second telematch operation, by 4/13/10.
5. DSMD-VEB delivers the production sample to DSMD-CHISPB and DSD by 4/21/10.
6. DSMD-CHISPB delivers the first (two) production SCIFs to TMO by 5/18/10.

VI. Contacts

DSMD-VEB: Leslie Flores, 301-763-5947, Leslie.R.Flores@census.gov
Stephen Ash, 301-763-4294, Stephen.Eliot.Ash@census.gov

VII. References

U.S. Census Bureau (2004). “CEQ FACT SHEET (2000 Design),” dated November 10, 2004. <http://cww.census.gov/demo/dsmd/Main/Surveys/major/ces.html>

U.S. Census Bureau (2009a). “Sample Date Reassignment for 2009-2015 CEQ Expansion (Doc. #5.3-S-12),” from Patrick E. Flanagan to James M. Lewis, dated December 3, 2009.

U.S. Census Bureau (2009b). “Specification for the 2010 Consumer Expenditure Quarterly Survey – Measurement Issues Study Instrument SCIF,” from Jimmie B. Scott to James M. Lewis, dated December 16, 2009.

Bureau of Labor Statistics (2009). "Measurement Issues Study Statement of Work," dated July 10, 2009.

U.S. Department of Education (2009). "An Evaluation of Bias in the 2007 National Households Education Surveys Program: Results from a Special Data Collection Effort," dated March 2009.

Attachment

cc:

| | |
|-------------|------|
| R. Schwartz | DSD |
| S. Ash | DSMD |
| L. Flores | " |
| J. Adams | " |

Final Codes for The CEQ Measurement Study Instrument

| Final | Subtype | Description |
|--------------|----------------|---|
| 1 | | Fully complete |
| 2 | | Sufficient Partial |
| 15 | | Complete/sufficient partial, Special Resolutions |
| 20 | | Sample Unit ineligible - out of scope |
| | 001 | HH/Respondent Deceased |
| | 004 | Mobile Phone |
| | 005 | No one uses as a usual residence |
| | 011 | HH under age |
| | 013 | Pay phone |
| 21 | | Sample unit eligible but unavailable through Closeout |
| | 001 | HH Institutionalized |
| | 002 | HH Unavailable through closeout |
| | 003 | HH Moved |
| 23 | | Coded out based on survey parameters |
| 24 | | Unconverted language problem |
| 25 | | Unconverted hearing barrier |
| 172 | | Sample Reduction |
| 176 | | Congressional refusal - delete case |
| 179 | | Hostile break off |
| 181 | | Refusal |

| | | |
|-----|--|--|
| 182 | | Hard Refusal |
| 183 | | Exceeded unproductive call max |
| 185 | | Sufficient Partial w/planned callbacks |
| 186 | | Pre-refusal -- explicit refusal or hostile break off |
| 188 | | Insufficient partial - callback |
| 191 | | Language barrier |
| 192 | | Hearing barrier |
| 193 | | Privacy Detector |
| 194 | | Never contacted -confirmed number |
| 195 | | Never contacted - unconfirmed number |
| 198 | | Other assessor pre-final- type 3 |
| 199 | | Never tried (new case) |

Appendix III

MIS Advance Letters

CEQ-303-L1 (MS1)
(6-2010)



UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001
OFFICE OF THE DIRECTOR

FROM THE DIRECTOR
U.S. CENSUS BUREAU

We are contacting you to request your help with an important survey—the Consumer Expenditure Telephone Survey. The U.S. Census Bureau is conducting this survey for the Bureau of Labor Statistics.

Your responses will help to improve the quality of data used to update the Consumer Price Index (CPI). The CPI is the most important tool used to measure how fast prices are rising or declining. It directly affects wages, pensions, and the cost of goods and services. Your response is a service to your community and the country.

Your participation in this survey is essential; however, you may choose to decline to answer any particular question. Federal law authorizes the collection of this information (Titles 13 and 29 of the United States Code), and Sections 9 and 214 of Title 13 require us to keep all information about you and your household strictly confidential, and to use that information for statistical purposes only.

Soon a Census Bureau interviewer will call you for an interview. We would appreciate your cooperation.

On the back of this letter are answers to questions you may have about this survey. We have also enclosed a factsheet about this survey. If you would like more information, please call us toll-free at 1-888-595-1257

Thank you for your cooperation in this important survey. We appreciate your help.

Sincerely,

Robert M. Groves

Enclosure

USCENSUSBUREAU

Helping You Make Informed Decisions

www.census.gov

FREQUENTLY ASKED QUESTIONS

What is the Census Bureau doing? I thought they only operated every ten years.

In addition to the decennial census, we collect data on a monthly basis. We collect these data to provide up-to-date information on topics such as unemployment, spending, family income, housing, manufacturing, and business activities to track the country's economy.

What is this survey about?

The survey you are participating in aims to improve the quality of data collected for the Consumer Expenditure Survey, which measures how people in the United States spend their money on items such as housing, food, education, transportation, and healthcare.

Why is this survey important?

The information you provide will affect the quality of national statistics on the cost of goods and services. In addition, people in your community and throughout the country use the Consumer Expenditure Survey statistics for planning public services and addressing consumer needs. Your voluntary response is essential for ensuring that this survey's results are complete and accurate.

What kinds of questions will I be asked?

We will ask you to recall information about recent purchases you have made. We also will ask you questions about the people who live in your household, such as their ages, and questions about things you already own like cars or property. A partial list includes:

- How much money do you spend on clothing?
- How much money do you spend on vehicle maintenance and repairs expenses?
- How much money do you spend on medical expenses?
- How much money do you spend education?

Why me? Why not interview someone else?

Through a scientific sampling process, we selected your address, not you personally. Your household represents hundreds of other households in your region, so it is important that we talk to you. Only you can accurately report how you spent your money.

What if I am retired, ill, unemployed, or just don't spend much money?

We are interested in how all Americans spend their money. We can only have a complete picture if we talk to people with different situations.

It is very important to know the purchasing habits of people of all ages and of all levels of spending. This is especially true for medical spending and spending on entertainment.

How long will the interview take?

The average interview takes about 25 minutes.

Where can I find out more about the survey?

You can learn more about the Consumer Expenditure Survey by writing to the Division of Consumer Expenditure Surveys, Room 3985, 2 Massachusetts Avenue, N.E., Washington, DC 20212. The Consumer Expenditure Survey Web site also has information about the survey. The address is <<http://www.bls.gov/cex/home.htm>>.

The U.S. Office of Management and Budget has approved this survey and assigned it Control Number 1220-0050 [UPDATE].

CE-303-L2 (MS2)
(6-2010)



UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001
OFFICE OF THE DIRECTOR

FROM THE DIRECTOR
U.S. CENSUS BUREAU

About two weeks ago, we sent you a letter informing you that you were selected to participate in the Consumer Expenditure Telephone Survey. We are contacting you again to remind you of the interview for this very important survey.

Your responses will help to improve the quality of data used to update the Consumer Price Index (CPI). The CPI is the most important tool used to measure how fast prices are rising or declining. It directly affects wages, pensions, and the cost of goods and services. Your response is a service to your community and the country.

To help shorten the interview, you may want to have some records of expenses available for reference. On the back of this letter are examples of records that may reduce the interview time and improve the quality of survey results.

We have also enclosed an Information Booklet with this letter. This booklet contains examples and definitions that you will find helpful during the interview.

Your participation in this survey is essential; however, you may choose to decline to answer any particular question. Federal law authorizes the collection of this information (Titles 13 and 29 of the United States Code), and Sections 9 and 214 of Title 13 require us to keep all information about you and your household strictly confidential, and to use that information for statistical purposes only.

Soon a Census Bureau interviewer will call you for an interview. If you would prefer to schedule an appointment for your interview, please call us toll-free at 1-888-595-1257.

Thank you for your cooperation in this important survey. We appreciate your help.

Sincerely,

Robert M. Groves

Enclosure

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

www.census.gov

The U.S. Census Bureau is conducting this voluntary survey
for the Bureau of Labor Statistics.

Records/Receipts Check List

**Collecting bills, receipts, or bank statements before the
interviewer contacts you may help reduce the interview time.
Examples of useful documents are shown below.**

Billing statements (paper or on-line) for:

- Medical expenses
- Education expenses
- Insurance payments

Receipts or bills of sales for:

- Appliances
- Clothing
- Travel or vacation expenses
- Gifts or charitable donations

Credit card statements for:

- Expenses charged automatically to your credit card
- Credit card membership fees
- Any expenditures with no receipt or statement

Bank statements or check books for:

- Expenses automatically deducted from your bank account
- ATM fees and other banking fees
- Any expenditures with no receipt or statement

The U.S. Office of Management and Budget has approved this survey and assigned it
Control Number 1220-0050 [UPDATE].

CE-303-L2 (MS3)
(6-2010)



UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001
OFFICE OF THE DIRECTOR

FROM THE DIRECTOR
U.S. CENSUS BUREAU

About three months ago, you completed an interview for the Consumer Expenditure Telephone Survey. We are contacting you again to request your ongoing help with this very important survey.

Your responses will help to improve the quality of data used to update the Consumer Price Index (CPI). The CPI is the most important tool used to measure how fast prices are rising or declining. It directly affects wages, pensions, and the cost of goods and services. Your response is a service to your community and the country.

To help shorten the interview, you may want to have some records of expenses available for reference. On the back of this letter are examples of records that may reduce the interview time and improve the quality of survey results.

We have also enclosed an Information Booklet with this letter. This booklet contains examples and definitions that you will find helpful during the interview.

Your participation in this survey is essential; however, you may choose to decline to answer any particular question. Federal law authorizes the collection of this information (Titles 13 and 29 of the United States Code), and Sections 9 and 214 of Title 13 require us to keep all information about you and your household strictly confidential, and to use that information for statistical purposes only.

Soon a Census Bureau interviewer will call you again for an interview. If you would prefer to schedule an appointment for your interview, please call us toll-free at 1-888-595-1257.

Thank you for your cooperation in this important survey. We appreciate your help.

Sincerely,

Robert M. Groves

Enclosure

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

www.census.gov

The U.S. Census Bureau is conducting this voluntary survey
for the Bureau of Labor Statistics.

Records/Receipts Check List

**Collecting bills, receipts, or bank statements before the
interviewer contacts you may help reduce the interview time.
Examples of useful documents are shown below.**

Billing statements (paper or on-line) for:

- Medical expenses
- Education expenses
- Insurance payments

Receipts or bills of sales for:

- Appliances
- Clothing
- Travel or vacation expenses
- Gifts or charitable donations

Credit card statements for:

- Expenses charged automatically to your credit card
- Credit card membership fees
- Any expenditures with no receipt or statement

Bank statements or check books for:

- Expenses automatically deducted from your bank account
- ATM fees and other banking fees
- Any expenditures with no receipt or statement

The U.S. Office of Management and Budget has approved this survey and assigned it
Control Number 1220-0050 [UPDATE].

CE-303-L2(MS4)
(6-2010)



UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001
OFFICE OF THE DIRECTOR

FROM THE DIRECTOR
U.S. CENSUS BUREAU

A month ago, you completed an interview for the Consumer Expenditure Telephone Survey. We are contacting you again to request your ongoing help with this very important survey.

Your responses will help to improve the quality of data used to update the Consumer Price Index (CPI). The CPI is the most important tool used to measure how fast prices are rising or declining. It directly affects wages, pensions, and the cost of goods and services. Your response is a service to your community and the country.

To help shorten the interview, you may want to have some records of expenses available for reference. On the back of this letter are examples of records that may reduce the interview time and improve the quality of survey results.

We have also enclosed an Information Booklet with this letter. This booklet contains examples and definitions that you will find helpful during the interview.

Your participation in this survey is essential; however, you may choose to decline to answer any particular question. Federal law authorizes the collection of this information (Titles 13 and 29 of the United States Code), and Sections 9 and 214 of Title 13 require us to keep all information about you and your household strictly confidential, and to use that information for statistical purposes only.

Soon a Census Bureau interviewer will call you again for an interview. If you would prefer to schedule an appointment for your interview, please call us toll-free at 1-888-595-1257.

Thank you for your cooperation in this important survey. We appreciate your help.

Sincerely,

Robert M. Groves

Enclosure

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

www.census.gov

The U.S. Census Bureau is conducting this voluntary survey
for the Bureau of Labor Statistics.

Records/Receipts Check List

**Collecting bills, receipts, or bank statements before the
interviewer contacts you may help reduce the interview time.**

Examples of useful documents are shown below.

Billing statements (paper or on-line) for:

- Medical expenses
- Education expenses
- Insurance payments

Receipts or bills of sales for:

- Appliances
- Clothing
- Travel or vacation expenses
- Gifts or charitable donations

Credit card statements for:

- Expenses charged automatically to your credit card
- Credit card membership fees
- Any expenditures with no receipt or statement

Bank statements or check books for:

- Expenses automatically deducted from your bank account
- ATM fees and other banking fees
- Any expenditures with no receipt or statement

The U.S. Office of Management and Budget has approved this survey and assigned it
Control Number 1220-0050 [UPDATE].

CE-303-L5(MS5)
(6-2010)



UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001
OFFICE OF THE DIRECTOR

FROM THE DIRECTOR
U.S. CENSUS BUREAU

Very soon a Census Bureau interviewer will call you again to conduct the final interview for your household for the Consumer Expenditure Telephone Survey. In this interview, as in previous ones, we will ask about your household expenditures.

The Consumer Price Index is the most important tool used to measure how fast prices are rising or declining and directly affects wages and pensions. Your response is a service to your community and the country.

To help shorten the interview, you may want to have some records of expenses available for reference. On the back of this letter are examples of records that may reduce the interview time and improve the quality of survey results.

We have also enclosed an Information Booklet with this letter. This booklet contains examples and definitions that you will find helpful during the interview.

Your participation in this survey is essential; however, you may choose to decline to answer any particular question. Federal law authorizes the collection of this information (Titles 13 and 29 of the United States Code), and Sections 9 and 214 of Title 13 require us to keep all information about you and your household strictly confidential, and to use that information for statistical purposes only.

If you would prefer to call to set up a specific time for your interview, we can be reached toll free at 1-888-595-1257.

Thank you for your cooperation and voluntary participation in this important survey. We appreciate your help.

Sincerely,

Robert M. Groves

Enclosure

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

www.census.gov

The U.S. Census Bureau is conducting this voluntary survey
for the Bureau of Labor Statistics.

Records/Receipts Check List

**Collecting bills, receipts, or bank statements before the
interviewer contacts you may help reduce the interview time.**

Examples of useful documents are shown below.

Billing statements (paper or on-line) for:

- Medical expenses
- Education expenses
- Insurance payments

Receipts or bills of sales for:

- Appliances
- Clothing
- Travel or vacation expenses
- Gifts or charitable donations

Credit card statements for:

- Expenses charged automatically to your credit card
- Credit card membership fees
- Any expenditures with no receipt or statement

Bank statements or check books for:

- Expenses automatically deducted from your bank account
- ATM fees and other banking fees
- Any expenditures with no receipt or statement

The U.S. Office of Management and Budget has approved this survey and assigned it
Control Number 1220-0050 [UPDATE].

Appendix IV

MIS Information Booklet



Consumer Expenditure Telephone Survey

Information Booklet

TABLE OF CONTENTS

| Page | |
|-------------|--|
| 2 | Hispanic Origin |
| 2 | Race |
| 2 | Asian Origin |
| 3 | Education |
| 3 | Armed Forces |
| 4 – 10 | Appliances, Household Equipment and Other Selected Items |
| 11 – 12 | Clothing |
| 13 – 14 | Vehicle Operating Expenses |
| 15 | Insurance Other than Health |
| 16 | Hospitalization and Health Insurance |
| 17 | Educational Expenses |
| 18 – 19 | Subscriptions, Memberships, Books, and Entertainment Expenses |
| 20 – 21 | Trips and Vacations |
| 22 | Privacy Act Statement |

Hispanic Origin

- 1 – Mexican
- 2 – Mexican-American
- 3 – Chicano
- 4 – Puerto Rican
- 5 – Cuban
- 6 – Cuban-American
- 7 – Central or South American
- 8 – Other

Race

(Please choose one or more)

- 1 – White
- 2 – Black or African American
- 3 – American Indian or Alaskan Native
- 4 – Asian
- 5 – Native Hawaiian or Other Pacific Islander

Asian Origin

- 1 – Chinese
- 2 – Filipino
- 3 – Japanese
- 4 – Korean
- 5 – Vietnamese
- 6 – Asian Indian
- 7 – Other group not listed



Education

- 0** – Never attended, preschool, kindergarten
- 1 - 11** – 1st grade through 11th grade
- 38** – 12th grade, no diploma
- 39** – High school graduate – high school diploma, or the equivalent (For example: GED)
- 40** – Some college but no degree
- 41** – Associate degree in college – Occupational/Vocational
- 42** – Associate degree in college – Academic program
- 43** – Bachelor's degree (For example: BA, AB, BS)
- 44** – Master's degree (For example: MA, MS, MEng, MEd, MSW, MBA)
- 45** – Professional School Degree (For example: MD, DDS, DVM, LLB, JD)
- 46** – Doctorate degree (For example: PhD, EdD)

Armed Forces

A person is considered to be in the armed forces if he or she serves in any branch of the U.S. military. This includes the Army, Navy, Marine Corps, Air Force and Coast Guard, their Reserve components, and the Air and Army National Guard.



APPLIANCES, HOUSEHOLD EQUIPMENT AND OTHER SELECTED ITEMS

Part A – Major Household Appliances

- 1** – Microwave oven
- 2** – Cooking stove, range, or oven
(including convection oven)
- 3** – Range hood
- 4** – Refrigerator or home freezer
- 5** – Built-in dishwasher
- 6** – Portable dishwasher
- 7** – Garbage disposal
- 8** – Clothes washer
(including washer/dryer combinations)
- 9** – Clothes dryer

APPLIANCES, HOUSEHOLD EQUIPMENT AND OTHER SELECTED ITEMS (continued)

Part B – Household Appliances and Other Selected Items

SMALL HOUSEHOLD APPLIANCES

1 – Small electrical kitchen appliances, including –

| | | |
|---------------------|-----------------------|----------------|
| blender | electric frying pan | mixer |
| breadmaker | electric iron | pizza oven |
| coffee grinder | electric knife | popcorn maker |
| coffee maker | electric timer | rice cooker |
| crockpot | electric wine chiller | sandwich grill |
| deep fryer | electric wok | slow cooker |
| electric barbecue | food processor | smoothie maker |
| electric can opener | hot plate | toaster |
| electric grill | ice cream maker | toaster oven |
| electric fondue set | juicer | waffle iron |

2 – Electrical personal care appliances, including –

| | | |
|-----------------------|---------------------|----------------|
| curling iron | electric toothbrush | make-up mirror |
| denture cleaner | facial sauna | massager |
| electric hair trimmer | foot bath | Waterpik |
| electric razor | hair dryer | |
| digital scale | heating pad | |

3 – Smoke detectors, including –

| | | | |
|-------|------------------|-------------------------|-----------------|
| wired | battery operated | ionization chamber type | photo-cell type |
|-------|------------------|-------------------------|-----------------|

4 – Electrical floor cleaning equipment, including –

| | | | |
|----------------|-------------|---------------|----------------|
| vacuum cleaner | hand vacuum | rug shampooer | floor polisher |
|----------------|-------------|---------------|----------------|

5 – Other household appliances, including –

| | |
|--------------------------|--|
| trash compactor | home security device (burglar alarm) including |
| air purifier | console, control modules, burglar alarm console, |
| water filters | door and window transmitters |
| carbon monoxide detector | |

6 – Sewing machines (with or without cabinet)

7 – Other office machines including fax machines and calculators, also including –

| | |
|-------------|---------------|
| typewriters | copy machines |
|-------------|---------------|

8 – Personal Digital Assistant or PDAs, including -

| | |
|------|------|
| Palm | iPaq |
|------|------|



APPLIANCES, HOUSEHOLD EQUIPMENT
AND OTHER SELECTED ITEMS
continued

APPLIANCES, HOUSEHOLD EQUIPMENT AND OTHER SELECTED ITEMS (continued)

Part B – Household Appliances and Other Selected Items (continued)

9 – Computers, computer systems, and related hardware for non-business use, including -

| | | |
|-------------------|------------|---|
| CD/DVD drive | cables | home computers with or without monitors |
| computer printers | fax modems | external hard drive |
| keyboards | scanner | CD/DVD burner |
| modems | memory | mouse |
| monitors | Micro PC | laptops |

10 – Computer software, including computer games, or accessories, for non-business use, including -

| | | |
|----------|--------------------|------------|
| PC games | printer cartridges | mouse pads |
|----------|--------------------|------------|

11 – Video game hardware, video games, or accessories, including -

| | | | |
|--------------|-------------|-------------|-----|
| Nintendo Wii | Gamecube | Nintendo DS | PSP |
| Gameboy | Playstation | Xbox | |

12 – Telephones or accessories, including –

| | | | |
|-----------------------|----------|-----------------------|--------------|
| telephones | headsets | cordless telephones | car chargers |
| beepers | chargers | cell phone covers | cell phones |
| phone jacks and cords | pagers | Bluetooth accessories | |

13 – Telephone answering machines, including – combinations of telephone/answering machines

14 – Photographic equipment, including –

| | | | |
|----------------|-----------|-------------------|-------------------------------|
| camera | filter | projection screen | battery pack for camera flash |
| digital camera | winder | electro flash | motor driven film advancer |
| lens | enlarger | strobe light | |
| tripod | projector | (for photographs) | |

Do not include film, film processing, or other photographic supplies.

15 – Musical instruments, supplies, and accessories, including –

| | | | |
|-------------------|---------------------------------|-------------|-------------------------------|
| piano | sheet music | saxophone | music books |
| guitar | strings for musical instruments | music stand | trumpet |
| woodwinds | | clarinet | any other musical accessories |
| brass instruments | stringed instruments | picks | carrying case |
| trombone | | rosin | Keyboards |
| reeds | valve oil | drums | |

Do not include music lessons.

APPLIANCES, HOUSEHOLD EQUIPMENT AND OTHER SELECTED ITEMS (continued)

Part B – Household Appliances and Other Selected Items (continued)

16 – Lawn mowing machinery and other yard equipment, including -

| | | | |
|---------------------------------|------------------|--------------------------|----------|
| lawn mowers | motorized tiller | snow blower | shovel |
| tractor (farm, garden, etc.) | wheelbarrow | electric lawn trimmer | spreader |
| | rake | | |

TOOLS FOR HOME USE

17 – Power tools, including –

| | | |
|----------------|-------------------|--|
| electric drill | sander | cordless circular saw |
| electric saw | lathe | electric swimming pool cleaning equipment |
| router | electric plane | air compressor |
| cordless drill | electric polisher | |

18 – Non-power tools including –

| | | | |
|--------|-------------|-------|---------------|
| wrench | axe | saw | drill |
| socket | screwdriver | level | trouble light |
| hammer | pliers | plane | caulking gun |

HEATING AND COOLING EQUIPMENT

19 – Window air conditioners

20 – Portable cooling and heating equipment, including –

| | | | |
|--------------|--------------|------------|-----|
| space heater | dehumidifier | humidifier | fan |
|--------------|--------------|------------|-----|



APPLIANCES, HOUSEHOLD EQUIPMENT
AND OTHER SELECTED ITEMS
continued

APPLIANCES, HOUSEHOLD EQUIPMENT AND OTHER SELECTED ITEMS (continued)

Part B – Household Appliances and Other Selected Items (continued)

TELEVISIONS, RADIOS, VIDEO AND SOUND EQUIPMENT (not installed on vehicles)

21 – Televisions, all types including combinations of TVs with DVD/video players, including –
flat screen TV plasma TV high definition TV

22 – DVD players, VCRs, DVRs, or video cameras, including –
TiVo unit digital TV converter box
Blu-ray disc player video tape recorder/player
high definition disc player video laser disc player
combination of VCR/DVD player video cassette recorder/player

23 – Satellite dishes, receivers, or accessories

24 – Handheld personal music players, including –
iPod personal mp3 players

25 – Radio, all types, including –
CB (not permanently clock radio short-wave transistor/portable
mounted in an console walky-talky Walkman (radio only)
automobile)

26 – Tape recorders or players (not permanently mounted in an automobile), including –
audio tape decks audio cassette players/recorders
reel-to-reel tape decks Walkman (cassette/radio combination or cassette only)

27 – Sound components, component systems, or compact disc sound systems, including –
speakers amplifier tape deck (not specified)
mixer turntable compact disc players
stereo receiver stereo rack system
tuner equalizer

28 – Other sound or video equipment, including –
earphones/headphones battery packs adapter for sound equipment
power converter power booster base station CB antenna
antenna (TV, radio, etc.) headset microphone

29 – Portable memory, such as flash drives, memory cards, and recordable discs and tapes, including –
thumb drives blank DVDs blank CDs
zip drives memory stick USB flash drive

APPLIANCES, HOUSEHOLD EQUIPMENT AND OTHER SELECTED ITEMS (continued)

Part B – Household Appliances and Other Selected Items (continued)

SPORTS, RECREATION, AND EXERCISE EQUIPMENT

30 – General sports equipment, including –

| | | | |
|----------------|--------------------|------------------------|-------------------|
| roller blades | baseball bat | table tennis | badminton set |
| sports uniform | football | equipment | soccer ball |
| sports shoes | basketball | lawn games | sports protective |
| tennis racket | racquetball | frisbee | equipment/gear |
| bowling ball | racquetball racket | boxing equipment | golf clubs |
| baseball glove | volleyball | karate equipment | basketball hoop |
| skateboard | golf shoes | golf cart (non-riding) | |

Include specialized athletic shoes such as for football, baseball, soccer, biking, and bowling, except if included in the rental or activity fee for the sport.

31 – Health and exercise equipment, including –

| | | |
|----------------|--------------|--------------|
| trampoline | exercise mat | weight bench |
| weights | home gym | treadmill |
| rowing machine | exercycle | pedometer |

32 – Camping equipment, including –

| | | |
|---------------------|--------------|------------------|
| tent | sleeping bag | camping stove |
| kerosene lamp | sleeping pad | camping cookware |
| frame packs and | air mattress | portable heater |
| other camping packs | canteen | |

33 – Hunting and fishing equipment, including –

| | | |
|------------------------|---------|------------------------|
| fishing rod and tackle | knife | BB/pellet gun |
| bow and arrow | rifle | ammunition |
| crossbow | shotgun | scopes (not specified) |

34 – Winter sports equipment, including –

| | | |
|---------------------|--------------------|----------|
| snow skis | ski poles | toboggan |
| ski boots | ice skates | sled |
| snowboard | ice boat | |
| snowboard equipment | sledding equipment | |

APPLIANCES, HOUSEHOLD EQUIPMENT
AND OTHER SELECTED ITEMS
continued

APPLIANCES, HOUSEHOLD EQUIPMENT AND OTHER SELECTED ITEMS (continued)

Part B – Household Appliances and Other Selected Items (continued)

35 – Water sports equipment, including –

| | | |
|-------------|------------------|-----------------|
| water skis | snorkel | surf board |
| life jacket | water ski vest | raft |
| wake board | diving equipment | wind surf board |

36 – Outboard motors

37 – Bicycles, including –

| | | | |
|-----------------|-------|-------|----------|
| bicycle helmets | stand | tires | tubes |
| bicycle parts | locks | rack | supplies |

38 – Tricycles or battery powered riders, including – big wheels

39 – Playground equipment, including –

| | | | |
|------------------------|-----------|----------|---------|
| portable swimming pool | swing set | sand box | gym set |
|------------------------|-----------|----------|---------|

40 – Other sports or recreation equipment, including –

| | |
|---------------------|----------------|
| telescope | metal detector |
| paintball equipment | |

CLOTHING

Clothing

1 – Coats, jackets, and furs, including -

| | | | |
|-------------|-----------|------------------|-------------|
| shawl | raincoat | fur coat | winter coat |
| windbreaker | down vest | outerwear jacket | |

2 – Sport coats and tailored jackets, including blazers

3 – Suits, including –

| | |
|--------------------------------------|-------------|
| woman's suit (of two or more pieces) | formal suit |
| man's suit (of two or more pieces) | |

4 – Vests (purchased separately, not with a suit), excluding sweater vests and down vests

5 – Sweaters and sweater sets, including –

| | | |
|--------------|-------------|----------------|
| cardigan | pullover | V-neck sweater |
| sweater vest | ski sweater | |

6 – Pants, slacks, or shorts, including –

| | | | |
|-------------|------------|-----------------|--------------|
| jump suit | blue jeans | maternity pants | dress slacks |
| dress pants | overalls | short sets | casual pants |

Do not include any athletic shorts.

7 – Dresses, including –

| | | | |
|-------------------|--------------|-------------------|-------------------------|
| two-piece dresses | wedding gown | maternity dresses | formals or semi-formals |
|-------------------|--------------|-------------------|-------------------------|

8 – Skirts, including skorts

Do not include any tennis skirts, golf skirts, or other athletic skirts.

9 – Shirts, blouses, and tops, including –

| | | |
|--------------|--------------|----------------|
| sport shirts | tops | maternity tops |
| dress shirts | knit blouses | T-shirts |

Do not include any sweat shirts or athletic shirts.

10 – Undergarments, including –

| | | |
|-----------|-------------|-------------------|
| bras | undershirts | slips |
| shapewear | underwear | thermal underwear |

11 – Hosiery, including –

| | | | |
|-------|------------|--------|-----------|
| socks | knee-highs | tights | pantyhose |
|-------|------------|--------|-----------|

12 – Nightwear and loungewear, including –

| | | |
|---------|-------------|---------------------------|
| pajamas | night shirt | night gown |
| robe | house coat | thermal sleeping garments |

CLOTHING
continued

CLOTHING (continued)

Clothing (continued)

13 – Accessories, including -

| | | | |
|-----------|---------|------------------|-----------------------------|
| umbrellas | gloves | apron | fold-up rain accessories |
| belts | mittens | ear muffs | bandannas |
| ties | purse | handkerchiefs | hair accessories |
| scarves | wallet | bridal headpiece | non-prescription sunglasses |

14 – Swimsuits or warm-up or ski suits, including –

| | | | |
|-----------------|---------------|------------|----------------------|
| athletic shorts | tennis outfit | sweatshirt | swimwear accessories |
| athletic shirt | jogging suit | swimwear | snow and ski suit |
| hunting wear | leotards | | |

Do not include any sports uniforms.

15 – Uniforms, other than sport, for which the cost is not reimbursed, including shirts, pants, suits, service apparel, such as: medical, barber, boy or girl scout, mechanic, waiter/waitress, plumber and lab smocks, and military apparel

16 – Costumes, including costumes for dance, ballet, Halloween, etc.

17 – Footwear, including –

| | | |
|--------------|---------|--|
| dress shoes | sandals | bedroom slippers |
| casual shoes | boots | sneakers, jogging, aerobic, basketball, tennis shoes |

Do not include specialized athletic shoes such as for football, soccer, bowling, biking, or baseball.

VEHICLE OPERATING EXPENSES

Vehicle Maintenance and Repair, Parts, and Equipment

1 – Oil change, lubrication, and oil filter

(Include oil only if purchased for an oil change)

2 – Motor tune-up, including –

| | | | |
|------------------------|-----------------|------------------|-----------------------------------|
| air/fuel filters | computer sensor | ignition wires | ignition timing or mixture |
| distributor cap, rotor | PCV valve | valve adjustment | adjustment |
| emission controls | spark plugs | | breather/vapor/air filter element |

3 – Brake work, including –

| | | | |
|------------------|------------------|----------------|----------------------|
| anti-lock brakes | hydraulic system | shoes or pads | bleed brake system |
| master cylinder | wheel calipers | wheel cylinder | machine drums/rotors |
| brake adjustment | parking brake | | |

4 – Battery purchase and installation

5 – Tire purchase and mounting

6 – Tire repairs

7 – Front end alignment, wheel balancing, wheel rotation

8 – Steering or front-end work, including –

| | | | |
|--------------------|-----------------|----------------|-----------------------------|
| axel bearing/seals | bushings | studs/lug nuts | power steering fluid/filter |
| axle shafts | CV joints/boots | tie rods | idler arms |
| rack and pinion | wheel hubs | ball joints | steering box/linkage |

9 – Electrical system work, including –

| | | | |
|----------------------|--------------|-----------------|--------------------|
| alternator belt | car computer | ignition system | voltage regulator |
| alternator/generator | coil | starter motor | gauges/instruments |
| battery charge | switches | wiring | |

10 – Engine repair or replacement, including –

| | | | |
|------------------------|---------------|--------------|----------------------------|
| carburetor | fuel injector | motor mounts | timing chain/gears or belt |
| fuel pump/lines/filter | turbo charge | pistons/rods | oil pump/cooler/hoses |
| crankshaft bearings | gaskets | choke | |

11 – Air conditioning work, including –

| | | | |
|------------|-----------|--------------|------------|
| compressor | condenser | motor/switch | recharging |
|------------|-----------|--------------|------------|

12 – Engine cooling system work, including –

| | | | |
|-------------------|-------------------|--------------------|------------------------|
| coolant or filter | heater core | radiator | fan or water pump belt |
| hoses | thermostat | water pump | cooling fan/controls |
| pressure cap | cooling fan relay | fan switch or belt | |

VEHICLE OPERATING EXPENSES
continued

VEHICLE OPERATING EXPENSES (continued)

Vehicle Maintenance and Repair, Parts, and Equipment (continued)

13 – Exhaust system work, including –

| | | | |
|---------------------|-----------------|---------|-----------|
| catalytic converter | hanger/clamps | muffler | resonator |
| exhaust pipe | manifold gasket | | |

14 – Clutch or transmission work, including –

| | | | |
|--------------|------------------|----------------------|---------------------|
| clutch cable | hydraulic system | rebuilt transmission | transaxle |
| clutch fork | master cylinder | safety switch | transmission filter |
| flywheel | pilot bearing | shaft seal | |

15 – Body work and painting, including –

| | | | |
|-------------------|-----------|---------------|---------------------------|
| convertible top | doors | T-roof | crash repairs |
| glass replacement | vinyl top | rust proofing | window repair/replacement |
| sanding | | | |

16 – Shock absorber replacement, including MacPherson struts

17 – Drive shaft or rear-end work, including –

| | | | |
|---------------------|-----------------|----------------------|----------------------|
| axle fluid | CV joints | rear axle | suspension |
| differential | tie rods | grommet | rear wheel axle seal |
| rear wheel bearings | universal joint | coil or leaf springs | axle mounts/bushings |

18 – Video equipment and installation, including –

| | |
|--|---|
| televisions and combinations of TVs with VCRs and/or DVD players Video game consoles | GPS navigational system with screen satellite receiver, In-Motion satellite receiver |
|--|---|

19 – Audio equipment and installation, including –

| | | | |
|------------|----------|------------------|---------------------------|
| antenna | CB radio | speakers | tape player |
| CB antenna | radio | stereo equipment | satellite radio equipment |

20 – Vehicle accessories and customizing, including –

| | | | |
|----------------|---------------|----------------|-----------------------|
| alarm system | carpeting | running boards | steering wheel covers |
| bike/ski racks | fender skirts | seat covers | spoilers |
| bumper guards | luggage rack | | |

21 – Vehicle cleaning services and cleaning supplies, including –

| | | | |
|------------------------|-------------------|--------------------|----------------------------|
| car washes | vacuuming | cleaning mitts | vehicle detailing services |
| waxes | upholstery sprays | protective coating | wheel cleaning supplies |
| boat cleaning services | | | |

22 – Other Vehicle Services, parts, or equipment, including –

| | | | |
|-------------------|--------------|---------------|---------------------------|
| battery cables | vent filters | jack | charcoal canister-filters |
| gas cable/cap/can | brake lights | gasket sets | windshield wipers |
| light bulbs | wheel lugs | headlights | speedometer cable |
| wheels/rims | hub caps | heater repair | tire/wheel combination |
| upholstery work | | | |

INSURANCE OTHER THAN HEALTH

1 – Long term care insurance

2 – Life insurance or other policies which provide benefits in case of death or disability, including –

| | |
|----------------------|--------------------------------|
| term insurance | income or disability insurance |
| whole-life insurance | group-life insurance |
| cash benefits | straight-life insurance |
| mortgage insurance | flight insurance |
| veterans insurance | life endowments |
| annuities | burial insurance |

HOME INSURANCE

Insurance protecting your home, furniture, personal effects, or other property against fire, theft, loss, natural disasters, or damage from other means.

3 – Homeowners' insurance, including any insurance covered in mortgage payments or flood insurance

| | |
|-----------------|----------------------------|
| flood insurance | fire and extended coverage |
|-----------------|----------------------------|

4 – Renters' insurance

5 – Automobile or other vehicle insurance, including –

| | |
|-------------------------|---------------------------|
| liability insurance | bodily injury insurance |
| collision insurance | property damage insurance |
| comprehensive insurance | no-fault insurance |

6 – Other types of non-health insurance, including –

- credit card insurance
- personal liability insurance
- mortgage guarantee insurance
- ambulance
- umbrella policies

Do not include malpractice insurance.

HOSPITALIZATION AND HEALTH INSURANCE

1 – Health Maintenance Organization

Expenses usually covered in full, or there may be a modest co-payment at the time of your visit.

- Group/staff type: You go to a central facility (group health center) to receive care.
- Independent practice association (IPA): Providers work from their individual offices (and are referred to as primary care physicians).

2 – Fee for Service Plan

You or your insurance company are generally billed after each visit.

- Traditional fee for service plan: You may go to any doctor or hospital you choose.
- Preferred Provider Organization (PPO): You are given a list of doctors from which to choose. If you go to a doctor on the PPO list, more expenses are covered than if you go to a doctor not on the list.

3 – Commercial Medicare Supplement

Voluntary contributory private insurance plan available to Medicare recipients. Covers the costs of deductibles, co-insurance, physician services, and other medical and health services.

4 – Special Purpose Plan

Covers only specific health needs. Examples of special purpose health insurance plans are:

dental insurance

vision insurance

prescription drug insurance

mental health insurance

dread disease policy

Do not include Medicare Prescription Drug plans.

EDUCATIONAL EXPENSES

1 – Recreational lessons or other instructions, including –

| | | | |
|-----------------|------------------|---------|-------------------------|
| golf | dancing | driving | tennis |
| music | photography | skiing | painting |
| sailing | swimming | sewing | instructional day camps |
| skydiving | needlepoint | cooking | self defense |
| driving lessons | horseback riding | | |

2 – Nursery school or child day care centers, including non-instructional day camps

3 – Tuition, including –

| | | | |
|-----------------------|-------------------|--------------------|---------------------------|
| college or university | business school | elementary school | high school |
| secretarial school | parochial school | seminary | technical school |
| adult education | vocational school | preparatory school | middle/junior high school |

*Include only those expenses paid directly to the school or to other educational facility.
Do not report student loans.*

4 – Housing while attending school, including –

| | | | |
|-------------------|----------|------------|------------------------------|
| student dormitory | sorority | fraternity | housing for married students |
|-------------------|----------|------------|------------------------------|

Include only those expenses paid directly to the school or to other educational facility.

5 – Food or board while attending school

Include only those expenses paid directly to the school or to other educational facility.

6 – Private school bus

7 – Test preparation or tutoring services

8 – Purchase of any school books, supplies, or equipment, which has not already been reported, including –

| | | | |
|-------------|--------------|--------------------|----------------------|
| text books | ruler | drafting equipment | laboratory equipment |
| microscopes | art supplies | technical books | cap and gown |

9 – Other school related expenses not already reported, including the rental of any school books or expenses

| | | | |
|----------------------|---------------------|----------------------------|--------------------------|
| laboratory fees | matriculation fees | transportation fees | conferences and seminars |
| registration fees | health fees | student union fees | |
| laundry fees | athletic fees | rental of school books | |
| cap and gown rentals | administration fees | rental of school equipment | |

Do not report student loans.

SUBSCRIPTIONS, MEMBERSHIPS, BOOKS, AND ENTERTAINMENT EXPENSES

Part A – Subscriptions and Memberships

- 1** – Subscriptions to newspapers, magazines, or periodicals. Include online subscriptions
- 2** – Books purchased from book club
- 3** – Season tickets to theater, concert series, opera, other musical series, or amusement parks
- 4** – Season tickets to sporting events
- 5** – Encyclopedias or other sets of reference books
- 6** – Golf courses, country clubs and other social organizations
- 7** – Health clubs, fitness centers, swimming pools, weight loss centers, or other sports and recreational organizations
- 8** – Vacation clubs
- 9** – Civic, service, or fraternal organization
- 10** – Credit card membership fees
- 11** – Shopping club memberships such as COSTCO and SAM'S
- 12** – Services that use Global Positioning Systems (GPS) such as OnStar, not already reported
- 13** – Direct or online dating services

SUBSCRIPTIONS, MEMBERSHIPS, BOOKS, AND ENTERTAINMENT EXPENSES (continued)

Part B – Books and Entertainment Expenses

Fees for participation in sports, including –

tennis golf bowling swimming
billiards

Single admissions to spectator sporting events, including –

football baseball hockey soccer
auto racing basketball

Single admissions to performances, including –

movies operas plays concerts

Single admissions to other entertainment activities, including –

museums zoos state parks amusement parks
historic sites

Books not purchased through book clubs, including –

paperbacks hardcover audio digital books

Exclude reference books or school books.

Single copies of newspapers, magazines, periodicals (non-subscription)

Compact discs, audio tapes, or records

Do not include blank or recordable CDs or blank or recordable audio tapes.

Photographic film, including disposable cameras

Photo processing, including -

digital photo processing video film processing

Purchase of video tapes or DVDs other than through a mail-order club

Do not include blank or recordable DVDs or blank or recordable video tapes.

Rental of video tapes or DVDs, including –

mail delivery DVD rentals

SUBSCRIPTIONS, MEMBERSHIPS, BOOKS, AND ENTERTAINMENT EXPENSES (continued)

Part C – Summary Questions

Subscriptions, Books, and Magazines

Subscriptions to newspapers, magazines, or periodicals. Include online subscriptions

Single copies of newspapers, magazines, periodicals (non-subscription)

Books purchased from book club

Books not purchased through book clubs, including –

paperbacks

hardcover

Audio

digital books

Exclude reference books or school books.

Encyclopedias or other sets of reference books

Tickets/Admission for Movies, Concerts, and Sporting Events

Season tickets to theater, concert series, opera, other musical series, or amusement parks

Season tickets to sporting events

Single admissions to spectator sporting events, including

Football

baseball

Hockey

soccer

auto racing

basketball

Single admissions to performances, including

Movies

operas

plays

concerts

Single admissions to other entertainment activities, including –

museums

zoos

state parks

amusement parks

historic sites

SUBSCRIPTIONS, MEMBERSHIPS, BOOKS, AND ENTERTAINMENT EXPENSES (continued)

Part C – Summary Questions (continued)

Membership Fees, including

Golf courses, country clubs and other social organizations

Health clubs, fitness centers, swimming pools, weight loss centers, or other sports and recreational organizations

Fees for participation in sports, including –

tennis golf bowling swimming
billiards

Vacation clubs

Civic, service, or fraternal organization

Credit card membership fees

Shopping club memberships such as COSTCO and SAM'S

DVDs, Photographic Film, and Audio Recordings, including

Purchase of video tapes or DVDs other than through a mail-order club

Do not include blank or recordable DVDs or blank or recordable video tapes.

Rental of video tapes or DVDs, including –

mail delivery DVD rentals

Photographic film, including disposable cameras

Photo processing, including -

digital photo processing video film processing

Compact discs, audio tapes, or records

Do not include blank or recordable CDs or blank or recordable audio tapes.

TRIPS AND VACATIONS

Types of Trips

- 1 – Visit friends or relatives
- 2 – Business trips
- 3 – Recreational trips, such as –
 - Sightseeing
 - Sports events
 - Club or organizational meetings
 - Outdoor recreation
- 4 – Any other trips that occur overnight or longer
- 5 – Any day trips to a place at least 75 miles away

Types of Transportation

COMMERCIAL

- 1 – local (taxi, etc.)
- 2 – airplane
- 3 – train
- 4 – bus
- 5 – ship

RENTED

- 6 – automobile
- 7 – truck, van
- 8 – motorcycle, moped
- 9 – private plane
- 10 – boat, trailer
- 11 – camper
- 12 – other vehicles

PRIVATE

- 13 – automobiles or other vehicles privately owned or leased by the household
- 14 – vehicle owned by someone else
- 15 – other transport

TRIPS AND VACATIONS (continued)

Rental of Sports Equipment

- Golf clubs
- Skis/snowboards
- Fishing equipment
- Boat
- Scuba/snorkeling equipment
- Other sports equipment

Fees for Playing Sports

- Golf
- Fishing
- Swimming
- Tennis
- Skiing/snowboarding
- Bowling
- Exercise classes
- Scuba/snorkeling
- Other sports

Entertainment or Admissions

- Movies
- Theater
- Concerts
- Museums
- Tours
- Sports events
- Other entertainment events

PRIVACY ACT STATEMENT

The U.S. Census Bureau is conducting the Consumer Expenditure Surveys for the Bureau of Labor Statistics of the U.S. Department of Labor under title 29, United States Code. The survey's purpose is to obtain information on what Americans are purchasing in order to update the Consumer Price Index (CPI). All survey information will be used for statistical purposes only.

Any information you provide for this survey is confidential, by law, under title 13, United States Code. Participation in this survey is voluntary and there are no penalties for refusing to answer any question(s). However, your cooperation is extremely important to help ensure the completeness and accuracy of these data.

Appendix V

MIS Survey Instrument Specifications - Housing

COVER SHEET

BHousing

Blaise Summary-Level Specifications Report

Name: H1

SAS Name: H1

Universe Description:

Question Text: Thinking about your primary residence, is it:

Answer List: TH1

| Value: | Mnemonic: | Description: |
|---------------|------------------|---|
| 1 | OwnMort | Owned by you or someone in this household with a mortgage or loan? |
| 2 | OwnNoMort | Owned by you or someone in this household without a mortgage or loan? |
| 3 | Rented | Rented? |
| 4 | NoRentMort | Occupied without payment of rent or mortgage? |

Skip Instructions: <1> [goto MORT]
<2,4,D,R> [exit block and if QTYPE = 2 then goto Section 9 -
Bsect9
else goto Section 6 - Bsect6PT]
<3> [goto RENT]

Name: MORT

SAS Name: MORT

Universe Description: Residence is owned, by someone in this household, with a mortgage or loan

Question Text: How much is your monthly mortgage payment?

Enter dollar amount

Skip Instructions: <1-99999999,D,R> [exit block and IF QTYPE = 2, goto Section
9 - BSECT9. Else goto Section 6- Bsect6PT]

Name: RENT

SAS Name: RENT

Universe Description: Primary residence is rented

Question Text: How much is your monthly rent?

Enter dollar amount

Skip Instructions: <1-99999999,D,R> [exit block and If QTYPE = 2, goto Section
9 - BSect9.
else goto Section 6 - BSect6PT]

Name: REC_STAT

SAS Name: REC_STAT

Universe Description:

Question Text: ** CREATED IN POST_PROCESSING **

Name: REC_ORIG

SAS Name: REC_ORIG

Universe Description:

Question Text: ** CREATED IN POST-PROCESSING **

Skip Instructions: <1,4,5>

Appendix VI

MIS Survey Instrument Specifications - Income

COVER SHEET

BIncome

Blaise Summary-Level Specifications Report

Name: I1

SAS Name: I1

Universe Description:

Question Text: In addition to knowing about the spending patterns in your household, it would also help us to have an idea of your household income. Last year, that is from January 1, ^YEAR to December 31st, ^YEAR, was your combined household income from all sources, above or below \$75,000 before taxes?

Income sources may include wages and salaries, self-employment income, social security, private and government retirement, interest, dividends, rental income, and other property income, public assistance, supplemental security income, food stamps, regular contributions for support and other sources.

Answer List: I1_AnswerList

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Below75 | Below \$75,000 |
| 2 | Above75 | \$75,000 or more |

Skip Instructions: <1> [goto I1a]
<2> [goto I1b]
<D,R> [exit block and goto BPSAQ]

Name: I1a

SAS Name: I1a

Universe Description: Household income was below \$75,000

Question Text: Was it . . . ?

Read answer list categories

Answer List: I1a_AnswerList

| Value: | Mnemonic: | Description: |
|---------------|------------------|--|
| 1 | Less25 | Less than \$25,000 |
| 2 | Least25Less50 | At least \$25,000 but less than \$50,000 |
| 3 | Least50Less75 | At least \$50,000 but less than \$75,000 |

Skip Instructions: <1,2,3,D,R> [exit block and goto PSAQ_Intro]

Name: I1b

SAS Name: I1b

Universe Description: Household income was \$75,000 or more

Question Text: Was it . . . ?

Read answer list categories

Answer List: I1b_AnswerList

| Value: | Mnemonic: | Description: |
|---------------|------------------|--|
| 1 | Least75Less100 | At least \$75,000 but less than \$100,000 |
| 2 | Least100Less150 | At least \$100,000 but less than \$150,000 |
| 3 | More150 | More than \$150,000 |

Skip Instructions: <1,2,3,D,R> [exit block and goto the PSAQ Block]

Appendix VII

MIS Survey Instrument Specifications – Global Questions

COVER SHEET

BGlobal

Blaise Summary-Level Specifications Report

Name: GLOBAL_INT

SAS Name: GLOBAL_INT

Universe Description:

Question Text: The next few questions are going to focus on the total amount you have spent on various expenditure categories. Instead of asking you to report individual items and their costs, we would like you to tell us the total amount your household spent on all items in that particular category.

The questions that follow refer to the past three months, that is, from the first day of ^REF_MONTH up through today.

Answer List: Tcontinue

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Continue | Enter 1 to Continue |

Skip Instructions: <1> [IF QTYPE = 2 then goto G9a]
[ELSE goto G6a]

Name: G9a

SAS Name: G9a

Universe Description: Subsample A cases in a 2nd or 3rd month interview

Question Text: 11 - 12 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent for clothing, shoes, or accessories?

Skip Instructions: <0-2000> [goto G9a_FR]
<2001-9997> [goto ERR1_G9a]
<D,R> [goto G12a]

Name: G9a_FR

SAS Name: G9a_FR

Universe Description: Spent \$0 - \$1,997 on clothing, shoes, or accessories.

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G12a]

Name: G12a

SAS Name: G12a

Universe Description: Subsample A cases in a 2nd or 3rd month interview

Question Text: 13 - 14 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent for maintaining or repairing a car, truck, motorcycle, or other vehicle?

Please do NOT include expenses for gasoline or other fuels.

Skip Instructions: <0-4000> [goto G12a_FR]
<4001-9997> [goto ERR1_G12a]
<D,R> [goto G13a]

Name: G12a_FR

SAS Name: G12a_FR

Universe Description: Spent \$0 - \$1,997 on maintaining or repairing a vehicle/motorcycle

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G13a]

Name: G13a

SAS Name: G13a

Universe Description: Subsample A cases in a 2nd or 3rd month interview

Question Text: 15 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent for auto insurance and

home or property insurance?

Skip Instructions: <0-1000> [goto G13a_FR]
<1001-9997> [goto ERR1_G13a]
<D,R> [goto G13b]

Name: G13a_FR

SAS Name: G13a_FR

Universe Description: Spent \$0 - \$9,997 on auto and property insurance

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G13b]

Name: G13b

SAS Name: G13b

Universe Description:

Question Text: 15 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent for other types of non-health insurance, such as life insurance or long-term care insurance? Include any non-health insurance premiums paid through payroll deductions.

Respondent's best estimate is fine.

Skip Instructions: <0-1000> [goto G13b_FR]
<1001-9997> [goto ERR1_G13b]
<D,R> [goto G17a]

Name: G13b_FR

SAS Name: G13b_FR

Universe Description: Spent \$0 - \$9,997 on other non-health insurance (e.g., life, long-term care)

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|--------|-----------|--------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G17a]

Name: G17a

SAS Name: G17a

Universe Description: Subsample A cases in a 2nd or 3rd month interview

Question Text: 20 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent for subscriptions, books, and magazines?

Skip Instructions: <0-500> [goto G17a_FR]
 <501-9997> [goto ERR1_G17a]
 <D,R> [goto G17b]

Name: G17a_FR

SAS Name: G17a_FR

Universe Description: Spent \$0 - \$9,997 on subscriptions, books, and magazines.

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|--------|-----------|--------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G17b]

Name: G17b

SAS Name: G17b

Universe Description:

Question Text: 20 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent for tickets or admissions for movies, concerts, or sporting events?

Skip Instructions: <0-500> [goto G17b_FR]
 <501-9997> [goto ERR1_G17b]

<D,R> [goto G17c]

Name: G17b_FR

SAS Name: G17b_FR

Universe Description: Spent \$0 - \$9,997 on movies, concerts, or sporting events

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G17c]

Name: G17c

SAS Name: G17c

Universe Description:

Question Text: 21 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent on membership fees for shopping or vacation clubs, fitness centers, or civic organizations?

Skip Instructions: <0-300> [goto G17c_FR]
<301-9997> [goto ERR1_G17c]
<D,R> [goto G17d]

Name: G17c_FR

SAS Name: G17c_FR

Universe Description: Spent \$0 - \$9,997 on membership fees

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G17d]

Name: G17d

SAS Name: G17d

Universe Description:

Question Text: 21 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent on DVD rentals or purchases, photographic film or processing, and purchases of audio CDs, records, or tapes?

Skip Instructions: <0-350> [goto G17d_FR]
<351-9997> [goto ERR1_G17d]

<D,R> [IF SQ_SPLIT = 2 goto GLBL_TRANS]
[ELSEIF (INTNMBR = 2 and SQ_SPLIT = 1) goto
BlkWebCatiBack.THANKYOU]
[ELSE goto BlIncome {INTNMBR = 3 and SQ_SPLIT = 1}]

Name: G17d_FR

SAS Name: G17d_FR

Universe Description: Spent \$0 - \$9,997 on DVDs, CDs, and film

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|--------|-----------|--------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [IF SQ_SPLIT = 2 goto GLBL_TRANS]
[ELSEIF (INTNMBR = 2 and SQ_SPLIT = 1) goto
BlkWebCatiBack.THANKYOU]
[ELSE goto BlIncome {INTNMBR = 3 and SQ_SPLIT = 1}]

Name: G6a

SAS Name: G6a

Universe Description: Subsample B cases in a 2nd or 3rd month interview

Question Text: 4 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent on purchases of major household appliances, such as refrigerators, ovens, and clothes dryers?

Skip Instructions: <0-6000> [goto G6a_FR]

<6001-9997> [goto ERR1_G6a]
<D,R> [goto G6b]

Name: G6a_FR
SAS Name: G6a_FR
Universe Description: Spent \$0 - \$9,997 on major HH appliances

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|--------|-----------|--------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G6b]

Name: G6b
SAS Name: G6b
Universe Description:

Question Text: 5 - 8 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent on small household appliances and electronics, including computers, telephones, and audio or video equipment?

Skip Instructions: <0-9997> [goto G6b_FR]
<D,R> [goto G6c]

Name: G6b_FR
SAS Name: G6b_FR
Universe Description: Spent \$0 - \$9,997 on small household appliances and electronics

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|--------|-----------|--------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G6c]

Name: G6c

SAS Name: G6c

Universe Description:

Question Text: 9 - 10 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent on equipment for sports, recreation, and exercising?

Skip Instructions: <0-9997> [goto G6c_FR]
<D,R> [goto G14a]

Name: G6c_FR

SAS Name: G6c_FR

Universe Description: Spent \$0 - \$9,997 on sports and exercise equipment

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|--------|-----------|--------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G14a]

Name: G14a

SAS Name: G14a

Universe Description: Subsample B cases in a 2nd or 3rd month interview

Question Text: 16 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent for health insurance, including health insurance premiums paid through payroll deductions?

Skip Instructions: <0-2000> [goto G14a_FR]
<2001-9997> [goto ERR1_G14a]
<D,R> [goto G16a]

Name: G14a_FR

SAS Name: G14a_FR

Universe Description: Spent \$0 - \$9,997 on health insurance

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G16a]

Name: G16a

SAS Name: G16a

Universe Description: Subsample B cases in a 2nd or 3rd month interview

Question Text: 17 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent for tuition, other educational expenses, or payments for recreational lessons?

Skip Instructions: <0-6000> [goto G16a_FR]
<6001-9997> [goto ERR1_G16a]
<D,R> [goto G18]

Name: G16a_FR

SAS Name: G16a_FR

Universe Description: Spent \$0 - \$9,997 on tuition, other educational expenses, or payments for recreational lessons

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G18]

Name: G18

SAS Name: G18

Universe Description:

Question Text: 22 ? [F1]

Have ^YOU_ANYMEM completed any trips or vacations since the first of ^REF_MONTH?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1> [goto G18a]
<2,D,R> [goto G20]

Name: G18a

SAS Name: G18a

Universe Description: Yes, the household has completed trips

Question Text: 22 - 23 ? [F1]

Since the first of ^REF_MONTH, how much have ^YOU_ANYMEM spent for trips or vacations?

Please do not include trips made entirely for business or which will be entirely reimbursed by someone else.

Skip Instructions: <0-7000> [goto G18a_FR]
<7001-9997> [goto ERR1_G18a]
<D,R> [goto G20]

Name: G18a_FR

SAS Name: G18a_FR

Universe Description: Spent \$0 - \$9,997 on trips or vacations

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [goto G20]

Name: G20

SAS Name: G20

Universe Description: Subsample B cases in a 2nd or 3rd month interview

Question Text: Since the first of ^REF_MONTH, what has been ^YR_YRHHs WEEKLY expense for grocery shopping?

Respondent's best estimate is fine.

Skip Instructions: <0-400> [goto G20_FR]
<401-9997> [goto ERR1_G20]
<D,R> [IF SQ_SPLIT = 2 goto GLBL_TRANS]
[ELSEIF (INTNMBR = 2 and SQ_SPLIT = 1) goto
BikWebCatiBack.THANKYOU]
[ELSE goto BIncome {INTNMBR = 3 and SQ_SPLIT = 1}]

Name: G20_FR

SAS Name: G20_FR

Universe Description: Spent \$0 - \$9,997 on grocery shopping

Question Text: Did respondent give any indication that the answer provided was only a rough estimate or a guess?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1,2> [IF SQ_SPLIT = 2 goto GLBL_TRANS]
[ELSEIF (INTNMBR = 2 and SQ_SPLIT = 1) goto
BikWebCatiBack.THANKYOU]
[ELSE goto BIncome {INTNMBR = 3 and SQ_SPLIT = 1}]

Name: GLBL_TRANS

SAS Name: GLBL_TRANS

Universe Description: Global questions are asked first and the last global question has been asked

Question Text: The remaining questions in this survey are going to ask you about expenses
^YOU_ANYMEM
had for individual purchases, not the total amounts spent in different categories.

Answer List: Tcontinue

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Continue | Enter 1 to Continue |

Skip Instructions: <1> [IF QTYPE = 2 goto Bsect6]
[ELSE goto Bsect9]

Name: G9a_

SAS Name: G9a_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G12a_

SAS Name: G12a_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G13a_

SAS Name: G13a_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G13b_

SAS Name: G13b_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G17a_

SAS Name: G17a_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G17b_

SAS Name: G17b_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G17c_

SAS Name: G17c_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G17d_

SAS Name: G17d_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G6a_

SAS Name: G6a_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G14a_

SAS Name: G14a_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G16a_

SAS Name: G16a_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G18a_

SAS Name: G18a_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: G20_

SAS Name: G20_

Universe Description:

Question Text: ** Output variable **

Skip Instructions: <blank, A1>

Name: REC_STAT

SAS Name: REC_STAT

Universe Description:

Question Text: ** CREATED IN POST_PROCESSING **

Name: REC_ORIG

SAS Name: REC_ORIG

Universe Description:

Question Text: ** CREATED IN POST-PROCESSING **

Skip Instructions: <1>

Appendix VIII

MIS Survey Instrument Specifications – Post-Survey Assessment Questions

COVER SHEET

BPSAQ

Blaise Summary-Level Specifications Report

Name: PSAQ_INTRO

SAS Name: PSAQ_INTRO

Universe Description:

Question Text: I would like to change topics and ask you a few questions about your experience completing this interview. Your responses will help us improve the survey.

Answer List: Tcontinue

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Continue | Enter 1 to Continue |

Skip Instructions: <1> [goto PSAQ_1]

Name: PSAQ_1

SAS Name: PSAQ_1

Universe Description:

Question Text: How interesting was this survey to you? Would you say it was ^PSAQ1_FILL?

Answer List: TPSAQ_1

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Option_1 | ^very |
| 2 | Option_2 | ^somewhat |
| 3 | Option_3 | ^notVery |
| 4 | Option_4 | ^NotAtAll |

Skip Instructions: <1-4> [goto PSAQ_1_SP]
<DK, RF> [goto PSAQ_2A]

Name: PSAQ_1_SP

SAS Name: PSAQ_1_SP

Universe Description:

Question Text: Enter the respondent's verbatim response if he/she does more than simply repeat one of the given response options.

Enter 99 if respondent only repeats response options and provides no additional comments.

Skip Instructions: [goto PSAQ_2A]

Name: PSAQ_2A

SAS Name: PSAQ_2A

Universe Description:

Question Text: How difficult or easy was it for you to answer the questions in this survey - would you say it was ^PSAQ2A_FILL?

Answer List: TPSAQ_2A

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Option_1 | ^EASY |
| 2 | Option_2 | ^someEasy |
| 3 | Option_3 | ^someDifficult |
| 4 | Option_4 | ^veryDifficult |

Skip Instructions: <1-4> [goto PSAQ_2B]
<D,R> [goto PSAQ_3]

Name: PSAQ_2b

SAS Name: PSAQ_2b

Universe Description:

Question Text: Can you tell me a little more about that?
What made it ^Fill_PSAQ_2a?

Skip Instructions: [goto PSAQ_3]

Name: PSAQ_3

SAS Name: PSAQ_3

Universe Description:

Question Text: How burdensome was this survey to you? Would you say it was ^PSAQ3_FILL?

Answer List: TPSAQ_3

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Very | ^VeryBurden |
| 2 | Somewhat | ^someBurden |

| | | |
|---|----------|-----------------|
| 3 | NotVery | ^notVeryBurden |
| 4 | NotAtAll | ^NotAtAllBurden |

Skip Instructions: <1-4> [goto PSAQ_3_SP]
<DK, RF> [goto PSAQ_3A]

Name: PSAQ_3_SP

SAS Name: PSAQ_3_SP

Universe Description:

Question Text: Enter the respondent's verbatim response if he/she does more than simply repeat one of the given response options.

Enter 99 if respondent only repeats response options and provides no additional comments.

Skip Instructions: [goto PSAQ_3A]

Name: PSAQ_3A

SAS Name: PSAQ_3A

Universe Description:

Question Text: Over the course of the survey, you were asked to participate in ^Fill_PSAQ3A interviews.

Would you say that this was too many interviews, or did it seem like a reasonable number?

Answer List: TPSAQ_3A

| Value: | Mnemonic: | Description: |
|--------|------------|---------------------|
| 1 | TooMany | Too many interviews |
| 2 | Reasonable | A reasonable number |

Skip Instructions: <1, 2> [goto PSAQ_3A_SP]
<DK, RF> [goto PSAQ_3B]

Name: PSAQ_3A_SP

SAS Name: PSAQ_3A_SP

Universe Description:

Question Text: Enter the respondent's verbatim response, if he/she does more than simply repeat one of the given response options.

Enter 99 if respondent only repeats response options and provides

no additional comments.

Skip Instructions: [goto PSAQ_3B]

Name: PSAQ_3B

SAS Name: PSAQ_3B

Universe Description:

Question Text: Thinking about the number of phone calls you received before each interview, would you say that it was too many, or did it seem like a reasonable number?

Answer List: TPSAQ_3B

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | TooManyCalls | Too many calls |
| 2 | Reasonable | A reasonable number |

Skip Instructions: <1, 2> [goto PSAQ_3B_SP]
<DK, RF> [goto PSAQ_4]

Name: PSAQ_3B_SP

SAS Name: PSAQ_3B_SP

Universe Description:

Question Text: Enter the respondent's verbatim response if he/she does more than simply repeat one of the given response options.

Enter 99 if respondent only repeats response options and provides no additional comments.

Skip Instructions: [goto PSAQ_4]

Name: PSAQ_4

SAS Name: PSAQ_4

Universe Description:

Question Text: Do you feel that the length of today's interview was too long, too short, or about right?

Answer List: TPSAQ_4

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | TooLong | Too long |
| 2 | TooShort | Too short |

Skip Instructions: <1-3> [goto PSAQ_4_SP]
<DK, RF> [goto PSAQ_5]

Name: PSAQ_4_SP
SAS Name: PSAQ_4_SP
Universe Description: ALL

Question Text: Enter the respondent's verbatim response if he/she does more than simply repeat one of the given response options.

Enter 99 if respondent only repeats response options and provides no additional comments.

Skip Instructions: [goto PSAQ_5]

Name: PSAQ_5
SAS Name: PSAQ_5
Universe Description: ALL

Question Text: How long do you think today's interview took?

Enter response in minutes

Skip Instructions: <1-90, DK, or RF> [goto PSAQ_7]

Name: PSAQ_7
SAS Name: PSAQ_7
Universe Description:

Question Text: During today's interview, did you look at bills, receipts, bank or credit card statements, or other types of records when reporting any of your household expenses?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1> [goto PSAQ_7B]
<2,D,R> [goto PSAQ_8]

Name: PSAQ_7B

SAS Name: PSAQ_7B

Universe Description: Yes, looked at bills, receipts, etc

Question Text: What types of record(s) did you look at most often?

If necessary, ready answer list categories

Enter all that apply, separate with commas.

Answer List: PSAQ_7B

| Value: | Mnemonic: | Description: |
|---------------|------------------|------------------------|
| 1 | Bills | Bills |
| 2 | Receipts | Receipts |
| 3 | Bank | Bank Statements |
| 4 | Credit_Card | Credit Card Statements |
| 5 | Other | Other (specify) |

Skip Instructions: <1-4, DK, RF> [goto PSAQ_8]
<5> [goto PSAQ_7B_SP]

Name: PSAQ_7B_SP

SAS Name: PSAQ_7B_SP

Universe Description:

Question Text: Specify:

Skip Instructions: <30 characters> [goto PSAQ_8]

Name: PSAQ_8

SAS Name: PSAQ_8

Universe Description:

Question Text: Do you do any financial activity online, such as checking your account balances or paying your bills?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: [goto PSAQ_9]

Name: PSAQ_9

SAS Name: PSAQ_9

Universe Description:

Question Text: Do you use any personal finance software, such as Quicken, Microsoft Money, or Mint.com?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1> [goto PSAQ_10]

<2,D,R> [IF PSAQ_8 = 1 goto PSAQ_10]
[ELSEIF BControlCard.IB_R = 1 goto PSAQ_11]
[ELSE goto WebCATIBack Block]

Name: PSAQ_10

SAS Name: PSAQ_10

Universe Description:

Question Text: One idea that has been proposed to shorten the length of the interview is to give people the option of sharing their electronic expenditure records with us ahead of time. This would allow us to ask fewer questions over the phone.

Imagine that you are contacted to participate in a similar survey in the future. If you could be assured that your data would remain secure and confidential, how willing would you be to share information from your online accounts or other personal finance software for the purposes of this survey?

Would you say that you would be very willing to share that information, somewhat willing, not very willing, or not at all willing?

Answer List: TPSAQ 10

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Very | Very Willing |
| 2 | Somewhat | Somewhat Willing |
| 3 | NotVery | Not Very Willing |
| 4 | NotAtAll | Not At All Willing |

Skip Instructions: [goto PSAQ_11]

Name: PSAQ_11

SAS Name: PSAQ_11

Universe Description:

Question Text: Thinking back to the Information Booklet that you received, did you use the Booklet to prepare in advance for today's interview - for example, to help you gather information about your household expenses?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1> [goto PSAQ_11A]
<2> [goto PSAQ_11D]
<D,R> [goto BlkWebCATIBack Block]

Name: PSAQ_11a

SAS Name: PSAQ_11a

Universe Description:

Question Text: Did you refer to the Information Booklet during today's interview?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
| 1 | Yes | Yes |
| 2 | No | No |

Skip Instructions: <1> [goto PSAQ_11B]
<2, RF> [goto PSAQ_11D]

Name: PSAQ_11B

SAS Name: PSAQ_11B

Universe Description:

Question Text: During today's interview, did you follow along in the booklet as the questions were read, or did you look at examples only when the meaning of the question was unclear?

Answer List: TPSAQ_11B

| Value: | Mnemonic: | Description: |
|---------------|------------------|---|
| 1 | Follow | To consistently follow along as the questions were read |

| | | |
|---|---------|--|
| 2 | Example | To only look at examples when the meaning of the question was unclear. |
|---|---------|--|

Skip Instructions: [goto PSAQ_11C]

Name: PSAQ_11C

SAS Name: PSAQ_11C

Universe Description:

Question Text: Did you use the Information Booklet in any other way to help you answer the questions I asked today?

Answer List: Tyesno

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
|---------------|------------------|---------------------|

| | | |
|---|-----|-----|
| 1 | Yes | Yes |
|---|-----|-----|

| | | |
|---|----|----|
| 2 | No | No |
|---|----|----|

Skip Instructions: <1> [goto PSAQ_11C_SP]
<2,R> [Exit block and goto the WebCATIBack Block]

Name: PSAQ_11C_SP

SAS Name: PSAQ_11C_SP

Universe Description:

Question Text: In what way did you use it?

Enter the respondent's verbatim response if he/she does more than simply repeat one of the given response options.

Enter 99 if respondent only repeats response options and provides no additional comments.

Skip Instructions: [Exit block and goto the WebCATIBack Block]

Name: PSAQ_11D

SAS Name: PSAQ_11D

Universe Description:

Question Text: Was there a particular reason you didn't use the Information Booklet today?

Answer List: TPSAQ_11D

| Value: | Mnemonic: | Description: |
|---------------|------------------|---------------------|
|---------------|------------------|---------------------|

| | | |
|---|------|------------------------------------|
| 1 | Lost | Lost or forgot Information Booklet |
|---|------|------------------------------------|

| | | |
|---|-----------|--|
| 2 | HardToUse | Information Booklet poorly designed or hard to use |
| 3 | Already | Already familiar with each expenditure category |
| 4 | Trouble | Too much trouble |
| 5 | NoReason | No reason |
| 6 | Other | Other - specify |

Skip Instructions: <1-5,D,R> [Exit block and goto the BlkWebCATIBack Block]
<6> [goto PSAQ_11D_SP]

Name: PSAQ_11D_SP

SAS Name: PSAQ_11D_SP

Universe Description:

Question Text: Enter the respondent's verbatim response if he/she does more than simply repeat one of the given response options.

Enter 99 if respondent only repeats response options and provides no additional comments.

Skip Instructions: [Exit block and goto the WebCATIBack Block]

Name: REC_STAT

SAS Name: REC_STAT

Universe Description:

Question Text: ** CREATED IN POST_PROCESSING **

Name: REC_ORIG

SAS Name: REC_ORIG

Universe Description:

Question Text: ** CREATED IN POST-PROCESSING **

Skip Instructions: <1>

APPENDIX IX

MIS Source Variables

Mapping the Variables for the Split Questionnaire (SQ)-Control Group (C) Comparisons

In the table below, we document the mapping between source variables and analysis variables used for SQ – C group comparisons: column 1 and column 2 list the source variables from detailed questions and global questions by section, respectively; column 3 identifies the SQ group to which the global question for the section was assigned; and column 4 shows the analysis variables created that were common to the Cand SQ groups. The new variables in column 4 were sourced from a variable in the detailed question by default, unless it was an SQ group that had been assigned to global questions for that section. For example: for section 6, the SQ-A group had TXG6 sourced from TX6 (the sum of detailed questions TX6A and TX6B), while the SQ-B group had TXG6 sourced from G6_NET (the sum of four global questions G6A, G6B, 66C, G6D). For the C group, all the new variables in column 4 were sourced from detailed questions.

Table SC1. Expenditure variables for C-SQ group comparison

| Expenditure Section | Source variables (variables in the monthly MI summary files) | | SQ group | New analysis variables created for C-SQ-A, C-SQ-B group comparison in Waves 1, 2, & 3 | New analysis variables created for C-SQ (combined) group comparison in Waves 1, 2, & 3 |
|---------------------|---|---|----------|---|---|
| | From detailed questions (default source) | From global questions | | | |
| 6 | TX6= TX6A + TX6B | G6_NET=(G6A + G6BC), where G6BC=(G6B+G6C) | B | TXG6 =TXG6A + TXG6BC, where TXG6A=G6A, TXG6BC=G6BC | TXG6 |
| 9 | TX9A | G9A | A | TXG9A | TXG9A |
| 12 | TX12A_NET | G12A | A | TXG12A | TXG12A |
| 13 | TX13B_NET | G13_NET=(G13A + G13B) | A | TXG13B | TXG13B |
| 14 | TX14B_NET | G14A | B | TXG14B | TXG14B |
| 16 | TX16A_NET | G16A | B | TXG16A | TXG16A |
| 17 | TX17_NET==TX17A_NET + TX17B_NET | G17_NET=[G17A + G17BCD, where G17BCD=(G17B+G17C+G17D) | A | TXG17=TXG17A + TXG17B Where TXG17A=G17A, TXG17B=G17BCD | TXG17 |
| 18 | TX18=TX18A+TX18B | G18A | B | TXG18 | TXG18 |
| 20 | TX20A | n/a | A | TXG20_SQA | GROCWEKX |
| 20 | GROCWEKX (section 20A) | G20 | B | TXG20_SQB | |

| Table SC2. Analysis variables of totals for group comparisons | | |
|--|---|---|
| | CON-SQA | CON-SQB |
| Total expenditures (including globals) | TOTXG_AC=SUM(TXG6, TXG9A, TXG12A, TXG13B, TXG14B, TXG16A, TXG17, TXG18, TXG20_SQA); TOTXG_BC=SUM(TXG6, TXG9A, TXG12A, TXG13B, TXG14B, TXG16A, TXG17, TXG18, TXG20_SQB); For CON_SQ: TOTXG_QC=SUM(TXG6, TXG9A, TXG12A, TXG13B, TXG14B, TXG16A, TXG17, TXG18, GROCWEKX); | |
| The following comparisons only involve non-global questions | | |
| Number of valid reports (excluding 0) | TOTNR_ACng =sum(NR6, NR14B, NR16A, NR18) | TOTNR_BCng =sum(NR9A, NR12A, NR13B, NR17B, NR17A) |
| Number of don't know/refused reports | TOTDR_ACng =sum(DR6, DR14B, DR16A, DR18) | TOTDR_BCng =sum(DR9A, DR12A, DR13B, DR17B, DR17A) |
| Number of aggregated (combined) reports | TOTCC_ACng =sum(CC6A, CC6B, CC16A) | TOTCC_BCng =sum(CC9A, CC12A, CC13B) |
| For global questions of SQ group | | |
| Frequency of reporting of global questions | NR_G9A NR_G12A NR_G13A NR_G13B NR_G17A NR_G17B NR_G17C NR_G17D | NR_G6A NR_G6B NR_G6C NR_G14A NR_G16A NR_G18A NR_G20 |
| For Wave 1, all treatment groups | | |
| Total expenditures | TOTXG_W1=TOTXG_AC | |
| Number of reports | TOTNR_W1=SUM(NR6, NR9A, NR12A, NR13B, NR14B, NR16A, NR17A, NR17B, NR18); | |
| Number of don't know/refused | TOTDR_W1=SUM(DR6, DR9A, DR12A, DR13B, DR14B, DR16A, DR17A, DR17B, DR18); | |
| Number of aggregated (combined) reports | TOTCC_W1=SUM(CC6A, CC6B, CC9A, CC12A, CC13B, CC16A); | |

Mapping Variables for the Global vs. Detailed Question Comparisons

The mapping between source variables and analysis variables used for group comparisons are documented in Table SC1: column 2 identifies the treatment group that is asked global questions for the section specified in column 1; column 3 lists the analysis variables created; column 4 and column 5 list the source variables from detailed questions and global questions by section, respectively, used to create the analysis variables in column 3. For example: to create the analysis variable TXG6 for section 6, for a sample unit in CONTROL or SQA, the analysis variable TXG6 will be assigned the value of TX6, while for a sample unit in SQB, the analysis variable TXG6 will be assigned the value of G6_NET. For the comparison of reporting rate of “don’t know/refused” responses, a similar mapping between source variables and analysis variables used for group comparisons are documented in Table SC2.

| Table SC3. Definition of analysis variables for comparison of SQ_COMBINED with CONTROL groups on expenditures | | | | |
|---|------------------------------------|--|--|--------------------------------|
| Expenditure Section | Global section for treatment group | Expenditure variables used in analysis | Source variables for creating analysis variables (variables in the monthly MI summary files) | |
| | | | From detailed questions | From global questions |
| [1] | [2] | [3] | [4] | [5] |
| 6 | SQB | TXG6 | TX6= TX6A + TX6B | G6_NET= (G6A + G6B + G6C) |
| 9 | SQA | TXG9A | TX9A | G9A |
| 12 | SQA | TXG12A | TX12A_NET | G12A |
| 13 | SQA | TXG13B | TX13B_NET | G13_NET= G13A + G13B |
| 14 | SQB | TXG14B | TX14B_NET | G14A |
| 16 | SQB | TXG16A | TX16A_NET | G16A |
| 17 | SQA | TXG17 | TX17_NET==TX17A_NET + TX17B_NET | G17_NET= G17A + G17B+G17C+G17D |
| 18 | SQB | TXG18 | TX18=TX18A+TX18B | G18A |
| 20 | SQB | GROCWEKX | TX20A= SUM(GROCWEKX, OTHSTUFX, OSTORWKX, ALC_HX4, ALC_OX4, DINE_WKX); | n/a |
| | | | GROCWEKX (section 20A) | G20 |

Table SC4. Definition of analysis variables for comparison of SQ_COMBINED with CONTROL groups on “don’t know/refused” responses

| Expenditure Section | Expenditure variables | Source variables for creating analysis variables (variables in the monthly MI summary files) | |
|---------------------|-----------------------|---|---------------------------|
| | | From detailed questions | From global questions |
| 6 | DR_SC6 | DR6A, DR6B | DR_G6A, DR_G6B, DR_G6C |
| 9 | DR_SC9 | DR9A | DR_G9A |
| 12 | DR_SC12 | DR12A | DR_G12A |
| 13 | DR_SC13 | DR13B | DR_G13A, DR_G13B |
| 14 | DR_SC14 | DR14B | DR_G14A |
| 16 | DR_SC16 | DR16A | DR_G16A |
| 17 | DR_SC17 | DR17A, DR17B | DR_G17B, DR_G17C, DR_G17D |
| 18 | DR_SC18 | DR18A, DR18B | DR_G18A |
| 20 | DR_SC20A | DR20A | DR_G20 |

APPENDIX X
Section-Level Expenditure Estimates for All MIS Treatment Groups

SQ-A – C: Wave 2 Expenditure Comparisons

| Variable | SQ-A | Control | Difference (R-C) | 95LCI diff | 95UCI diff | SE | p-value for T-test |
|------------------------------------|--------|---------|------------------|------------|------------|--------|--------------------|
| Appliances (6) | 325.8 | 360.77 | -34.97 | -145.1 | 75.17 | 56.129 | 0.5296 |
| Maj. app. (6A) | 72.926 | 61.64 | 11.286 | -29.04 | 51.613 | 20.551 | 0.5867 |
| Small app. (6B) | 252.87 | 299.13 | -46.26 | -148.3 | 55.811 | 52.015 | 0.3681 |
| Clothing (9A) ¹ | 274.91 | 233.5 | 41.407 | -12.61 | 95.42 | 27.525 | 0.1378 |
| Vehicle operating expenses (12A) | 304.59 | 223.65 | 80.948 | 19.36 | 142.54 | 31.386 | 0.0112 |
| Non-health insurance (13B) | 839.1 | 488.34 | 350.76 | 232 | 469.51 | 60.518 | <.0001 |
| Health insurance (14B) | 158.37 | 117.89 | 40.48 | -9.67 | 90.62 | 25.55 | 0.1196 |
| Education (16A) | 890.82 | 672.62 | 218.2 | -221.3 | 657.74 | 223.99 | 0.3302 |
| Subscriptions & Entertainment (17) | 206.24 | 247.42 | -41.19 | -99.01 | 16.641 | 29.469 | 0.1584 |
| Subscriptions (17A) | 37.495 | 121.8 | -84.31 | -120 | -48.6 | 18.199 | <.0001 |
| Entertainment (17B) | 168.74 | 125.62 | 43.124 | 2.5823 | 83.666 | 20.66 | 0.0406 |
| Trips (18) | 130.75 | 215.62 | -84.88 | -238.2 | 68.458 | 78.14 | 0.2660 |
| Weekly Grocery (20A) | 187.42 | 192.37 | -4.951 | -22.5 | 12.603 | 8.9457 | 0.5819 |

SQ-A – C: Wave 3 Expenditure Comparisons

| Variable | SQ-A | Control | Difference (R-C) | 95LCI diff | 95UCI diff | SE | p-value for T-test |
|------------------------------------|--------|---------|------------------|------------|------------|--------|--------------------|
| Appliances (6) | 346.61 | 375.37 | -28.76 | -127.9 | 70.331 | 50.495 | 0.5687 |
| Maj. app. (6A) | 83.274 | 63.874 | 19.4 | -23.44 | 62.239 | 21.829 | 0.3746 |
| Small app. (6B) | 263.33 | 311.5 | -48.16 | -133.4 | 37.032 | 43.412 | 0.267 |
| Clothing (9A) | 357.82 | 249.24 | 108.58 | 8.4501 | 208.7 | 51.021 | 0.034 |
| Vehicle operating expenses (12A) | 319.12 | 177.27 | 141.85 | 59.565 | 224.13 | 41.927 | 0.0008 |
| Non-health insurance (13B) | 831.96 | 520.79 | 311.16 | 183.56 | 438.77 | 65.023 | <.0001 |
| Health insurance (14B) | 162.48 | 140.4 | 22.076 | -33.06 | 77.208 | 28.093 | 0.4326 |
| Education (16A) | 491.24 | 559.76 | -68.52 | -428.3 | 291.28 | 183.34 | 0.7084 |
| Subscriptions & Entertainment (17) | 233.37 | 241 | -7.631 | -72.97 | 57.704 | 33.292 | 0.819 |
| Subscriptions (17A) | 45.31 | 122.57 | -77.26 | -107.3 | -47.18 | 15.324 | <.0001 |
| Entertainment (17B) | 188.06 | 118.43 | 69.627 | 13.623 | 125.63 | 28.538 | 0.0152 |
| Trips (18) | 35.766 | 65.017 | -29.25 | -75.04 | 16.541 | 23.334 | 0.2096 |
| Weekly Grocery (20A) | 189.8 | 187.81 | 1.9883 | -14.67 | 18.65 | 8.4905 | 0.8149 |

¹ For the SQ-A and SQ-B tables, the yellow-highlighted rows indicate use of global questions.

SQ-B – C: Wave 2 Expenditure Comparisons

| Variable | SQ-B | Control | Difference (R-C) | 95LCI diff | 95UCI diff | SE | p-value for T-test |
|------------------------------------|--------|---------|------------------|------------|------------|--------|--------------------|
| Appliances (6) | 271.32 | 360.77 | -89.46 | -186.2 | 7.3055 | 49.315 | 0.0748 |
| Maj. app. (6A) | 65.153 | 61.64 | 3.5137 | -35.44 | 42.464 | 19.851 | 0.8582 |
| Small app. (6B) | 206.16 | 299.13 | -92.97 | -181.3 | -4.669 | 45.003 | 0.0435 |
| Clothing (9A) | 253.17 | 233.5 | 19.671 | -24.64 | 63.987 | 22.586 | 0.384 |
| Vehicle operating expenses (12A) | 269.86 | 223.65 | 46.217 | -7.241 | 99.675 | 27.245 | 0.0881 |
| Non-health insurance (13B) | 544.65 | 488.34 | 56.309 | -36.98 | 149.6 | 47.545 | 0.2344 |
| Health insurance (14B) | 613.35 | 117.89 | 495.46 | 422.02 | 568.9 | 37.43 | <.0001 |
| Education (16A) | 839.46 | 672.62 | 166.84 | -289.7 | 623.33 | 232.66 | 0.4704 |
| Subscriptions & Entertainment (17) | 264.86 | 247.42 | 17.436 | -41.25 | 76.125 | 29.911 | 0.5601 |
| Subscriptions (17A) | 127.64 | 121.8 | 5.8313 | -39.99 | 51.657 | 23.355 | 0.8029 |
| Entertainment (17B) | 137.22 | 125.62 | 11.605 | -16.79 | 40.002 | 14.473 | 0.4257 |
| Trips (18) | 778.98 | 215.62 | 563.36 | 371.95 | 754.77 | 97.555 | <.0001 |
| Weekly Grocery (20A) | 119.97 | 114.42 | 5.5489 | -4.212 | 15.31 | 4.9749 | 0.2649 |

SQ-B – C: Wave 3 Expenditure Comparisons

| Variable | SQ-B | Control | Difference (R-C) | 95LCI diff | 95UCI diff | SE | p-value for T-test |
|------------------------------------|--------|---------|------------------|------------|------------|--------|--------------------|
| Appliances (6) | 368.79 | 375.37 | -6.581 | -113.7 | 100.5 | 54.571 | 0.9040 |
| Maj. app. (6A) | 104 | 63.874 | 40.122 | -11.29 | 91.534 | 26.2 | 0.1126 |
| Small app. (6B) | 264.79 | 311.5 | -46.7 | -133.3 | 39.924 | 44.147 | 0.2986 |
| Clothing (9A) | 304.14 | 249.24 | 54.9 | -3.071 | 112.87 | 29.543 | 0.0634 |
| Vehicle operating expenses (12A) | 263.19 | 177.27 | 85.912 | 28.667 | 143.16 | 29.173 | 0.0026 |
| Non-health insurance (13B) | 596.36 | 520.79 | 75.568 | -24.51 | 175.65 | 51.001 | 0.1323 |
| Health insurance (14B) | 725.67 | 140.4 | 585.27 | 447.39 | 723.15 | 70.267 | <.0001 |
| Education (16A) | 966.5 | 559.76 | 406.75 | -139.1 | 952.62 | 278.19 | 0.1289 |
| Subscriptions & Entertainment (17) | 265.88 | 241 | 24.877 | -31.26 | 81.01 | 28.606 | 0.3754 |
| Subscriptions (17A) | 125.25 | 122.57 | 2.6845 | -34.66 | 40.031 | 19.032 | 0.8887 |
| Entertainment (17B) | 140.62 | 118.43 | 22.193 | -15.49 | 59.876 | 19.204 | 0.2263 |
| Trips (18) | 430.21 | 65.017 | 365.19 | 256.32 | 474.06 | 55.481 | <.0001 |
| Weekly Grocery (20A) | 124.68 | 114.01 | 10.675 | -0.629 | 21.979 | 5.7608 | 0.0642 |

RP – C: Expenditure Comparisons (RP Data Aggregated from Waves 2 – 4)

| Variable | RP | Control | Difference (R-C) | 95LCI diff | 95UCI diff | SE | p-value for T-test |
|------------------------------------|--------|---------|------------------|------------|------------|--------|--------------------|
| Appliances (6) | 477.27 | 360.77 | 116.50 | -19.67 | 252.67 | 69.39 | 0.1003 |
| Maj. app. (6A) | 120.58 | 61.64 | 58.94 | 11.73 | 106.14 | 24.06 | 0.0199 |
| Small app. (6B) | 356.69 | 299.13 | 57.56 | -67.52 | 182.64 | 63.74 | 0.3725 |
| Clothing (9A) | 231.00 | 233.50 | -2.50 | -46.25 | 41.26 | 22.30 | 0.9082 |
| Vehicle operating expenses (12A) | 437.74 | 223.65 | 214.09 | 59.98 | 368.21 | 78.53 | 0.0147 |
| Non-health insurance (13B) | 486.71 | 488.34 | -1.63 | -103.00 | 99.71 | 51.64 | 0.9752 |
| Health insurance (14B) | 386.22 | 117.89 | 268.33 | 182.85 | 353.81 | 43.56 | <.0001 |
| Education (16A) | 679.87 | 672.62 | 7.25 | -396.90 | 411.37 | 205.92 | 0.9709 |
| Subscriptions & Entertainment (17) | 192.86 | 247.42 | -54.56 | -109.80 | 0.71 | 28.17 | 0.0394 |
| Subscriptions (17A) | 88.37 | 121.80 | -33.43 | -74.65 | 7.79 | 21.00 | 0.0868 |
| Entertainment (17B) | 104.49 | 125.62 | -21.13 | -50.62 | 8.36 | 15.03 | 0.1446 |
| Trips (18) | 41.61 | 215.62 | -174.00 | -324.80 | -23.18 | 76.86 | 0.0123 |
| Weekly Grocery (20A) | 11.38 | 5.62 | 5.76 | -4.84 | 16.35 | 5.40 | 0.2947 |
| | 30.24 | 210.00 | -179.80 | -330.30 | -29.24 | 76.70 | 0.0095 |
| | 173.83 | 192.37 | -18.54 | -34.07 | -3.02 | 7.91 | <.0154 |

Appendix XI

Respondents' Self-Reported Use of Financial Applications and Willingness to Share Electronic Records

| | Control | Recall | SQA | SQB | Chi-sq p-value 4 groups | Chi-sq p-value CON, SQA, SQB | Chi-sq p-value CON, RP |
|---|---------|--------|------|------|----------------------------|---------------------------------------|---------------------------------|
| Any online financial activity (e.g., online banking)? | | | | | 0.3532 | 0.8182 | 0.5194 |
| . | 0.2 | 0.2 | 0.4 | 0.5 | | | |
| Yes | 51.8 | 48.3 | 54.6 | 53.2 | | | |
| No | 48.0 | 51.5 | 44.9 | 46.3 | | | |
| Use personal finance software? | | | | | 0.1742 | 0.1987 | 0.9284 |
| . | 0.2 | 0.3 | 0.6 | 1.4 | | | |
| Yes | 10.9 | 10.9 | 12.7 | 12.5 | | | |
| No | 88.9 | 88.8 | 86.7 | 86.1 | | | |
| Willingness to provide electronic expenditure records? | | | | | 0.4241 | 0.7413 | 0.3255 |
| . | 47.0 | 50.2 | 43.9 | 45.4 | | | |
| Very willing | 3.4 | 4.0 | 3.2 | 3.0 | | | |
| Somewhat willing | 10.3 | 6.9 | 7.4 | 9.1 | | | |
| Not very willing | 9.2 | 9.5 | 9.9 | 9.3 | | | |
| Not at all willing | 30.2 | 29.4 | 35.7 | 33.3 | | | |

Appendix XII

Measurement Issues Study Weighting Documentation



Draft – July 19, 2011

MEMORANDUM FOR Scott Fricker
Research Psychologist
Office of Survey Methods Research
Bureau of Labor Statistics

Through: Cheryl R. Landman
Chief, Demographic Surveys Division
U. S. Census Bureau

From: Ruth Ann Killion
Chief, Demographic Statistical Methods Division
U. S. Census Bureau

Prepared by: Franklin Silberstein
Victimization and Expenditures Branch
Demographic Statistical Methods Division

Subject: Base Weights and Non-Interview Adjustments for the 2010
Consumer Expenditures Quarterly, Measurement Issues
Study

I. Introduction

This memorandum documents the base-weight calculation for the Consumer Expenditures Quarterly – Measurement Issues Study (CEQ-MIS). The Demographic Statistical Methods Division (DSMD) will deliver the base weights to the Bureau of Labor Statistics on a unit control file (UCF) specified in Census (2010.) This file, like the CEQ-MIS sample, only contains unit-frame cases.

The final weight is calculated as $FW = BW_{CEQ-MIS} \cdot NAF$.

The Bureau of Labor Statistics is conducting the CEQ-MIS, as part of an ongoing effort to examine alternative data collection strategies for the CEQ Survey that would improve data quality, maintain or increase response rates, and reduce data collection costs. This study focuses on the effects of a shorter questionnaire on respondent burden, data quality, and nonresponse error. In addition, a separate condition in this study assesses the effects of a one-month versus three-month reference period on underreporting due to recall errors. The results of this study are being used to inform future CEQ research activities and decisions about how to redesign the production survey.

II. Calculating the Base Weights

Base-weight (BW) formula for the CEQ-MIS:

$$BW_{CEQ-MIS} = BW_{CE} \times F1 \times F2 \times F3$$

Following are definitions of the base-weight-formula components:

- BW_{CE} : Original Base Weights

The original base weights for the Consumer Expenditures Surveys differ by the stratification primary sampling unit.

- $F1$: Universe Creation Using Reduction Groups

This factor differs by sample designation because the number of reduction groups we used varied by sample designation. To create a universe of cases that we would use later for sample selection, we selected specific reduction groups, within each sample designation. See Census(2010c). (Each sample designation is evenly divided into 101 reduction groups. For a more detailed description of how these codes were used in the CEQ-MIS sample, see Census (2010b).)

We selected enough reduction groups in each sample designation to yield the desired number of first interviews after inflating that number for the expected ineligibility, non-response, and non-telematch rates. Ineligibility rates reflected cases with non-working telephone numbers and numbers that belonged to a business instead of a residence. The non-telematch rate refers to cases where the address did not match to a telephone number. Below is the calculation for this base-weight factor:

$$F1 = \begin{cases} \frac{\text{reductiongroupsavailable}=101}{\text{reductiongroupsused}=96} & \text{if } X11 \\ \\ 101/101 & \text{if } X17 \\ \\ 101/83 & \text{if } X18 \\ \\ 101/74 & \text{if } X19 \\ \\ 101/17 & \text{if } X20 \end{cases}$$

- *F2: Telematch*

This factor differs depending on whether a record was in the first or second telematch operation. The first telematch operation resulted in a lower rate of matches than our assumption of 50%. As a result, we decided to obtain more cases by creating a supplemental sample universe and sending it through a second telematch operation. The set of cases we initially selected for telematch was referred to as the original universe and the second set of cases as the supplemental universe. Below is the factor depending on whether a record was selected for the first or second telematch operation:

$$F2 = \begin{cases} \frac{\text{recordsenttotelematch}=20,007}{\text{recordsmatched}=6,203} & \text{if 1st telematch (X17, X18, X19)} \\ \\ 14,226/6,131 & \text{if 2nd telematch (X11, X18, X19, X20)} \end{cases}$$

- *F3: Final Subsampling to Reduce Sample Size*

This factor differs by the CEQ-MIS condition (e.g., study treatment). The conditions were control-group (CG), split-questionnaire (SQ) and recall-

period (RP). We subsampled to remove extra cases resulting from the telematch operations. Since the different conditions and corresponding sample designations varied in sample size, we customized the subsampling to each condition. We used reduction groups within sample designations to define the subsampling rates. For the recall-period condition, the smallest sample, we did not have to subsample because the telematch process returned the number of cases we wanted to send out for interview. Below are the definitions of this factor based on subsampling rates and conditions.

$$F3 = \begin{cases} 2,411 = \text{sample received from telematch} / 2,344 = \text{sample sent for interview} & \text{if } CG \\ 6,688 = \text{sample received from telematch} / 4,734 = \text{sample sent for interview} & \text{if } SQ \\ 1/2 & \text{if } RP \end{cases}$$

III. Calculating the NAF: Nonresponse Adjustment Factor

Prior to developing nonresponse adjustments, it's important to identify the factors that determine nonresponse adjustments. Considering the overall areas needing nonresponse adjusting, we first break out sample into four basic groups: (1) the control group sample (CG), (2) the split questionnaire sample covered by sample designation X11(SQ1), (3) the split questionnaire sample covered by sample designation X17 (SQ2), and (4) the recall period sample group (RP). We then examine the data by panel. Each of the groups is interviewed in three equal panels. According to panel and within each of those four groups, we identified the following factors for nonresponse adjustment: (1) interview number, (2) renter/owner status, and (3) region.

Regarding CG and SQ, each household was scheduled for interview three times. For RP, each household was scheduled for interview four times. In determining renter/owner status, in general, about 84% of households were identified clearly as owner. Of the remainder, the rest were primarily identified as renter. A small percentage were identified as indeterminate and so we consolidated all those nonowners as renters. The regions were broken down into the four general BLS regions: (1) Northeast (NE), (2) Midwest (MW), (3) South (S), and (4) West (W).

For CG and SQ1 and SQ2, we start out with 24 possible cells. First, we cross interview number with owner renter status (yielding the combinations: Interview no. 1/ Owner, Interview No. 1/ Renter, Interview No. 2/ Owner, ... Interview No. 3/ Renter). We examine within each of those six cross-combinations, and there

exists cells representing each of the four possible regions. So prior to collapsing, we start out with the same cells in Table 1 for all three panels.

Table 1

| | Interview No. 1 | | | | Interview No. 2 | | | | Interview No. 3 | | | |
|--------|-----------------|----|---|---|-----------------|----|---|---|-----------------|----|---|---|
| Owner | NE | MW | S | W | NE | MW | S | W | NE | MW | S | W |
| Renter | NE | MW | S | W | NE | MW | S | W | NE | MW | S | W |

For CG, we only collapsed renters for all panels. We collapsed the cells of the four regions into one merged cell per each interview. As a result, we were left with the cells in Table 2 after collapsing.

Table 2

| | Interview No. 1 | | | | Interview No. 2 | | | | Interview No. 3 | | | |
|--------|-----------------|----|---|---|-----------------|----|---|---|-----------------|----|---|---|
| Owner | NE | MW | S | W | NE | MW | S | W | NE | MW | S | W |
| Renter | Merged cell 1 | | | | Merged cell 2 | | | | Merged cell 3 | | | |

Regarding SQ1, we saw no need to collapse since each cell had at least 29 eligible interviews. However, for SQ2, for each panel we collapsed all regions so that there was only one merged cell per each interview. The resulting collapsed cells would be identical to what we have in CG. Therefore, once again SQ2's collapsed cell combinations are noted in Table 2 for all panels.

For RP, the pre-collapsed format covers 32 possible cells consisting of eight basic cross-combinations broken down by owner/renter status, interview number, and region. See Table 3 for RP pre-collapsed data for all panels. All renter cells in Interview Number 4 collapse all regions into a single cell only during the fourth interview and during the second panel. The first, second and third interviews of panel two do not collapse. See Table 4 for results of that collapsing. Otherwise, cells remained the same as shown in Table 3.

Table 3

| | Interview No. 1 | | | | Interview No. 2 | | | | Interview No. 3 | | | | Interview No. 4 | | | |
|--------|-----------------|----|---|---|-----------------|----|---|---|-----------------|----|---|---|-----------------|----|---|---|
| Owner | NE | MW | S | W | NE | MW | S | W | NE | MW | S | W | NE | MW | S | W |
| Renter | NE | MW | S | W | NE | MW | S | W | NE | MW | S | W | NE | MW | S | W |

Table 4 (panel 2 only)

| | Interview No. 1 | | | | Interview No. 2 | | | | Interview No. 3 | | | | Interview No. 4 | | | |
|--------|-----------------|----|---|---|-----------------|----|---|---|-----------------|----|---|---|-------------------------|----|---|---|
| Owner | NE | MW | S | W | NE | MW | S | W | NE | MW | S | W | NE | MW | S | W |
| Renter | NE | MW | S | W | NE | MW | S | W | NE | MW | S | W | Renters for Interview 4 | | | |

Defining Outcome Status

Table 5 below lists the outcomes of the MIS and how the outcomes were grouped into the basic AAPOR outcome categories.

Table 5: Outcome Codes for the MIS

| Eligibility Status | Outcome of Interview | Outcome Code | Description of Outcome |
|--------------------|----------------------|--------------|---|
| Eligible | Completed Interviews | 001 | Fully completed |
| | | 002 | Sufficient Partial |
| | | 004 | Sufficient Partial set at Closeout ¹ |
| | Non-Interviews | 021 | Sample unit eligible but unavailable through Closeout |
| | | 024 | Unconverted language problem |
| | | 025 | Unconverted hearing barrier |
| | | 176 | Congressional refusal – delete case |
| | | 179 | Hostile break off |
| | | 181 | Refusal |
| | | 182 | Hard Refusal |
| | | 183 | Exceeded unproductive call max |
| | | 188 | Insufficient partial - callback |
| | | 193 | Privacy Detector |
| | | 194 | Never contacted - confirmed number |
| | | 195 | Never contacted – unconfirmed number |
| | | 199 | Never tried (new case) |
| Ineligible | Ineligible | 020 | Sample Unit ineligible – out of scope |

Some outcomes identified the sample units that were not sent for additional interviews. These cases only had one or two records instead of the full set of three interviews. So that we could calculate the non-interview adjustment factor, we copied the last outcome of a case and repeated it for all subsequent interviews. For example, if a sample unit was ineligible in the first interview and never attempted in the second or third interview, we made a copy of the ineligible

¹ Note: In Attachment A of the memorandum “Sample Design of the 2010 Consumer Expenditures Quarterly Measurement Issues Study (CEQ-MIS)”, July 23, 2010, a final code for 004 was omitted. However, a few times code 004 occurred in this survey and we therefore list it as “Sufficient Partial set at Closeout”, hence a completed interview.

record for the second and third interviews. By making a complete dataset for each qtype/panel/interview combination, we can calculate weights that represent the same total.

Non-Interview Adjustment Factor

We defined the non-interview adjustment factor as

$$NAF_c = \frac{A_c + B_c}{A_c}$$

where

| | |
|-------|--|
| c | Index on the weighting cells |
| A_c | Sum of the base weights for the Completed Interviews |
| B_c | Sum of the base weights for the Non-Interviews |

IV. Miscellaneous

This memorandum is stored in the directory “M:\ADC-LEDSP\VEB\CE\Final Memos\2010-07 CE Measurement Study Base Weights v4.0.docx”.

IV. References

U.S. Census Bureau (2010). Census Bureau Memorandum, “2010 Consumer Expenditure Quarterly Measurement Study - Specification for Creating the Unit Control File (UCF) for the Bureau of Labor Statistics (BLS),” from Scott to Lewis, dated June 2010.

U.S. Census Bureau (2010b). Census Bureau Memorandum, “Sample Design of the 2010 Consumer Expenditures Quarterly Measurement Issues Study (CEQ-MIS),” from Ash to Fricker, dated July 2010.

U.S. Census Bureau (2010c). Census Bureau Memorandum, “Sample Date Reassignment for the Consumer Expenditure Quarterly Measurement Issues Study,” from Flanagan to Montgomery.

cc:

P. Flanagan (DSMD)

S. Ash

M. Wilson

L. Flores

J. Adams

A. Spencer

D. Castello

B. Dumbacher

D. Swanson (BLS)