

# **ATTACHMENT A: GEMINI CE LARGE SCALE FEASIBILITY TEST OF ONLINE DIARIES**

## **I. STUDY OVERVIEW**

### **1. BACKGROUND**

In 2009, the Bureau of Labor Statistics (BLS) Consumer Expenditure Survey (CE) initiated the multi-year Gemini Project for the purpose of researching, developing, and implementing an improved survey design. The objective of the redesign is to improve the quality of the survey estimates through a verifiable reduction in measurement error.

The current Consumer Expenditure Diary Survey (CED) at the Bureau of Labor Statistics (BLS) uses a pencil-and-paper instrument (PAPI) to collect expenditure information from respondents. PAPI diary data collection has a number of inherent drawbacks, which limit the ability of respondents to contemporaneously enter expenditure data. The Gemini redesign involves the introduction of an online diary which can be accessed on a computer or mobile device which can provide easy access to the diary for data entry on the go. Previous versions of online diaries have been tested in the Individual Diary Feasibility Test and Proof of Concept Test.

BLS will be realizing the CE survey redesign through a phased implementation of key design elements. This involves implementing an online household-level diary into CED production in January 2022 (pending testing) and implementing a streamlined questionnaire and redesigned interviewing instrument into CEQ production April 2023.

### **2. OBJECTIVES**

In order to confidently implement an online diary, the CE program must first successfully complete a large-scale feasibility test of online diaries. This test will build on information learned from prior tests and additionally will have sufficient sample to make statistical inferences. The objective of this work is to develop an online diary and field a test from October 2019 to April 2020. The target number of completed interviews for the large-scale

diary test is 1,200, assuming a response rate similar to current production at the time of planning for the test. This will require a starting sample size of 2,500 needed cases to obtain the desired 1,200 complete cases.

*Objectives for Large-scale Online Diary Test:*

1. Develop an online diary instrument which can be accessed on a computer and mobile device. Integrate the online and paper diaries with a case management system to capture and integrate data. The case management system will have the additional capabilities of managing the sample and tracking response status.
2. Develop systems, field procedures and training materials for the LSF. The test will involve a small pilot test of 9 interviews, followed by a six-month field test with a starting sample of 2,500 cases. Interviewer debriefing after data collection will be conducted to get feedback on procedures, instruments, and materials. Respondent feedback on the online diary procedures, instrument, materials and overall experience will be obtained through a self-administered respondent debriefing.
3. Analyze the data and paradata from the feasibility test to make decisions on the implementation of online diaries in the CED in 2022. Research questions include:
  - Are there any operational issues with implementing online diaries?
  - How do token incentives impact response rates and number of contact attempts?
  - What technical issues do respondents have with using online diaries?
  - Does receiving the pre-notification postcard and advance letter affect response?
  - How useful are respondent materials like the user guide and online tutorials?
  - How effective are the diary placement and pickup procedures?
  - Are there differences in data quality between online and production (paper) diaries? Data quality can be measured as number of items reported, expenditure amounts, item nonresponse, and extent of rounding of expenditures.
  - Does using an online diary affect the incidence or amount of data entered at the diary pickup visit by recall?
  - Are there differences in week 1 and week 2 completion rates for online diaries compared to paper diaries?

- Are there differences in data entry patterns for online diaries compared to paper diaries?
- Are there demographic differences between online diary users and those who are online-eligible but chose to complete the paper diary?

## **II. STUDY DESIGN**

### **1. FIELDING PERIOD**

The large scale feasibility test of online diaries will be fielded from October 2019 to March 2020 (April 2020 closeout). The target number of completed interviews for the large-scale diary test is 1,200 with a response rate similar to current production at the time of the test. BLS estimates a starting sample size of 2,500 needed cases to obtain the desired 1,200 complete cases.

### **2. RESPONDENTS**

The study sample will serve as the main test group. The control group will come from the production sample for the quarters used for fielding. In the test group, one respondent in each Consumer Unit (CU) will complete a two week online diary or two one-week paper diaries. Respondents will be provided an online diary if they qualify based on responses to screening questions. If they do not qualify based on the screening questions or are otherwise unwilling or unable to use the online diary, then they will be provided with two one-week paper diaries. The control group from the production sample will be completing two one-week paper diaries. The paper diaries in the test group and control will be identical.

### **3. METHODOLOGY**

#### **Recruitment and incentives:**

The selected initial sample will receive a pre-notification postcard by U.S. mail followed by an advance letter via priority mail. The advance letter will provide information that is currently included in the CE diary survey (CED) advance letter, modified to streamline the language and include an FAQ of the LSF test and diary structure.

For the large-scale online diary test, the overall study sample will be divided between two, equally distributed test groups:

- **Group 1** respondents will receive a \$5 bill, an unconditional token incentive with their advance letter via U.S. Priority Mail.
- **Group 2** respondents will receive no incentive.

### **Diary placement procedures**

The diary placement visit protocol will be similar to current CED placement procedures, except that every household respondent will be screened to see if they have in-home internet access and access the internet at least a few times a week (by self-report). Respondents will be provided an online diary if they qualify based on responses to screening questions. If they do not qualify based on the screening questions or are otherwise unwilling or unable to use the online diary, then they will be provided with a paper diary which will be identical to the paper diary used in the CED. Members that keep an online diary will be provided with login-credentials and instructive materials for the online diary. The Census field representative (FR) will review the diary-keeping procedures for the online or paper version (if applicable) with the respondent and provide materials and information that he or she will need during the diary-keeping weeks. The FR will ask the respondent to login during the diary pickup visit, if possible. The FR will ask for the diary keeper's email address and phone number for contacting the respondent mid-week. Online diary respondents who provided an email address, will receive a welcome email with links to the diary, the user guide, and video tutorials.

### **Diary keeping procedures**

The respondent will keep the diary recording expenditures for all household members beginning the day following the diary placement interview. During the diary-keeping week, paradata will be collected on the logins for the online diary. A report will be generated for FRs that will identify the number of successful and unsuccessful logins by respondents and summary information of expenditures entered in the diary. FRs will use this information to email or call the diarist and encourage him/her to record their expenditures or to solve any technical issues. Regardless of logins, the FR will call the respondent mid-week to further encourage the diary keeping task.

A help desk will be required to assist respondents using online diaries with technical issues during data collection. Respondents will be able to contact the help desk by phone or email during extended business hours and on the weekend. The help desk will keep a record of respondent contact and reasons for contact.

### **Diary pickup procedures**

The diary pickup visit will begin with a review of the diary expenditures if the respondent is willing to log in with the FR. The FR will have guidelines for picking up the diaries, similar to current production. The FR will also collect any additional expenditures that may have been missed during the diary-keeping week through a recall process similar to production CED. Additional expenditures will be entered into the recall tab of the CAPI instrument. The diary pickup visit also includes the administration of the income section of the CAPI instrument. In the LSF, we will allow the second visit to be conducted over the phone if necessary, as in production.

The diary completion requirement is one household diary per week (which includes at least one expenditure or indication of a week without expenditures) from the Consumer Unit. Since each week is treated independently, if a respondent completes one week's diary and does not complete the second week's diary, it will count as one completed diary and one Type A non-completed diary. Total recall diaries (i.e. expenditures for the diary week entered by the interviewer into the CAPI instrument recall tab based on information given by the respondent at the time of diary pickup and not entered in the online or paper diary) will satisfy the diary completion requirement if the majority of expenses are provided through receipts. Total recall diaries *without* the majority of expenses provided through receipts (i.e., diaries are being completed by memory at pickup) only satisfy the diary completion requirement under certain conditions depending on the day of pickup.

### **Respondent debriefing**

After the diary pickup is completed, the respondent will be asked a series of debriefing questions. The debriefing questions will be self-administered by turning the FR's laptop over to the respondent

to complete. FRs will provide instructions and go through a simple example with the respondent on how to respond to the questions on the laptop. If respondents are unwilling or unable to self-administer the questions, the FR can administer the respondent debriefing. The questions will cover various topics including:

- perceptions of burden
- helpfulness of the instructions and materials
- reactions to advance letter (including token incentive for the incentive group)
- perceived task difficulty
- frequency of shopping trips and reasons for any deviations from normal shopping frequency in the diary keeping period
- preferences for online vs. other methods of data collection such as using an app, collecting receipts, or photographing and emailing receipts

### 3. CAPI INSTRUMENT CHANGES

Census will modify the CED CAPI instrument to include:

1. Screening questions: Two questions have been added to the CAPI instrument to screen respondents for the online diary. The first question is about internet access and the second asks about frequency of internet use. If the respondent has in-home internet access and accesses the internet at least a few times a week, they qualify for the online household diary.
2. Respondent debriefing questions: At the end of the survey, respondents will be asked to complete a self-administered module of questions as described in the section on respondent debriefing.
3. CAPI expense entry by FRs: A recall tab is added to the CED CAPI instrument for FRs to enter expenses based on information given by the respondent at the time of diary pickup which have not been entered in the online or paper diary
4. FR debriefing questions: Questions have been added to the CAPI instrument to ask FRs about their experience with the online diary and get feedback on online diary procedures, instruments, and materials.

#### 4. SAMPLING AND SAMPLE SIZE

The sample for the LSF will be drawn from the CED reserve sample. The sample will not include Alaska and Hawaii. The final sample will include respondents living in group quarters. The projected total size of the final sample is 2,502 units, consisting of 2,478 unit frame sample units and 22 group quarters frame units. Half of all sample units will be randomly selected to receive a \$5 token incentive, regardless of whether they actually participate in the test. The other half of the sample units are to receive no compensation.

#### 6. RESPONDENT MATERIALS

1. Pre-notification postcards (Attachment 1)
2. Advance letters (Attachment 2)
3. Respondent emails (Attachment 3)
4. Online diary user guide (Attachment 4)
5. Online diary tutorial script (Attachment 5)
6. Online diary screenshots (Attachment 6)
7. CED 2019 Paper Diary (Attachment 7)
8. CAPI instrument specifications (Attachment 8)

### III. BURDEN HOURS BY CALENDAR YEAR

Burden estimates for Calendar years 2019 and 2020 are shown in Tables 1 and 2 below.

**Table 1: Burden estimates for Calendar Year 2019 (October-December 2019)**

Response Burden for the LSF Online Diary Test					
Form	Total Respondents	Frequency	Total Responses	Average Time per Response	Estimated Total Burden Hours
LSF Online Diary Record-keeping	600	2	1200	70	1400
LSF Online Diary Interview including	600	2	1200	25	500

Respondent Debriefing					
TOTALS	600	////////	2400	////////	1900
<p>(1) The total number of respondents for the LSF is 1200. The respondents who do the record-keeping (entering expenses in diaries) are the same people who participate in the interviews (diary placement and pickup).</p> <p>(2) Respondent debriefing is an additional self-administered module added to the CAPI instrument to take place at the diary pickup.</p>					

**Table 2: Burden estimates for Calendar Year 2020 (January-March/April 2020)**

Response Burden for the LSF Online Diary Test					
Form	Total Respondents	Frequency	Total Responses	Average Time per Response	Estimated Total Burden
LSF Online Diary Record-keeping	600	2	1200	70	1400
LSF Online Diary Interview including Respondent Debriefing	600	2	1200	25	500
TOTALS	600	////////	2400	////////	1900
<p>(1) The total number of respondents for the LSF is 1200. The respondents who do the record-keeping (entering expenses in diaries) are the same people who participate in the interviews (diary placement and pickup).</p> <p>(2) Respondent debriefing is an additional self-administered module added to the CAPI instrument to take place at the diary pickup.</p>					