

---

# **RESULTS FROM TESTING A WEB MODE FOR THE CONSUMER EXPENDITURE DIARY SURVEY**

---

**VERSION DATE: 05/01/2014**

**Ian Elkin  
Laura Erhard  
Brett McBride**

**U.S. Bureau of Labor Statistics  
Division of Consumer Expenditure Survey  
Branch of Research and Program Development**

**Dawn Nelson**

**U.S. Bureau of the Census  
Center for Survey Measurement**

## Executive Summary

### 1. Overview – Purpose of Web Diary Field Test

The Consumer Expenditure (CE) Diary Survey relies on a pencil-and-paper instrument to collect data from multiple household members. There are a number of drawbacks to this method, such as limiting entry to a single diarist in one location, who must carry the diary with them throughout the day as well as the respondents' limited access to the paper diary from multiple locations if the diary is left at home. A web diary (WD) has the potential to address these limitations and, over time, the feasibility and potential benefits of using a web diary have increased.

### 2. Objectives

The Web Diary Feasibility Test endeavored to:

- a. Finalize a “usable” CE web diary, with “usable” meaning that the diary:
  - i. can be administered to respondents.
  - ii. can collect respondents’ data.
  - iii. can retrieve and integrate respondents’ data into standard CE data files.
  - iv. has a secure, workable log-in.
  - v. can be turned around for Production work (given certain documented changes are made) or another research test (in case the team finds there are some operational issues that need further investigation).
  - vi. has been tested and found to be implementable by Census.
- b. Understand the operational issues regarding implementing a web mode for the CE, including respondent and interviewer reactions to the process
- c. Understand the data quality and response rates associated with a web mode for the CE

### 3. Success Criteria

The final CE web diary is considered to have met the chartered objectives. The instrument developed and fielded (1) can be administered to respondents, (2) can collect respondents’ data, (3) can retrieve and integrate respondents’ data into standard CE data files, (4) has a secure log-in, (5) can be used as part of another research test (Individual Diaries Feasibility Test), and (6) has been tested by Census. Through multiple types of analysis, the team understands the operational issues regarding the implementation of a web mode for the CE as well as the data quality and response rates associated with it and what adjustments can be implemented as part of future research endeavors.

While the web diary, in its current state, is not ready for production, the issues are surmountable and valuable lessons have been learned. The web diary test was successful in generating useful data to inform and improve the individual diaries feasibility test. In addition, the web diary test was successful in moving CE on the path towards a web diary for eventual implementation.

## 4. Findings

### a. Sample Performance

Web diary sample performance rates, such as response rates, were lower than corresponding rates in the production diary. Lower sample performance rates are a concern; however, through the web diary test, we have determined a number of factors that may have had an impact on sample performance and are working on methods to address them.

### b. Expenditure Comparisons

Overall, there were 13 fewer median expenditures for web diary respondents when compared to corresponding respondents in CE production. Lower expenditure totals are a concern; however, strengths of a multimodal diary approach are exhibited as web diary respondents:

- reported higher median expenditure amounts for certain sections
- exhibited lower item nonresponse rates, as defined for a diary instrument, when compared with CE production
- week-to-week drop-off rates validate a shift to a one week collection period
- that logged in within the first three days and completed both diary weeks, had comparable median expenditure totals, when compared with restricted CE production
- that logged in within the first three days of the reporting period, had greater median number of entries, when compared with restricted CE production

### c. Paradata Analysis

Among respondents who successfully logged into the web diary, approximately 17 percent of respondents logged in on one day and one day only compared to less than one percent of respondents who successfully logged into the instrument on all 14 days. However, a number of factors, such as issues with the transcription of usernames and passwords, have been identified and will be addressed in upcoming feasibility tests.

### d. Operational Issues

Throughout the testing of the web diary a large number of potential challenges for future implementation were uncovered. These challenges have been analyzed and strategies have been developed to mitigate these challenges in the individual diaries feasibility test and other future feasibility tests.

## 5. Recommendations

The Web Diary Analysis Team will schedule meetings with the Individual Diary Feasibility Team and CE managers to discuss the following action items, based on the web diary test:

### *Low Response Rates*

- a. [Create flexible, but secure, username and password requirements](#)
- b. [Restrict default usernames and passwords to specific characters](#)
- c. [Print User Guide on non-glossy paper to facilitate username/password transcription](#)
- d. [Change the test protocol to allow for FR collection and input of records, particularly receipts with numerous items](#)

### *Lower Expenditure Amounts*

- e. [Increase FR follow-up contacts and record information, through the CHI instrument, to assess disposition/outcome of follow-up contact](#)

### *Higher Rate of Total Recall*

- f. [Allow entry, into the instrument, of in-scope expenditures after the final day of collection](#)
- g. [Assess different protocols for allowing the FR access to a summary of the respondent's web diary entries](#)
- h. [Set Date field to default to current day](#)
- i. [Run daily analysis on the paradata to determine which respondents have not logged into the instrument as well as keeping a log of how many days have passed since the last successful logon by the respondent and, after a specific time period has lapsed, FRs will be notified to contact the respondent to remind them to enter any expenditures](#)

### *Low Placement Rates*

- j. [Create classroom training for any feasibility test that features enormous departures from past protocols](#)
- k. [Implement the most recent security protocols, displaying security assurances in a prominent place within the instrument, and educate FRs on the most relevant data security concerns](#)

### *Higher Ineligible Rates*

- l. [Complete future research to determine what operating systems were categorized as "Other"](#)
- m. [Include multiple survey modes in order to determine which modal offering is optimal and the most effective](#)

### *Other*

- n. [Employ a dialog box that opens prompting the FR that they are about to leave the password assignment screen](#)
- o. [Paradata should be formatted in a manner that is easily analyzed and readable](#)

# TABLE OF CONTENTS

|                                                                                        |                                                                   |    |
|----------------------------------------------------------------------------------------|-------------------------------------------------------------------|----|
| 1.                                                                                     | Overview – Purpose of Web Diary Field Test .....                  | 7  |
| 2.                                                                                     | Background .....                                                  | 7  |
| a.                                                                                     | High-level Timeline .....                                         | 8  |
| b.                                                                                     | Web Diary Feasibility Test Team Membership .....                  | 8  |
| 3.                                                                                     | Organization of this Report .....                                 | 8  |
| 4.                                                                                     | Sample Performance .....                                          | 8  |
| a.                                                                                     | Eligible Cases .....                                              | 9  |
| b.                                                                                     | Sample Demographics and Characteristics .....                     | 9  |
| c.                                                                                     | Overall Response Rate .....                                       | 11 |
| d.                                                                                     | Type B/C and Ineligibility Rates .....                            | 11 |
| e.                                                                                     | Placement Rates .....                                             | 13 |
| f.                                                                                     | Completion Rates .....                                            | 14 |
| g.                                                                                     | Total Recall Rates .....                                          | 15 |
| 5.                                                                                     | Expenditure Comparisons .....                                     | 16 |
| a.                                                                                     | Consumer Unit Characteristics .....                               | 16 |
| b.                                                                                     | Expenditure Results .....                                         | 17 |
| Number of Entries .....                                                                | 17                                                                |    |
| Amount of Expenditures Reported .....                                                  | 18                                                                |    |
| Entries and Expenditure per Week .....                                                 | 19                                                                |    |
| Entries and Expenditures by Section Type .....                                         | 20                                                                |    |
| c.                                                                                     | Expenditure Results by Characteristics of Respondent Log-in ..... | 22 |
| Association of Respondent Log-in Patterns with Entries and Reported Expenditures ..... | 22                                                                |    |
| Association of FR Follow-up Contacts with Entries and Reported Expenditures .....      | 24                                                                |    |
| Expenditures by Frequency of Respondent Log-ins .....                                  | 26                                                                |    |
| Comparison of Daily Expenditures between Web and Restricted Production Diaries .....   | 27                                                                |    |
| d.                                                                                     | Non-Expenditure Results .....                                     | 28 |
| Missing Data .....                                                                     | 28                                                                |    |
| Extent of Partial Recall Conducted .....                                               | 29                                                                |    |
| 6.                                                                                     | Paradata Analysis .....                                           | 29 |
| a.                                                                                     | Log-in Data .....                                                 | 30 |
| Successful Log-in Days .....                                                           | 30                                                                |    |

|            |                                                 |    |
|------------|-------------------------------------------------|----|
|            | <b>Respondent Log-ins by Day</b> .....          | 30 |
| <b>b.</b>  | <b>Access Device and Operating System</b> ..... | 31 |
| <b>7.</b>  | <b>Operational Issues</b> .....                 | 32 |
| <b>a.</b>  | <b>Respondent Experience</b> .....              | 32 |
|            | <b>Logging-In</b> .....                         | 32 |
|            | <b>Data Entry</b> .....                         | 33 |
|            | <b>Record Usage</b> .....                       | 34 |
|            | <b>Security</b> .....                           | 35 |
| <b>b.</b>  | <b>Respondent Device Information</b> .....      | 36 |
|            | <b>Browser</b> .....                            | 36 |
|            | <b>Operating System</b> .....                   | 36 |
| <b>c.</b>  | <b>Field Representative Experience</b> .....    | 37 |
|            | <b>Cooperation</b> .....                        | 37 |
|            | <b>CAPI</b> .....                               | 38 |
|            | <b>Receipts and Recall</b> .....                | 38 |
|            | <b>Follow-up Contacts</b> .....                 | 39 |
|            | <b>Training</b> .....                           | 39 |
| <b>8.</b>  | <b>Limitations</b> .....                        | 39 |
| <b>9.</b>  | <b>Recommendations</b> .....                    | 41 |
| <b>10.</b> | <b>References</b> .....                         | 43 |

## 1. Overview – Purpose of Web Diary Field Test

The Consumer Expenditure (CE) Diary Survey relies on a pencil-and-paper instrument to collect data from multiple household members. There are a number of drawbacks to this method, such as limiting entry to a single diarist in a one location, who must carry the diary with them throughout the day. A web diary (WD) has the potential to address these limitations and, over time, the feasibility and potential benefits of using a web diary have increased.

The potential advantages of a web diary over a paper form include:

1. Improved Reporting. There is evidence to suggest that accuracy of reporting may be improved through the use of a web diary (Couper, 2008).
2. Improved Response Rates. A web mode also has the potential to improve unit and item response rates.
3. Improved Access. A web mode would allow for easier access across consumer unit (CU) members since respondents can enter expenditure data from any internet-enabled location and would not be tied to a single instrument kept in one location.
4. Cost Savings. Finally, a web component has potential cost savings over paper-and-pencil interviewing (PAPI) due to reduced or eliminated materials, scanning, and data entry expenses.

## 2. Background

Beginning in January 2013, CE fielded a 3-month test to determine the feasibility of collecting expenditure data via a web diary. For analysis purposes, internet access was used to match these data with CE production data and analyzed to determine differences using (1) age, (2) race, (3) sex, (4) education, (5) income, and (6) household size. The purpose of this report is to present findings and provide recommendations for implementation and further research regarding sample performance and other associated operational issues. In particular, to address whether increased usability of a web diary results in higher response rates, if the web diary led to an increase in respondents accepting and completing the survey, and if difficulties reported by data collectors in administering the web diary are prohibitive.

Using the production sample, the web diary test sample was drawn of addresses (n=1,200) from the six Census regions: Atlanta, Chicago, Denver, Los Angeles, New York, and Philadelphia. Within the six regions, a total of 197 Field Representatives (FR) collected data in 82 Federal Information Processing Standards (FIPS) counties. At each address, the FR asked if the respondent had access to the internet at home and, if so, through what device they accessed the internet. Only CUs who reported access to the internet at home through either a personal computer or tablet were eligible to participate in the feasibility test. After determining that the CU was eligible, the FR attempted to place the web diary with the respondents. All web diaries in the field test were “double-placed” meaning a field representative visited the CU at the beginning of the first week’s diary to explain diary procedures and provide the respondent with the needed user names and passwords and then did not return until after the second week’s diary.

### a. High-level Timeline

- Project Start 09/01/2011
- Kick-off Meeting 10/18/2011
- Usability Test: Round One 01/26/2012 – 04/24/2012
- Internal Pilot Test 04/25/2012 – 06/08/2012
- Usability Test: Round Two 04/25/2012 – 07/16/2012
- Training 12/17/2012 – 12/31/2012
- Main Data Collection 01/01/2013 – 04/30/2013

### b. Web Diary Feasibility Test Team Membership

- Kathy Downey – BLS
- Ian Elkin – BLS
- Michelle Freeman – BLS
- John Gloster – Census
- Lauren Kirkpatrick – Census
- Bret McBride – BLS
- Chris Stringer – Census

## 3. Organization of this Report

This report focuses on three aspects of the web diary feasibility test: (1) sample performance, (2) expenditure comparisons, (3) paradata analysis and (4) operational issues. Due to the small sample size of the test group, significance testing was not conducted for any of the analyses.

## 4. Sample Performance

This section will present and analyze the demographics and other characteristics and sample performance rates between the web diary sample, a restricted production sample, and the full production sample. The web diary sample consists of two potential diaries from all addresses that were drawn for the web diary feasibility test. Each potential diary could have resulted in a Type A, Type B, or Type C diary, an ineligible diary based on web diary sample criteria, or a completed diary. This sample will be compared to the full production sample of potential diaries from the same time period (January – March 2013) and a restricted production sample of the same time period. The restricted production sample excludes cases from the full production sample that would have been identified as ineligible given based on the selection criteria of the web diary feasibility test. This section will present a description of how each sample was formed, an analysis of a selection of demographics and characteristics for each sample, and sample performance rates, including overall response rates, type B/C and ineligibility rates (for the web diary and restricted production sample, placement rates, week 1 and week 2 completion rates, and total recall rates).



### **a. Eligible Cases**

A CU was only eligible to participate in the web diary field test if certain conditions were met. First, the interview would be conducted in English. Second, the household could not include more than one CU (“spawned cases”). Lastly, CU’s were required to have home internet access via a PC or tablet device. If any of these conditions were not met, then the case would be considered ineligible and the diary could not be placed. These diaries are considered “ineligible diaries” and were included with Type B/C non-interviews.

The full production sample used in comparisons consisted of all cases that were used to field the Diary for CE’s production during the same January through March 2013 time period. This included Type A and Type B/C cases.

From the full production sample, a restricted sample of production cases was created to match the same eligibility criteria of the web diary. First, interviews that were conducted in Spanish or another language were excluded, reducing the sample by 170 diaries. Second, diaries from CU in multiple-CU households were excluded, further reducing the sample by 114 diaries. Lastly, households that reported no home internet access or internet access only through a mobile phone, during placement, were excluded, reducing the sample by 1,041 diaries.

For the purpose of looking at the sample performance, the web diary test sample was compared to a restricted production sample. The production sample was not limited to cases that were double-placed since a double-placed diary should not affect willingness to complete the diary. However, in the web diary feasibility test subsequent analyses, focusing on reporting and expenditures, the effect of double-placement will need to be taken into consideration.

### **b. Sample Demographics and Characteristics**

Most of the demographics and characteristics of web diary, the restricted production sample, and the full production sample were similar, though several major differences were found. The average age of the reference person was roughly the same for all the samples (around 50 years old). The average size of the CU was within 0.1 between the samples. There was a slightly higher percentage of Whites and Asians in the web diary sample compared to the other two samples by about two percentage points. Both the web diary sample and the restricted production sample had similar rates of home ownership that were much higher than the full production sample (72-73 percent compared to 64 percent). The major difference in the samples was in education level of the reference person. The level of education of the web diary sample was much higher than that of both the full and restricted production samples. To a certain extent, a higher education level for those reporting home internet access is not surprising; however, even when restricting the production sample to those with internet access, the percent of the sample with a college education was still 5.3 percent lower than that of the web diary sample. Another striking differences is the percent of homeownership between the three samples. In these preliminary stages of analysis, processed income data were not yet available. Higher incidence of college education and home ownership tends to be correlated with higher incomes, which would suggest that the web diary sample also had a higher average income than the restricted and full production samples.

Section 5 will focus on reporting and expenditures. As part of this analysis, weighting will be used so that the demographics match the web diary sample for more meaningful comparisons.

*Select Demographics and Characteristics*

|                               | Web Diary<br>Sample<br>(N = 456) | Restricted<br>Production<br>(N = 2,261) | Full<br>Production<br>(N = 3,251) |
|-------------------------------|----------------------------------|-----------------------------------------|-----------------------------------|
| <i>Average Age</i>            | 50.1                             | 49.6                                    | 50.1                              |
| <i>Average CU Size</i>        | 2.6                              | 2.5                                     | 2.4                               |
| <b>Race (percent)</b>         |                                  |                                         |                                   |
| <i>White &amp; Other Race</i> | 87.1                             | 86.5                                    | 85.1                              |
| <i>Black</i>                  | 6.8                              | 9.2                                     | 10.9                              |
| <i>Asian</i>                  | 6.1                              | 4.2                                     | 4.0                               |
| <b>Ethnicity (percent)</b>    |                                  |                                         |                                   |
| <i>Hispanic</i>               | 12.7                             | 9.8                                     | 14.6                              |
| <b>Gender (percent)</b>       |                                  |                                         |                                   |
| <i>Male</i>                   | 48.5                             | 51.0                                    | 49.6                              |
| <i>Female</i>                 | 51.5                             | 49.0                                    | 50.4                              |
| <b>Education (percent)</b>    |                                  |                                         |                                   |
| <i>Elementary</i>             | 0.9                              | 1.4                                     | 3.6                               |
| <i>High school</i>            | 21.1                             | 26.7                                    | 33.2                              |
| <i>College</i>                | 78.1                             | 71.8                                    | 62.8                              |
| <i>Never attended</i>         | -                                | 0.1                                     | 0.4                               |
| <b>Tenure (percent)</b>       |                                  |                                         |                                   |
| <i>Homeowner</i>              | 73.0                             | 72.3                                    | 64.4                              |
| <i>Renter</i>                 | 27.0                             | 27.7                                    | 35.6                              |

### c. Overall Response Rate

The overall response rate was calculated using the CE production response rate definition: the total number of good diaries divided by the total number of eligible diaries (good diaries plus the total Type A non-response diaries). Type A non-response includes diaries that were not completed due to respondent refusal, inability to contact the respondent, and blank diaries (or majority of the diary recalled without receipts).

On average the response rate for the web diary sample was very poor at only 32 percent completing overall. The best achieved response rate for the test was by the Philadelphia Regional Office at 41.8 percent, which was still much lower than the full production response rates. The restricted production sample shows only a slightly lower response rate compared to the full production sample (60.8% compared to 64.5%) indicating that the sample restrictions only minimally contributed to lack of response.

Response rates for the web diary sample by specific Regional Offices (ROs) varied from 27.6 percent to 35.3 percent -- all lower than the both their restricted and full production counterparts. The Atlanta RO, which obtained the highest response rates overall for both the restricted production sample and the full production sample (70.0% and 70.8%, respectively) showed almost the lowest response rate for the web diary test (28.0%).

| <b>Overall Response Rates</b> |                         |                |              |                              |                |              |                        |                |              |
|-------------------------------|-------------------------|----------------|--------------|------------------------------|----------------|--------------|------------------------|----------------|--------------|
| <b>RO</b>                     | <b>Web Diary Sample</b> |                |              | <b>Restricted Production</b> |                |              | <b>Full Production</b> |                |              |
|                               | <b>Total Eligible</b>   | <b>Diaries</b> | <b>Rate</b>  | <b>Total Eligible</b>        | <b>Diaries</b> | <b>Rate</b>  | <b>Total Eligible</b>  | <b>Diaries</b> | <b>Rate</b>  |
| New York                      | 178                     | 54             | 30.3%        | 536                          | 301            | 56.2%        | 708                    | 435            | 61.4%        |
| Philadelphia                  | 239                     | 100            | 41.8%        | 722                          | 410            | 56.8%        | 1,008                  | 609            | 60.4%        |
| Chicago                       | 273                     | 86             | 31.5%        | 560                          | 373            | 66.6%        | 761                    | 550            | 72.3%        |
| Atlanta                       | 286                     | 80             | 28.0%        | 604                          | 423            | 70.0%        | 845                    | 598            | 70.8%        |
| Denver                        | 156                     | 55             | 35.3%        | 584                          | 329            | 56.3%        | 809                    | 503            | 62.2%        |
| Los Angeles                   | 294                     | 81             | 27.6%        | 710                          | 425            | 59.9%        | 910                    | 556            | 61.1%        |
| <b>Overall</b>                | <b>1,426</b>            | <b>456</b>     | <b>32.0%</b> | <b>3,716</b>                 | <b>2,261</b>   | <b>60.8%</b> | <b>5,041</b>           | <b>3,251</b>   | <b>64.5%</b> |

### d. Type B/C and Ineligibility Rates

Type B/C rates were calculated by taking the total number type B/C diaries and dividing by the total number of potential diaries in the sample (e.g., all drawn addresses). Type B/C diaries for the web diary sample included diaries that would be considered Type B/C under normal production conditions (e.g., vacant home, demolished, located on military base, etc.) and also diaries that were ineligible because the CU did not meet the web diary criteria. For the restricted production data, the Type B/C diaries were those that would be Type B/C under normal production conditions plus diaries that would have been ineligible if the web diary selection criteria (e.g., no internet access, non-English speaking household, or more than one CU per household) had been imposed on the CU. The full production sample type B/C diaries only reflect those that are actually type B/C in the production

environment. Ineligibility rates are only applicable for the web diary sample and restricted production sample and are calculated by dividing the number of ineligible or would-be ineligible diaries based on the web diary selection criteria by the total number of potential diaries in the respective sample. Because ineligibility rates are not applicable for the full production sample, the analysis focuses on a comparison between the web diary sample and restricted production sample, using the full production sample to contrast what the type B/C rates would have been given no ineligibility criteria.

Type B/C rates between the web diary and restricted production sample are similar at 46.1 percent and 41.7 percent respectively. The web diary Type B/C rates are slightly higher due to ineligibility rates about seven percent higher for the whole population. The likelihood of an address falling into a Type B/C category is independent of the selection criteria for the web diary test, so as expected, excluding the ineligible cases leads to a normal Type B/C rate of 18.8 percent for the web diary, close to the full production rate of 20.9 percent (which is the same as the restricted production, excluding added ineligibles).

Looking across RO's, similar patterns were seen as in the overall picture. In most RO's the ineligibility rate was roughly four to five percent higher in the web diary sample than in the restricted production sample. Though Denver showed a greater dispersion from the web diary test to the restricted production sample with 9 percent more diaries determined ineligible during the diary test.

| <b>TYPE B/C Rates (incl. ineligible)</b> |                         |                    |              |                              |                    |              |                        |                    |              |
|------------------------------------------|-------------------------|--------------------|--------------|------------------------------|--------------------|--------------|------------------------|--------------------|--------------|
| <b>RO</b>                                | <b>Web Diary Sample</b> |                    |              | <b>Restricted Production</b> |                    |              | <b>Full Production</b> |                    |              |
|                                          | <b>Total</b>            | <b>B/C Diaries</b> | <b>Rate</b>  | <b>Total</b>                 | <b>B/C Diaries</b> | <b>Rate</b>  | <b>Total</b>           | <b>B/C Diaries</b> | <b>Rate</b>  |
| New York                                 | 310                     | 132                | 42.6%        | 894                          | 358                | 40.0%        | 894                    | 186                | 20.8%        |
| Philadelphia                             | 400                     | 161                | 40.3%        | 1,242                        | 520                | 41.9%        | 1,242                  | 234                | 18.8%        |
| Chicago                                  | 538                     | 265                | 49.3%        | 974                          | 414                | 42.5%        | 974                    | 213                | 21.9%        |
| Atlanta                                  | 596                     | 310                | 52.0%        | 1,150                        | 546                | 47.5%        | 1,150                  | 305                | 26.5%        |
| Denver                                   | 302                     | 146                | 48.3%        | 1,032                        | 448                | 43.4%        | 1,032                  | 223                | 21.6%        |
| Los Angeles                              | 502                     | 208                | 41.4%        | 1,084                        | 374                | 34.5%        | 1,084                  | 174                | 16.1%        |
| <b>Overall</b>                           | <b>2,648</b>            | <b>1,222</b>       | <b>46.1%</b> | <b>6,376</b>                 | <b>2,660</b>       | <b>41.7%</b> | <b>6,376</b>           | <b>1,335</b>       | <b>20.9%</b> |

| <b>Ineligibility Rates</b> |                         |                           |              |                              |                           |              |                        |                           |             |
|----------------------------|-------------------------|---------------------------|--------------|------------------------------|---------------------------|--------------|------------------------|---------------------------|-------------|
| <b>RO</b>                  | <b>Web Diary Sample</b> |                           |              | <b>Restricted Production</b> |                           |              | <b>Full Production</b> |                           |             |
|                            | <b>Total</b>            | <b>Ineligible Diaries</b> | <b>Rate</b>  | <b>Total</b>                 | <b>Ineligible Diaries</b> | <b>Rate</b>  | <b>Total</b>           | <b>Ineligible Diaries</b> | <b>Rate</b> |
| New York                   | 310                     | 73                        | 23.5%        | 894                          | 172                       | 19.2%        | N/A                    | N/A                       | N/A         |
| Philadelphia               | 400                     | 109                       | 27.3%        | 1,242                        | 286                       | 23.0%        | N/A                    | N/A                       | N/A         |
| Chicago                    | 538                     | 171                       | 31.8%        | 974                          | 201                       | 20.6%        | N/A                    | N/A                       | N/A         |
| Atlanta                    | 596                     | 156                       | 26.2%        | 1,150                        | 241                       | 21.0%        | N/A                    | N/A                       | N/A         |
| Denver                     | 302                     | 88                        | 29.1%        | 1,032                        | 225                       | 21.8%        | N/A                    | N/A                       | N/A         |
| Los Angeles                | 502                     | 126                       | 25.1%        | 1,084                        | 200                       | 18.5%        | N/A                    | N/A                       | N/A         |
| <b>Overall</b>             | <b>2,648</b>            | <b>723</b>                | <b>27.3%</b> | <b>6,376</b>                 | <b>1,325</b>              | <b>20.8%</b> | <b>N/A</b>             | <b>N/A</b>                | <b>N/A</b>  |

**e. Placement Rates**

A placement is defined as a diary with placement codes indicating that the diary was successfully left with the household (or the equivalent username/password for the web diary) or the CU was temporarily absent (i.e., PLCECODE = ‘201’ or ‘217’), and so placement was not possible. The placement rate is the total number of placements divided by the total number that could have been placed. The total number of diaries that could have been placed are those that were not Type B/C non-interviews and are referred to in the tables as “Eligible Diaries”.

Placement rates were fairly similar between the samples. Full production had the highest placement rate overall at 74.3 percent. Accounting for eligibility differences with the restricted populations reduced the rate to 67.9 percent overall. The web diary sample rate was 62.4 percent overall. Looking at the individual RO’s, the web diary placement rates were higher than the restricted production sample for Denver and Philadelphia. Further analysis on the number of contact attempts and other information about the contact history will be conducted for the second part of the report that may provide insight on the differences in placement rates between the test and production samples.

| <b>Placement Rates</b> |                         |                          |              |                              |                          |              |                        |                          |              |
|------------------------|-------------------------|--------------------------|--------------|------------------------------|--------------------------|--------------|------------------------|--------------------------|--------------|
| <b>RO</b>              | <b>Web Diary Sample</b> |                          |              | <b>Restricted Production</b> |                          |              | <b>Full Production</b> |                          |              |
|                        | <b>Total Eligible</b>   | <b>Total Place-ments</b> | <b>Rate</b>  | <b>Total Eligible</b>        | <b>Total Place-ments</b> | <b>Rate</b>  | <b>Total Eligible</b>  | <b>Total Place-ments</b> | <b>Rate</b>  |
| New York               | 178                     | 96                       | 53.9%        | 536                          | 327                      | 61.0%        | 708                    | 481                      | 67.9%        |
| Philadelphia           | 239                     | 157                      | 65.7%        | 722                          | 466                      | 64.5%        | 1,008                  | 722                      | 71.6%        |
| Chicago                | 273                     | 153                      | 56.0%        | 560                          | 388                      | 69.3%        | 761                    | 587                      | 77.1%        |
| Atlanta                | 286                     | 191                      | 66.8%        | 604                          | 480                      | 79.5%        | 845                    | 711                      | 84.1%        |
| Denver                 | 156                     | 111                      | 71.2%        | 584                          | 375                      | 64.2%        | 809                    | 583                      | 72.1%        |
| Los Angeles            | 294                     | 182                      | 61.9%        | 710                          | 486                      | 68.5%        | 910                    | 659                      | 72.4%        |
| <b>Overall</b>         | <b>1,426</b>            | <b>890</b>               | <b>62.4%</b> | <b>3,716</b>                 | <b>2,522</b>             | <b>67.9%</b> | <b>5,041</b>           | <b>3,743</b>             | <b>74.3%</b> |

#### f. Completion Rates

The completion rate looks at the total number of completed diaries over the number of successful placements. A completed diary is one that had expenditure entries, or was completed through recall by the FR with receipts, or the day after the diary week ended. Any diaries that did not meet a minimum number of entries/expenditures would have been converted to a Type A non-interview during the Initial Edit System's Minimum Expenditure Edit and were not considered complete.

There are two points that are apparent when comparing the completion rates between the web diary sample and the production samples. First, overall the completion rates were much lower in the web diary sample. Second, the completion rates from week one to week two dropped off by about six percent, whereas in the production samples, completion rates remained relatively constant between the two weeks. Interestingly, the completion rate for the restricted production sample was even higher than the full production sample overall and across all RO's.

The low completion rates is a cause for concern with the web diary instrument; however, operational issues associated with the web diary test may have led to these lower rates.

| Week 1 Completion Rates |                   |                  |              |                       |                  |              |                   |                  |              |
|-------------------------|-------------------|------------------|--------------|-----------------------|------------------|--------------|-------------------|------------------|--------------|
| RO                      | Web Diary Sample  |                  |              | Restricted Production |                  |              | Full Production   |                  |              |
|                         | Total Place-ments | Complete Diaries | Rate         | Total Place-ments     | Complete Diaries | Rate         | Total Place-ments | Complete Diaries | Rate         |
| New York                | 46                | 27               | 58.7%        | 160                   | 147              | 91.9%        | 236               | 210              | 89.0%        |
| Philadelphia            | 76                | 50               | 65.8%        | 229                   | 200              | 87.3%        | 348               | 295              | 84.8%        |
| Chicago                 | 72                | 44               | 61.1%        | 193                   | 186              | 96.4%        | 292               | 274              | 93.8%        |
| Atlanta                 | 92                | 40               | 43.5%        | 235                   | 207              | 88.1%        | 348               | 293              | 84.2%        |
| Denver                  | 55                | 29               | 52.7%        | 183                   | 161              | 88.0%        | 286               | 245              | 85.7%        |
| Los Angeles             | 85                | 42               | 49.4%        | 236                   | 209              | 88.6%        | 319               | 272              | 85.3%        |
| <b>Overall</b>          | <b>426</b>        | <b>232</b>       | <b>54.5%</b> | <b>1,236</b>          | <b>1,110</b>     | <b>89.8%</b> | <b>1,829</b>      | <b>1,589</b>     | <b>86.9%</b> |

| Week 2 Completion Rates |                   |                  |              |                       |                  |              |                   |                  |              |
|-------------------------|-------------------|------------------|--------------|-----------------------|------------------|--------------|-------------------|------------------|--------------|
| RO                      | Web Diary Sample  |                  |              | Restricted Production |                  |              | Full Production   |                  |              |
|                         | Total Place-ments | Complete Diaries | Rate         | Total Place-ments     | Complete Diaries | Rate         | Total Place-ments | Complete Diaries | Rate         |
| New York                | 50                | 27               | 54.0%        | 167                   | 154              | 92.2%        | 245               | 225              | 91.8%        |
| Philadelphia            | 81                | 50               | 61.7%        | 237                   | 210              | 88.6%        | 374               | 314              | 84.0%        |
| Chicago                 | 81                | 42               | 51.9%        | 195                   | 187              | 95.9%        | 295               | 276              | 93.6%        |
| Atlanta                 | 99                | 40               | 40.4%        | 245                   | 216              | 88.2%        | 363               | 305              | 84.0%        |
| Denver                  | 56                | 26               | 46.4%        | 192                   | 168              | 87.5%        | 297               | 258              | 86.9%        |
| Los Angeles             | 97                | 39               | 40.2%        | 250                   | 216              | 86.4%        | 340               | 284              | 83.5%        |
| <b>Overall</b>          | <b>464</b>        | <b>224</b>       | <b>48.3%</b> | <b>1,286</b>          | <b>1,151</b>     | <b>89.5%</b> | <b>1,914</b>      | <b>1,662</b>     | <b>86.8%</b> |

**g. Total Recall Rates**

The total recall rate is defined as the number of diaries that were collected through total recall over the total number of completed interviews. Total recall is determined upon pick-up of the Diary. If the respondent did not fill in any expenditures for the week, but had expenditures, the FR could collect those expenditures through recall.

Overall, recall was much higher in the web diary sample with 21.3 percent of completed diaries (456 diaries) collected through total recall (97 diaries). This again, could be explained by some issues with the test – if logging

in was a problem for respondents, then having the FR complete the diary upon pick-up would have been the best option for collecting data.

| <b>Total Recall Rates (week 1 and week 2)</b> |                         |                     |              |                              |                     |             |                         |                     |              |
|-----------------------------------------------|-------------------------|---------------------|--------------|------------------------------|---------------------|-------------|-------------------------|---------------------|--------------|
| <b>RO</b>                                     | <b>Web Diary Sample</b> |                     |              | <b>Restricted Production</b> |                     |             | <b>Full Production</b>  |                     |              |
|                                               | <b>Complete Diaries</b> | <b>Total Recall</b> | <b>Rate</b>  | <b>Complete Diaries</b>      | <b>Total Recall</b> | <b>Rate</b> | <b>Complete Diaries</b> | <b>Total Recall</b> | <b>Rate</b>  |
| New York                                      | 54                      | 6                   | 11.1%        | 301                          | 2                   | 0.7%        | 435                     | 11                  | 2.5%         |
| Philadelphia                                  | 100                     | 19                  | 19.0%        | 410                          | 30                  | 7.3%        | 609                     | 51                  | 8.4%         |
| Chicago                                       | 86                      | 22                  | 25.6%        | 373                          | 63                  | 16.9%       | 550                     | 112                 | 20.4%        |
| Atlanta                                       | 80                      | 21                  | 26.3%        | 423                          | 56                  | 13.2%       | 598                     | 81                  | 13.5%        |
| Denver                                        | 55                      | 8                   | 14.5%        | 329                          | 19                  | 5.8%        | 503                     | 28                  | 5.6%         |
| Los Angeles                                   | 81                      | 21                  | 25.9%        | 425                          | 40                  | 9.4%        | 556                     | 53                  | 9.5%         |
| <b>Overall</b>                                | <b>456</b>              | <b>97</b>           | <b>21.3%</b> | <b>2,261</b>                 | <b>210</b>          | <b>9.3%</b> | <b>3,251</b>            | <b>336</b>          | <b>10.3%</b> |

## 5. Expenditure Comparisons<sup>1</sup>

This section describes comparisons of the data collected by the test web diary and comparable paper production diaries. Included is a description of the universe of CUs that were analyzed (compared to those identified in Section 4 of this report) as well as how the comparable CUs in the production sample were derived. The analysis describes CU characteristics of these groups, and information on the extent to which data was entered in each group through recall. The next set of analysis focuses on the number of entries and expenditure totals, which is presented by week and by expenditure category. Finally, missing data rates, by CU and by total entries are examined. At reported expenditure levels, there were not sufficient completed diaries to allow for statistically significant difference to be detected; however, even without statistical significance for inferential statistics, there is practical significance in the examination of trends that can yield information regarding the substantive differences between web diary data and comparable paper production data.

### a. Consumer Unit Characteristics

Since this analysis excluded Type A non-respondents and removed diaries that FRs indicated to have no entries upon pick-up from both the test and comparison datasets, the number of diary weeks presented in this section is lower than that reported in Section 4 of this report. These removed diaries that were picked up would have been

<sup>1</sup> For all expenditure analyses, the small sample sizes prevented the web diary team from controlling for age, income or household size characteristics, factors typically associated with differences in expenditure reporting levels.



considered completes because of the recall/receipts process, in which all of the entries were provided during the FR visit via ‘total recall’)<sup>2</sup>. This analysis retained diaries in which respondents provided some entries, but the FR augmented those entries using the recall/receipts process, a process which will be referred to as ‘partial recall’.

For test cases, there were 200 CUs with 355 completed diary-weeks (about 78% completed entries both weeks, 19 percent only having Week 1 entries and 4% only having Week 2 entries). For production cases, there were 1,435 CUs with 2,611 completed diary-weeks. The ‘Restricted Production’ (RP) cases included only CUs with English interviews, no spawned CUs, and CUs reporting Internet service (through a computer or tablet). For RP cases, there were 948 CUs with 1,749 completed diary-weeks (about 85 percent completed entries both weeks, 6 percent only having Week 1 entries and 10 percent only having Week 2 entries).

#### *Diary completion*

|                     | <b>Test<br/>CUs</b> | <b>RP CUs</b> | <b>Test<br/>Percent</b> | <b>RP<br/>Percent</b> | <b>Difference<br/>(Test-RP)</b> | <b>Change</b> |
|---------------------|---------------------|---------------|-------------------------|-----------------------|---------------------------------|---------------|
| Week 1 entries only | 38                  | 52            | 19.0                    | 5.5                   | 13.5                            | 245.5%        |
| Week 2 entries only | 7                   | 95            | 3.5                     | 10.0                  | -6.5                            | -65.0%        |
| Entries both weeks  | 155                 | 801           | 77.5                    | 84.5                  | -7.0                            | -8.3%         |
| Total               | 200                 | 948           | 100.0                   | 100.00                |                                 |               |

Test CUs had median pre-tax income \$5,000 greater the median income for RP CUs. Both samples had an average of approximately 2.5 CU members. 81 percent of test CU respondents completed at least some college education compared to 73 percent of RP CUs. Please note that because 2013 CE data has not yet been released, information on expenditure amounts has been suppressed.

#### *CU characteristics*

|                         | <b>Test<br/>(n=200)</b> | <b>RP<br/>(n=948)</b> | <b>Change</b> |
|-------------------------|-------------------------|-----------------------|---------------|
| Pre-tax income (mean)   | -                       | -                     | \$55.58       |
| Pre-tax income (median) | -                       | -                     | \$5,000.00    |
| CU size                 | 2.55                    | 2.56                  | -.01          |
| % with college educ.    | 81.0%                   | 73.2%                 | 7.8%          |

## **b. Expenditure Results**

### **Number of Entries**

Among cases that completed entries (including partial recall) in both diary weeks, the mean and median diary entries were calculated. Test CUs had 8 fewer median entries recorded than RP CUs. The 155 test CUs had an average of 66.9 entries across both weeks compared to an average of 75.6 entries for the 801 RP CUs.

<sup>2</sup> Those who had no expenditures despite FR indication there were entries were removed. A small proportion of diaries that were reported to have no entries but did have them were not removed from analysis.

*Total entries (among CUs entering 2 weeks of data without total recall)*

|                | <b>Test<br/>(n=155)</b> | <b>RP<br/>(n=801)</b> | <b>Difference<br/>(Test-RP)</b> | <b>Change</b> |
|----------------|-------------------------|-----------------------|---------------------------------|---------------|
| Mean entries   | 66.9                    | 75.6                  | -8.7                            | -11.5%        |
| Median entries | 61.0                    | 69.0                  | -8.0                            | -11.6%        |

When comparing Test CUs with RP CUs for which diaries were double-placed (meaning there were no FR visits between the first and second weeks) the difference increased to 13 fewer median entries (61 for test CUs and 74 for double-placed RP CUs).

*Total entries (among CUs entering 2 weeks of data without total recall), compared against double-placed RP CUs*

|                | <b>Test<br/>(n=155)</b> | <b>Double-<br/>Placed<br/>RP<br/>(n=281)</b> | <b>Difference<br/>(Test-RP)</b> | <b>Change</b> |
|----------------|-------------------------|----------------------------------------------|---------------------------------|---------------|
| Mean entries   | 66.9                    | 79.4                                         | -12.5                           | -15.7%        |
| Median entries | 61.0                    | 74.0                                         | -13.0                           | -17.6%        |

### Amount of Expenditures Reported

Using the same universe as above, total reported expenditures (ZTOTAL) were examined. In addition to reporting fewer entries, test cases had smaller total expenditure amounts than RP cases. Test cases had an average expenditure total that was 15.7 percent less than expenditure totals for RP cases.

|                | <b>Test<br/>(n=155)</b> | <b>RP<br/>(n=801)</b> | <b>Change</b> |
|----------------|-------------------------|-----------------------|---------------|
| Mean amounts   | -                       | -                     | -15.1%        |
| Median amounts | -                       | -                     | -17.0%        |

Test cases had median expenditure totals that were 17.6 percent less than the median expenditure totals for RP cases and 15.1 percent less than RP cases that were double-placed.

|                | <b>Test<br/>(n=155)</b> | <b>Double-<br/>Placed<br/>RP<br/>(n=281)</b> | <b>Change</b> |
|----------------|-------------------------|----------------------------------------------|---------------|
| Mean amounts   | -                       | -                                            | -7.6%         |
| Median amounts | -                       | -                                            | -15.0%        |

## Entries and Expenditure per Week

### *Number of Entries per Week*

To compare entries and expenditures across diary weeks within a CU, cases were subset to CUs that:

- completed both diary weeks
- had double-placed diaries, and
- were eligible for the web diary test (based on the RP eligibility criteria indicated above).

Additionally, CUs were subset to those not having any entries from partial recall.

Among these CUs, the 126 test CUs had provided an average of 37.8 entries the first week and 31.7 entries the second.

#### *Entries by CUs completing both weeks (no recall)*

|                        | <b>Test<br/>(n=126)</b> | <b>RP<br/>(n=244)</b> | <b>Difference<br/>(Test-RP)</b> | <b>Change</b> |
|------------------------|-------------------------|-----------------------|---------------------------------|---------------|
| Week 1 mean entries    | 37.8                    | 41.7                  | -3.9                            | -9.4%         |
| Week 1 median entries  | 35.0                    | 37.0                  | -2.0                            | -5.4%         |
| Week 2 mean entries    | 31.7                    | 38.3                  | -6.6                            | -17.2%        |
| Week 2 median entries  | 29.0                    | 32.5                  | -3.5                            | -10.8%        |
| Overall mean entries   | 69.5                    | 80.0                  | -10.5                           | -13.1%        |
| Overall median entries | 62.0                    | 75.5                  | -13.5                           | -17.9%        |

The average percent of entries coming in the first week of data collection for test CUs was 54.8 percent. The 244 RP CUs had an average of 41.7 entries the first week and 38.3 entries the second. The average percent of entries in the first week of data collection for RP CUs was 53.1 percent. The larger drop-off in entries among test CUs validates comments in the FR debriefing that respondents were less diligent about doing entries toward the end of the diary period.

#### *Within-CU percent of entries entered in first week by CUs completing both weeks (no recall)*

|                  | <b>Test<br/>(n=126)</b> | <b>RP (n=244)</b> |
|------------------|-------------------------|-------------------|
| Percent (Week 1) | 54.8%                   | 53.1%             |

### *Amount of Expenditures Reported per Week*

Using the same universe as above, total reported expenditures (ZTOTAL) were examined and compared. Test cases had a median total that were 12.8percent in the first week less than RP cases and 15.9 percent less than RP cases in the second week<sup>3</sup>.

|                        | <b>Test<br/>(n=126)</b> | <b>RP<br/>(n=244)</b> | <b>Change</b> |
|------------------------|-------------------------|-----------------------|---------------|
| Week 1 mean amounts    | -                       | -                     | -0.4%         |
| Week 1 median amounts  | -                       | -                     | -12.8%        |
| Week 2 mean amounts    | -                       | -                     | -25.1%        |
| Week 2 median amounts  | -                       | -                     | -15.9%        |
| Overall mean amounts   | -                       | -                     | -13.2%        |
| Overall median amounts | -                       | -                     | -15.5%        |

The average percent of total expenditures coming in the first week of data collection for test CUs was 52.4 percent compared to 52.1 percent for RP CUs.

*Within-CU percent of expenditure amounts entered in first week of CUs completing both weeks (no recall)*

|                  | <b>Test (n=126)</b> | <b>RP (n=244)</b> |
|------------------|---------------------|-------------------|
| Percent (Week 1) | 52.4%               | 52.1%             |

### **Entries and Expenditures by Section Type**

#### *Number of Entries by Section Type*

Test and RP samples had differences in expenditure entries in varying amounts by category. For part 2, ‘Food and Drinks Away from Home,’ and part 3, ‘Clothing, Shoes, Jewelry, and Accessories’, the differences were minimal. For part 2, ‘Food and Drinks Away from Home’, there was a test average of 10.0 entries compared to a RP average of 11.7 entries; for clothing, there was a test average of 2.2 entries compared to a RP average of 2.4 entries. Larger differences were found between the test and RP cases for part 4, ‘All Other Products, Services, and Expenses’ – 18.6 and 21.5 average entries respectively – and part 1, ‘Food and Drinks for Home Consumption’ in which test cases averaged 36.0 entries compared to 40.1 for the RP cases.

---

<sup>3</sup> This may appear to suggest increasing expenditure amounts, but these median amounts were skewed by dramatic increases among certain CUs. On average, 52% of a RP CUs’ expenditures were during the first week.

*Total entries by section (among CUs entering 2 weeks of data without total recall<sup>4</sup>)*

|                                                              | <b>Test<br/>(n=155)</b> | <b>RP<br/>(n=801)</b> | <b>Difference<br/>(Test-RP)</b> | <b>Change</b> |
|--------------------------------------------------------------|-------------------------|-----------------------|---------------------------------|---------------|
| 1) Food and Drinks for Home Consumption Mean Entries         | 36.0                    | 40.1                  | -4.1                            | -10.2%        |
| 1) Food and Drinks for Home Consumption Median Entries       | 33.0                    | 36.0                  | -3.0                            | -8.3%         |
| 2) Food and Drinks Away from Home Mean Entries               | 10.0                    | 11.7                  | -1.7                            | -14.5%        |
| 2) Food and Drinks Away from Home Median Entries             | 7.0                     | 8.0                   | -1.0                            | -12.5%        |
| 3) Clothing, Shoes, Jewelry, and Accessories Mean Entries    | 2.2                     | 2.4                   | -0.2                            | -8.3%         |
| 3) Clothing, Shoes, Jewelry, and Accessories Median Entries  | 1.0                     | 1.0                   | 0.0                             | 0.0%          |
| 4) All Other Products, Services, and Expenses Mean Entries   | 18.6                    | 21.5                  | -2.9                            | -13.5%        |
| 4) All Other Products, Services, and Expenses Median Entries | 17.0                    | 19.0                  | -2.0                            | -10.5%        |

**Amount of Expenditures Reported by Section Type**

In contrast to entries, the differences between test and RP cases were not as large for expenditure amounts with the exception of the ‘All Other Products, Services, and Expenses’ category.

*Expenditure totals by section (among CUs entering 2 weeks of data without total recall)*

|                                                                   | <b>Test<br/>(n=155)</b> | <b>RP<br/>(n=801)</b> | <b>Change</b> |
|-------------------------------------------------------------------|-------------------------|-----------------------|---------------|
| 1) Food and Drinks for Home Consumption Mean Expenditures         | -                       | -                     | 4.1%          |
| 1) Food and Drinks for Home Consumption Median Expenditures       | -                       | -                     | 10.0%         |
| 2) Food and Drinks Away from Home Mean Expenditures               | -                       | -                     | -6.3%         |
| 2) Food and Drinks Away from Home Median Expenditures             | -                       | -                     | -11.6%        |
| 3) Clothing, Shoes, Jewelry, and Accessories Mean Expenditures    | -                       | -                     | -8.1%         |
| 3) Clothing, Shoes, Jewelry, and Accessories Median Expenditures  | -                       | -                     | 47.7%         |
| 4) All Other Products, Services, and Expenses Mean Expenditures   | -                       | -                     | -18.8%        |
| 4) All Other Products, Services, and Expenses Median Expenditures | -                       | -                     | -36.3%        |

Test CUs reported a median total that was approximately 36.3 percent less in ‘other’ expenditures compared to RP CUs. However, the non-‘other’ categories had differences that were much smaller. Test CUs reported a median amount that was 11.6 percent less for ‘Food and Drinks Away from Home’ compared to RP CUs. There was very little difference in median total amounts between groups for the ‘Clothing, Shoes, Jewelry, and Accessories’ and

<sup>4</sup> For entries, this reflects the parts of the diary where respondents classified their expense entries.

For expenditure totals, (table titled “Mean, median expenditure totals: by section (among CUs entering 2 weeks of data without total recall)”) this reflects where entry amounts have been re-classified into the ‘correct’ parts of the diary by phase 3 processing.

A common example of this would be a respondent entering something in other (e.g., ‘purses’) that phase 3 processing put in a different diary part (e.g., ‘clothing & accessories’), so for the expenditure table, the amount would be in clothing, while for the entries table it would be in ‘all other products.’

‘Food and Drinks for Home Consumption’ categories. Test CUs actually reported 47.7 percent more in median totals for clothing and 10.0 percent more for ‘Food and Drinks for Home Consumption.’ Combined with the entry information, this suggests that test CUs entered larger average expenditure entries per clothing and grocery item than the equivalent CUs who used paper diaries.

### c. Expenditure Results by Characteristics of Respondent Log-in

#### Association of Respondent Log-in Patterns with Entries and Reported Expenditures

The absence of a log-in-week indicator on the paradata file and unexpected log-in dates complicated efforts to analyze the relation of log-in patterns with weekly expenditure totals. As a solution, placement date information from the expenditure data file was applied to the paradata file to identify expected reporting periods in which Week 1 and Week 2 log-ins would have occurred (e.g., placement day+1 through +8 for Week 1). However, examining a subsample of cases indicated almost all to have log-in dates falling outside of these expected periods, and almost a third did not have any log-in dates within the expected periods<sup>5</sup>. To compensate for this, the reporting periods were extended to include all dates up to the pick-up date, which effectively matched most log-in dates within diary reporting weeks; only two cases still had log-ins occurring outside of the reporting periods using this approach. In this manner, paradata information was matched to the expenditure file, although not all cases had a fixed range of log-ins dates (e.g., some had log-in dates 16 days apart matched to one diary week compared to the expected 7 day range).

Following the matching of paradata cases with those on the expenditure file not involving ‘total recall,’ there were a total of 187 CUs that had matched paradata, or 307 diary weeks. The characteristics of these CUs are presented below.

#### *Cases Matched with Paradata*

|                        | <b>CUs</b> | <b>Diaries</b> |
|------------------------|------------|----------------|
| Week 1 paradata only   | 35         | 35             |
| Week 2 paradata only   | 32         | 32             |
| Both weeks of paradata | 120        | 240            |
| <b>Total</b>           | <b>187</b> | <b>307</b>     |

Among these 307 matched diaries, the number of days elapsed from the start of the expected reporting period until diaries had initial log-ins was calculated, indicated below for the first 3 days.

---

<sup>5</sup> There are two possible explanations: 1) the field test’s process allowed respondents to indicate a start date for their expenditure reporting, which may not have conformed to the requested date they were supposed to begin reporting expenditures, and consequently led to logins occurring after the end of the expected reporting period; 2) If a FR indicated only one of the two diary weeks to have no entries but only recall, that ‘total recall’ diary week would not have been included in this analysis. However, respondents may have actually logged in to the instrument during that excluded week (even if they didn’t enter expenses), creating login dates outside of the expected reporting period.

*Period before Initial Log-in (Among 307 Diaries)*

|                                | <b>Diaries</b> | <b>Percent of matched</b> |
|--------------------------------|----------------|---------------------------|
| Log-in on first day            | 98             | 31.9%                     |
| Log-in within first two days   | 176            | 57.3%                     |
| Log-in within first three days | 212            | 69.1%                     |

Over two-thirds of complete diaries that had matched paradata involved a log-in within the first three days of the reporting period. It was interesting to note that, of CUs logging-in within the first three days, 11 percent logged-in on the same day the Web diary was placed with them (prior to the expected start of the reporting period).

At the CU-level, cases deemed to be ‘compliant’ CUs were the focus, in which there was a log-in within the first three days of the expected reporting period. These CUs had the following weekly completion characteristics:

*Whether CU Logged-in Within First 3 Days ('Compliant') by Weekly Completion Characteristics*

|              | <b>CUs</b> | <b>Percent 'Compliant'</b> |
|--------------|------------|----------------------------|
| Week 1 only  | 35         | 45.7%                      |
| Week 2 only  | 32         | 25.0%                      |
| Both weeks   |            |                            |
| In Week 1    | 120        | 78.3%                      |
| In Week 2    | 120        | 78.3%                      |
| <b>Total</b> | <b>307</b> | <b>69.1%</b>               |

CUs having a Web log-in only in Week 2 were the least likely to be ‘compliant’ CUs, or those logging in early in the diary week (25 percent). In contrast, CUs that had log-ins in both weeks had consistently high rates of logging in to the diary early in the expected reporting periods (78 percent for each week)<sup>6</sup>. The mean and median number of diary entries among the 120 CUs that completed both weeks of data entry were calculated by whether they were ‘compliant’ in logging-in to their Web diaries shortly after Week 1 placement.

*Number of 2-Week Entries by Whether CU Logged-in Within First 3 Days Of Week 1*

|        | <b>No (CU=26)</b> | <b>Yes/Compliant (CU=94)</b> | <b>Difference</b> |
|--------|-------------------|------------------------------|-------------------|
| Mean   | 65.4              | 76.9                         | -11.5             |
| Median | 61.5              | 72.5                         | -11.0             |

<sup>6</sup> Among CUs logging-in both weeks, there were 10 (8.3%) that were not compliant both weeks, 16 (13.3%) that were compliant in Week 1 but not Week 2, 16 (13.3%) that were compliant in Week 2 but not Week 1, and 78 (65%) that were compliant in both Week 1 and Week 2.

CUs completing two weeks of data entry and logging-in early in the reporting period had a greater average number of entries (77) compared to those that did not (65).

When comparing the number of entries with the comparable Restricted Production subsample, those logging-in within the first three days of the reporting period had similar average number of entries (77) as those in the paper diary RP sample (76).

*Comparison of 2-Week Entries between 'Compliant' CUs and RP CUs*

|        | <b>Yes/Compliant<br/>(CU=94)</b> | <b>RP (CU=801)</b> | <b>Difference</b> |
|--------|----------------------------------|--------------------|-------------------|
| Mean   | 76.9                             | 75.6               | 1.3               |
| Median | 72.5                             | 69.0               | 3.5               |

Additionally expenditure totals were calculated for the groups mentioned above.

*2-Week Expenditure Amounts by Whether CU Logged-in Within First 3 Days of Week 1*

|        | <b>No (CU=26)</b> | <b>Yes/Compliant<br/>(CU=94)</b> | <b>Difference</b> |
|--------|-------------------|----------------------------------|-------------------|
| Mean   | -                 | -                                | \$89.57           |
| Median | -                 | -                                | -\$146.06         |

CUs that logged-in within the first three days and completed both diary weeks had a smaller average expenditure total compared to those that did not, \$90 less, although they had larger median expenditure amounts, \$150 more.

*Comparison of 2-Week Expenditure Amounts between 'Compliant' CUs and RP CUs*

|        | <b>Yes/Compliant<br/>(CU=94)</b> | <b>RP (CU=801)</b> | <b>Difference</b> |
|--------|----------------------------------|--------------------|-------------------|
| Mean   | -                                | -                  | -\$127.31         |
| Median | -                                | -                  | -\$26.44          |

CUs that logged-in within the first three days and completed both diary weeks had comparable median expenditure totals as those in the RP subsample.

**Association of FR Follow-up Contacts with Entries and Reported Expenditures**

FRs enter contact attempts with sampled diary households in the Contact History Instrument (CHI). For the web diary test, FRs were instructed to follow-up with CUs on the 3<sup>rd</sup> and 8<sup>th</sup> days to remind them to enter expenses in their diaries. Although FRs are expected to enter every contact attempt they make, little is known about how diligently contact information is entered by FRs for diary cases. Furthermore, for this test, there was no way to clearly distinguish reminder contact attempts from other contact attempts made by FRs.



In this analysis, CHI records for CUs with a successful contact attempt were merged with expenditure data for the 200 Web CUs that had completed diaries and matching paradata. ‘Reminder’ contacts were identified by using placement date information to identify expected reporting periods in which Week 1 and Week 2 reminder contacts would have been made (using the same methods described in the Association of Respondent Log-in Patterns with Entries and Reported Expenditures section). CHI contacts were further subset to ensure that only telephone contacts with a member of the sample unit that occurred within the expected reporting period would be defined as ‘reminder’ contacts. The number of Web cases that had contact records is indicated below.

*Universe of Test CUs and CHI data*

|                                               | <b>CUs</b> |
|-----------------------------------------------|------------|
| Test CUs                                      | 200        |
| CUs with CHI records                          | 151        |
| CUs with a successful FR contact <sup>7</sup> | 92         |
| CUs with 'reminder' contacts                  | 18         |

As shown above, almost 25 percent of the 200 CUs did not have any CHI contact records (successful or otherwise). Almost half of CUs did receive a successful FR contact between placement and pick-up. There were only 18 CUs that had successful telephone contacts with a sample unit member within that time frame. This corresponded to 21 ‘reminder’ contacts that could be analyzed against reported diary expenditures and paradata on log-in dates.

There was no established mechanism by which to monitor a CU’s diary log-ins, and so the reminder contacts were not targeted to CUs that had not yet logged-in to the diary. This is evident in the examination of when the reminder contacts occurred relative to the CUs’ log-ins.

*Relation of 21 FR Contacts (CU=18) with Log-in Behaviors*

|                               | <b>CUs</b> | <b>Percent</b> |
|-------------------------------|------------|----------------|
| 1st Contact Before 1st Log-in | 3          | 16.7%          |
| 1st Contact Day of 1st Log-in | 2          | 11.1%          |
| 1st Contact After 1st Log-in  | 13         | 72.2%          |
| Total                         | 18         | 100.0%         |

72 percent of CUs had already logged-in once before receiving an initial FR reminder contact. To control for differences in the number of diary weeks completed by CUs with FR contacts, entry and expenditure amounts were analyzed at a weekly level, below.

---

<sup>7</sup> Within the expected reporting period.

*Weekly Entry Amounts by FR Contact*

|                 | <b>No FR contact<br/>(N=286)</b> | <b>FR contact<br/>(N=21)</b> | <b>Difference</b> |
|-----------------|----------------------------------|------------------------------|-------------------|
| Average Entries | 32.88                            | 33.52                        | -0.64             |
| Median Entries  | 29                               | 33                           | -4                |

*Weekly Reported Expenditure Amounts by FR Contact*

|                           | <b>No FR contact<br/>(N=286)</b> | <b>FR contact<br/>(N=21)</b> | <b>Difference</b> |
|---------------------------|----------------------------------|------------------------------|-------------------|
| Average Expenditure Total | -                                | -                            | -\$16.16          |
| Median Expenditure Total  | -                                | -                            | -\$48.63          |

CUs receiving a FR reminder contact had a slightly greater median number of entries than those that did not (33 and 29 respectively). They also had a slightly higher median weekly amount of reported expenditures, \$50 more. However, these comparisons did not control for CU characteristics due to small sample sizes.

**Expenditures by Frequency of Respondent Log-ins**

Web diary log-in patterns were analyzed among the 187 CUs that had matched paradata. Among the 120 CUs that had completed 2 weeks of diary entries, 53 percent had logged-in to the instrument on 6 or fewer days, and 47 percent logged-in on 7 or more days. In addition, whether or not CUs that frequently accessed the diary were reporting higher expenditure amounts than those who logged in only a few times during the reporting period<sup>8</sup> was examined.

*Expenditures among CUs with Above-Median Log-in Days (>6) versus Median or Below (<=6)  
(2-Week Diary-Completing CUs with Matched Paradata, CU=120)*

|                           | <b>Infrequent<br/>Log-in (N=63)</b> | <b>Frequent<br/>Log-in (N=57)</b> | <b>Difference</b> |
|---------------------------|-------------------------------------|-----------------------------------|-------------------|
| Average Expenditure Total | -                                   | -                                 | -\$509.66         |
| Median Expenditure Total  | -                                   | -                                 | -\$135.68         |

The data suggest that CUs that regularly log-in to enter expenditures over the course of the diary reporting week have higher expenditure reports than those that do not. The median expenditure total for CUs with above-median log-in days was \$130 higher than those logging-in less frequently.

<sup>8</sup> Ideally comparisons would involve controls for CU characteristics such as age, income and household size, but this was not possible here given the small sample sizes.

In addition to examining the relation between log-ins and expenditures, whether or not FR contacts had an impact on the daily frequency with which CUs logged-in to the instrument was studied. Diaries weeks were compared in which there had been a FR reminder contact with those in which there had not be any reminder contact.

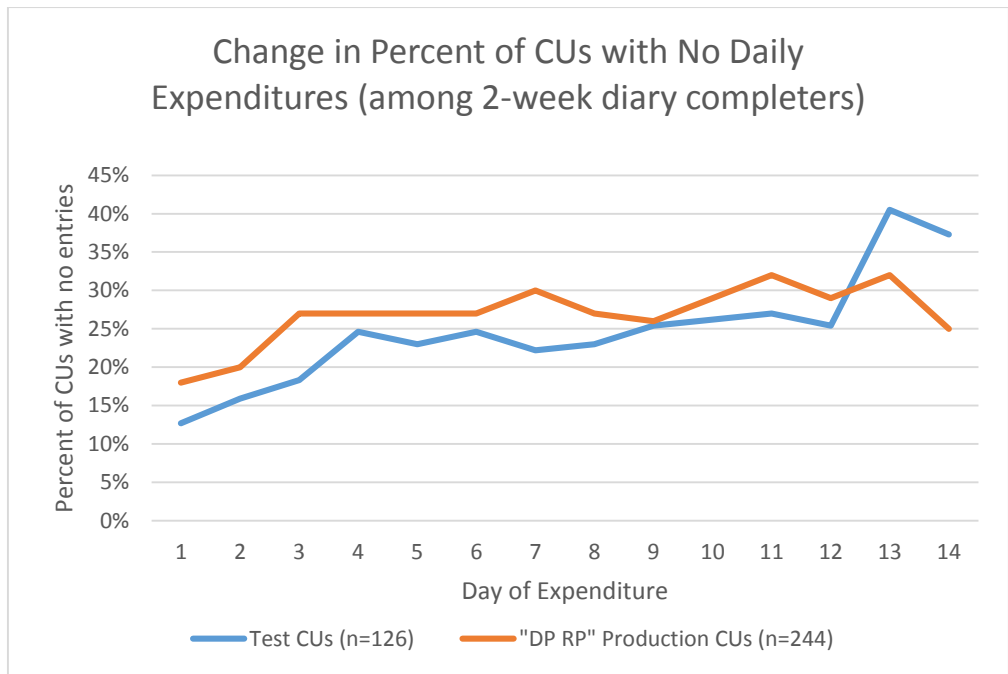
*Log-in Days per Week by FR Reminder Contact*

|                     | No FR contact<br>(N=286) | FR contact<br>(N=21) | Difference |
|---------------------|--------------------------|----------------------|------------|
| Average Log-in Days | 3.09                     | 3.05                 | 0.04       |
| Median Log-in Days  | 3                        | 3                    | 0          |

There was no notable difference in the number of log-in days by whether CUs received a reminder contact.

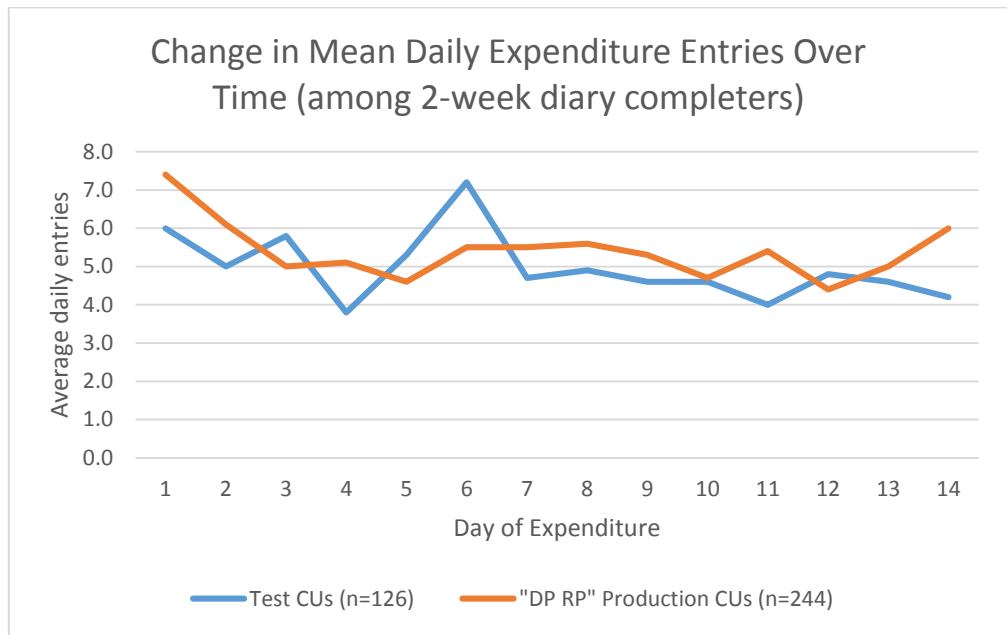
**Comparison of Daily Expenditures between Web and Restricted Production Diaries**

Expenditure data include an indication of the sequential day (1-7, or 8 if unspecified, for each week) on which an expenditure occurred. These data were used to calculate the percentage of respondents not having any expenditures per day of the week, among CUs recording two weeks of expenditures. This statistic was graphed below for the 126 Web diary CUs and the 244 double-placed RP CUs with no recall.



The graph shows that, despite the RP CUs having a higher number of entries on average than Web test CUs, a greater percentage of RP CUs had days with no expenditures recorded. The exception was on the last two days, in which 41 percent and 37 percent of Web CUs had no expenditures, compared to 32 percent and 25 percent RP CUs, respectively. This was consistent with FR debriefings indicating a drop-off in reporting among Web CUs at the end

of the reporting period. The average number of expenditures per day by reporting day for these groups was also examined, illustrated in the graph below.



Similar to the graph of the percent of cases with no expenditures per day, Web CUs had a comparable number of expenditures on each day as did double-placed RP CUs. It should be noted that RP CUs did have additional expenditures recorded on unspecified days which would not be reflected in the graph above (an average of 2.4 entries unassigned to a day in Week 1 diaries and 1.8 entries unassigned in Week 2).

**d. Non-Expenditure Results**

**Missing Data**

Among CUs with at least some entries reported upon pick-up (including partial recall CUs), there were lower item non-response rates among test CUs than among RP CUs. 8.5 percent of test CUs had at least one missing cost field (entries for which there was a description or a day reported), compared to 12.5 percent of RP CUs.

*CUs with any item non-response, by response field*

|             | <b>Test (n=200)</b> |                         | <b>Production (n=1,435)</b> |                         |
|-------------|---------------------|-------------------------|-----------------------------|-------------------------|
|             | <b>CUs</b>          | <b>Percent of total</b> | <b>CUs</b>                  | <b>Percent of total</b> |
| Cost        | 17                  | 8.5%                    | 180                         | 12.5%                   |
| Description | 0                   | 0.0%                    | 0                           | 0.0%                    |
| Day         | 1                   | 0.5%                    | 255                         | 17.8%                   |

The largest difference was for day of purchase, as 0.5 percent of test CUs had some missing day information, compared to 17.8 percent of RP CUs (since there was a higher RP percentage of missing day fields for ‘Food and Drinks for Home Consumption’, this suggests the difference was due to RP entries in the back-pages of the paper diary for which no day is specified). Otherwise, there were no notable differences in item non-response rates by

expenditure category.

*Item Non-Response at Day Field, by section*

|                                                   | <b>Test<br/>(n=200)</b> | <b>Production<br/>(n=1,435)</b> |
|---------------------------------------------------|-------------------------|---------------------------------|
| <b>Food and Drinks for Home Consumption</b>       |                         |                                 |
| Missing entries                                   | 32                      | 3,237                           |
| Total entries                                     | 5,947                   | 46,990                          |
| Percent                                           | 0.50%                   | 6.90%                           |
| <b>Food and Drinks Away from Home</b>             |                         |                                 |
| Missing entries                                   | 0                       | 18                              |
| Total entries                                     | 1,661                   | 13,593                          |
| Percent                                           | 0.00%                   | 0.10%                           |
| <b>Clothing, Shoes, Jewelry, and Accessories</b>  |                         |                                 |
| Missing entries                                   | 0                       | 5                               |
| Total entries                                     | 435                     | 3,141                           |
| Percent                                           | 0.00%                   | 0.20%                           |
| <b>All Other Products, Services, and Expenses</b> |                         |                                 |
| Missing entries                                   | 0                       | 131                             |
| Total entries                                     | 3,078                   | 24,910                          |
| Percent                                           | 0.00%                   | 0.50%                           |

**Extent of Partial Recall Conducted**

Among CUs with at least some diary entries, 20 percent of test cases had partial recall carried out compared to 16 percent of RP cases.

*Proportion of CUs having recorded entries with some recall carried out*

|                        | <b>Test<br/>Percent<br/>(n=200)</b> | <b>RP<br/>Percent<br/>(n=948)</b> | <b>Difference<br/>(Test-RP)</b> | <b>Change</b> |
|------------------------|-------------------------------------|-----------------------------------|---------------------------------|---------------|
| Percentage with recall | 20.0%                               | 16.1%                             | 3.9                             | 24.2%         |

**6. Paradata Analysis**

This section describes findings from a review of the web diary paradata file, with focus on access and usage. In some instances, the web diary paradata is coupled with FR recorded Contact History Instrument (CHI) data as well as unprocessed CE production data, matched at the case level. There were two key delays affecting the omission of a full paradata analysis section. First, the suite of paradata, provided by Census, was delivered in a format, without a corresponding data dictionary, that made analyzing the content problematic. In addition, there was a delay in

receiving the paradata draft narrative and tables leading to additional time being necessary for analysis and interpretation.

### a. Log-in Data

#### Successful Log-in Days

Among respondents who successfully logged into the instrument, approximately 17 percent of respondents logged in on one day and one day only compared to less than one percent of respondents who successfully logged into the instrument on all 14 days. The median number of successful respondent log-in days was four days.

#### *Total Number of Successful Log-in Days (Respondent Level)*

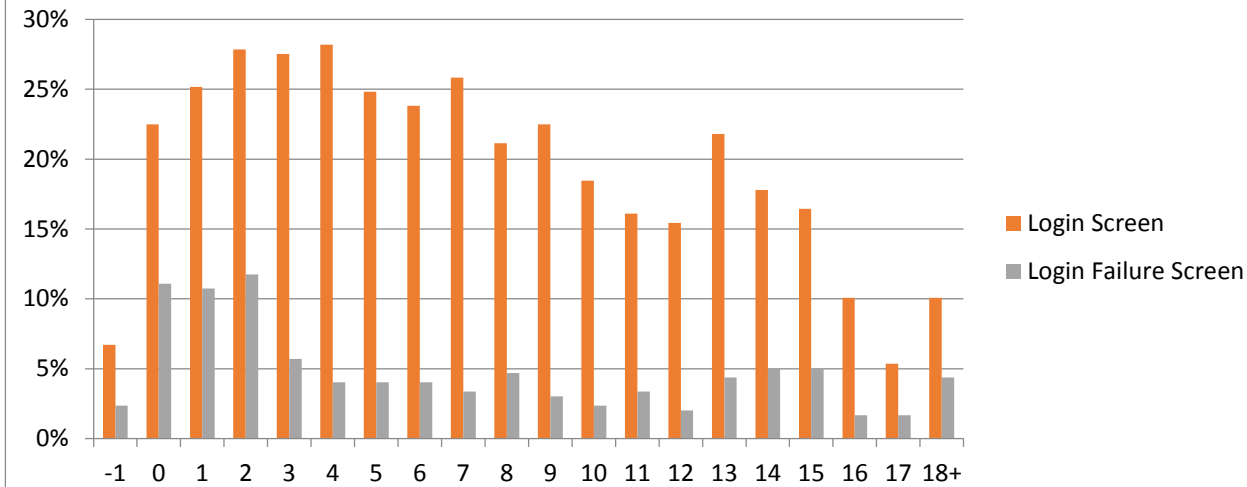
| <b>Successful Log-in Days</b> | <b>Frequency</b> | <b>Percent</b> | <b>Cumulative Percent</b> |
|-------------------------------|------------------|----------------|---------------------------|
| 1                             | 45               | 17.0%          | 17.0%                     |
| 2                             | 49               | 18.6%          | 35.6%                     |
| 3                             | 34               | 12.9%          | 48.5%                     |
| 4                             | 31               | 11.7%          | 60.2%                     |
| 5                             | 17               | 6.4%           | 66.7%                     |
| 6                             | 23               | 8.7%           | 75.4%                     |
| 7                             | 14               | 5.3%           | 80.7%                     |
| 8                             | 13               | 4.9%           | 85.6%                     |
| 9                             | 4                | 1.5%           | 87.1%                     |
| 10                            | 9                | 3.4%           | 90.5%                     |
| 11                            | 12               | 4.5%           | 95.1%                     |
| 12                            | 5                | 1.9%           | 97.0%                     |
| 13                            | 6                | 2.3%           | 99.2%                     |
| 14                            | 2                | 0.8%           | 100.0%                    |
| Total                         | 264              | 100.0%         |                           |

#### Respondent Log-ins by Day

Approximately 20 percent of respondents attempted to log into the web diary instrument on any given day during the collection period. The largest percentage of respondents successfully logged into the web diary instrument on days two, three, and four. Log -in rates steadily decline throughout most of the collection period with a noticeable increase at the end of the period, likely directly related to respondents logging in to enter expenditures before their diary is picked up by an FR.

The most log-in failures occur during the first three days of the collection period then falling to about five percent each day, with approximately ten percent of respondents failing to log-in on at least one attempt. Analysis of the paradata also shows that approximately ten percent of respondents attempt to log-in days after the collection period ends. These attempts likely occur during pickup as respondents and FRs attempt to log into the instrument to enter expenditure forgotten during the diary collection period.

## Percent of Respondents Who Reach the Login or Login Failure Screens: Difference in their Attempt Minus the Placement Date



### b. Access Device and Operating System

The web diary paradata file included information on the operating system, device, and browser the respondents used while accessing the web diary. Of the 1,576 instances, virtually all respondents log-ins to the web diary instrument (96 percent, n=1,513) were through a desktop computer. Furthermore, a Windows operating system was detected in nearly 9 out of 10 desktop computer log-ins (86 percent, n=1,297), while the remaining 216 desktops log-ins were through a “Mac OS X” operating system.

#### Access Device and Operating System

| OS       | Device     | Percent | Freq | Android Webkit | Mobile Safari | Safari | Chrome | Firefox | Internet Explorer | AOL Browser |
|----------|------------|---------|------|----------------|---------------|--------|--------|---------|-------------------|-------------|
| Android  | SmartPhone | 0.3%    | 5    | 4              |               |        | 1      |         |                   |             |
|          | Tablet     | 0.6%    | 9    | 9              |               |        |        |         |                   |             |
| iOS      | iPhone     | 0.4%    | 7    |                | 7             |        |        |         |                   |             |
|          | iPad       | 2.7%    | 42   |                | 42            |        |        |         |                   |             |
| Mac OS X | Desktop    | 13.7%   | 216  |                |               | 141    | 30     | 45      |                   |             |
| Windows  |            | 82.3%   | 1297 |                |               | 1      | 305    | 259     | 706               | 26          |
|          |            | TOTAL   | 1576 | 13             | 49            | 142    | 336    | 304     | 706               | 26          |
|          |            |         |      | 0.8%           | 3.1%          | 9.0%   | 21.3%  | 19.3%   | 44.8%             | 1.6%        |

Three browsers were the definitive selection for the 1,297 Windows operating system desktop instances: Internet Explorer (54 percent, n=706), Chrome (24 percent, n=304), and Firefox (20 percent, n=259). The remaining 2 percent (n=26) used an “AOL Browser” when accessing the web diary. Three selections were detected with the

216 Mac OS X operating system desktops: Safari (65 percent, n=141), Firefox (21 percent, n=45), and Chrome (14 percent, n=30).

The remaining 3 percent (n=63) of operating system detections were predominantly iOS (n=49) with a small handful (n=13) detected as Android. Within this small group of 63, the tablets (iPad n=42, Android tablet n=9) far outnumbered the Smartphones (iPhone n=7, Android Smartphone n=5).

## 7. Operational Issues

After the web survey was completed in April 2013, respondents were given the opportunity enter comments on different aspects of the web diary experience through debriefing questions; however, not all respondents chose to answer all of the debriefing questions. In addition, selected FRs took part in several debriefing sessions that took place while the test was being fielded as well as after its completion. Given the variety of experiences and opinions of the respondents and FRs, it is not surprising that many comments and suggestions conflict which underscores the importance of a flexible instrument and survey experience.

It is also important to note that some respondent and FR suggestions may focus on functionalities that already exist in the web diary instrument and protocols. Rather than dismissing these opinions, they should be highlighted as they indicate that some respondents and FRs did not completely comprehend some aspects of the instrument functionality and/or test protocols. This may suggest that additional training, revised survey materials, and/or an overhaul of the placement procedures is necessary.

A qualitative analysis was conducted on open-ended comments/questions and the main points summarized. For all open-ended questions, the complete set of open-ended comments is available upon request.

### a. Respondent Experience

#### Logging-In

Respondents were asked with what ease they were able to log-in to the web diary instrument. Almost 70 percent reported that logging into the instrument was straightforward and uncomplicated remarking that it was “easy to access” (WD\_LIKE). However, one in five respondents reported that they found logging in to be at least difficult if not very difficult. Some respondents noted that they “couldn’t get into the web diar[y]” (WD\_LIKE) and that they were “unable to log in” (WD\_LIKE).



*How easy or difficult was it to log in? (WD\_DFLOG)*

|                            | Percent<br>(n= 296) |
|----------------------------|---------------------|
| Very easy                  | 38.5%               |
| Easy                       | 28.0%               |
| Neither easy nor difficult | 13.2%               |
| Difficult                  | 8.1%                |
| Very difficult             | 12.2%               |

During the field test two operational issues were discovered that made logging into the instrument more difficult. First, the FRs were instructed to transcribe the username and password from the computer-assisted personal interviewing (CAPI) instrument to the user guide, which was then distributed to the respondent; however, the user guide was printed on glossy paper, which was difficult to write on (WD\_USGDE). This issue led to illegible usernames and password marring the log-in process. In addition, the default font in the CAPI instrument made it difficult for the FRs to distinguish between “1,” “l,” “I,” “O,” and “0” leading to FRs incorrectly transcribe usernames and passwords. FRs noted that up to 30 percent of the passwords did not work correctly. These issues were compounded by the lack of a dedicated helpdesk with hours extending into the evening. FRs noted that the survey materials were not always sufficient in preventing respondent log-in issues and they recommended that the placement protocol be modified to allow FRs to demonstrate the log-in process to the respondent.

*Recommendations.* Future iterations of the web diary instrument should allow for the respondent to change their username and password to facilitate logging into the instrument. In addition, since a default username and password must be created to tie the respondent to the control number, “1,” “l,” “I,” “O,” and “0” should not be allowed as part of the default usernames and passwords. Finally, the respondent user guide should be printed on paper that is not as glossy to facilitate transcription of default usernames and passwords and reduce inaccurate transcription.

### **Data Entry**

Similarly, respondents were asked to rate how difficult it was for them to record expenses in the web diary instrument. Approximately 65 percent responded it was easy to very easy to record expenses in the instrument with slightly over 15 percent responding that it was difficult to very difficult. While some respondents noted that the instrument was “easy to use” and “faster and more convenient than paper” (WD\_LIKE), others “didn’t like it” noting that “it was to[o] time consuming” (WD\_LIKE).

*How easy or difficult was it to record your expenses? (WD\_DFREC)*

|                            | Percent<br>(n = 297) |
|----------------------------|----------------------|
| Very easy                  | 30.0%                |
| Easy                       | 33.6%                |
| Neither easy nor difficult | 19.5%                |
| Difficult                  | 5.4%                 |
| Very difficult             | 11.5%                |

*Recommendations.* Although no operational or technical issues pertaining to the recording of expenses were reported by respondents, during the debriefing session, FRs noted that respondents were frustrated by having to repeatedly enter the date after each purchase and had trouble navigating through the different tabs. Currently, the web diary instrument allows for expenditures to be entered for any day during the collection period and defaults to the first day of period. Respondents are asked to select a day through a dropdown dialog box for each expense they enter. Repetitive entering of the same information into the survey is not user friendly design. Future iterations of the web diary instrument should have the date field default to the current day; this will make entering numerous items less difficult and, hopefully, spur same day diary maintenance. In addition, the survey materials should be improved as well as FR training in order to bolster respondent understanding of the instrument.

**Record Usage**

When respondents were asked whether or not they used any records or online statements when recording expenses, almost half responded that they did. FRs stated that this process would have been facilitated if the test protocol included guidance for the collection of records and if user guide had included a tab to facilitate the collection of records similar to what exists in the paper diary.

*Did you use any records or online statements when recording your expenses? (WD\_RCRD)*

|     | Percent<br>(n = 311) |
|-----|----------------------|
| Yes | 44.7%                |
| No  | 55.3%                |

*Recommendations.* Future research should include changes to diary protocol to allow for FR collection and input of records, particularly receipts with numerous items. This will help to determine the appropriate balance between respondent and FR burden as well as note in improvements in data quality. In addition survey materials should be altered to include language detailing the record collection protocol and a flap or tab to facilitate record collection.

When asked how the process of recording expenses could be improved (WD\_IMPRV), many respondents noted the web diary was “pretty simple already.” However, some thought the instrument was “too detailed,” needed “fewer categories,” and desired “more clarity in what to itemize.” FRs also noted lost expenditures due to the respondent being locked out of the instrument after the final day of the collection period. Certain design features of the web instrument as well as the test protocol are required based on CE data collections rules; however, streamlined design features, such as defaulting to the current day, and allowing for the entry into the instrument of in-scope expenditures after the final day of collection will improve the user experience.

## Security

Data security is an overriding concerns for many internet users; subsequently, it is paramount that respondents are provided with an instrument that both appears secure and is secure. Accordingly, respondents were asked about how secure they felt their data was when entering expenditures into the instrument.

*When using the web diary to record your expenses, how secure do you feel your data is? (WD\_SCRE)*

|                             | Percent<br>(n = 291) |
|-----------------------------|----------------------|
| Completely secure           | 48.5%                |
| Somewhat secure             | 32.6%                |
| Neither secure nor insecure | 13.4%                |
| Somewhat insecure           | 3.1%                 |
| Completely insecure         | 2.4%                 |

Overwhelmingly, 80 percent of respondents felt the data collected by the web diary instrument was secure while slightly over five percent disagree, some stating they “didn’t trust it” (WD\_LIKE). With an average age of 48 years, respondents who felt the data collected by the web diary instrument was secure were slightly younger than those who believed the data to be insecure (55 years).

As part of the test protocol, respondents were assigned a username and password by Census. Passwords assigned met Census security standards for length and composition. In addition, respondents were not given the option to change their username or password from the ones that they were initially assigned. In addition to making logging into the instrument more difficult, assigning usernames and passwords, especially through the protocol of transcribing them to a user guide, can give the appearance of an absence of robust data security. Respondents reported “having problems with the password” (WD\_IMPRV), asked to be able to “shorten the passwords” (WD\_IMPRV), and requested to be able to “customize passwords” (WD\_IMPRV). During debriefing, FRs reiterated this sentiment requesting that respondents be able to customize their usernames and passwords.

*Recommendations.* It is recommended that future iterations of the web diary continue to focus on data security by implementing the most recent security protocols, displaying security assurances in a prominent place within the instrument, and educating FRs on the most relevant data security concerns. In addition, respondents should be able to customize their usernames and passwords as long as they meet Census security requirements. Customizable usernames and passwords will make it more likely that respondents will log into the instrument, log-in from multiple locations, and should bolster the perception of data security by allowing the respondent more control of their log-in information.

## b. Respondent Device Information

### Browser

It is critical that the web diary instrument perform optimally on a myriad of different browser and operating platforms. Although many browsers and operating systems are coded using the same programming languages (C++, SpiderMonkey, JavaScriptCore, etc.), web pages can render differently depending on the program a respondent is using. Subsequently, the instrument programming must be flexible enough to account for these differences. In addition, the proliferation of mobile and tablet only respondents reinforces the need for a flexible instrument.

Respondents were asked which browser they typically use at home which resulted in a distribution similar to the national browser market share<sup>9</sup>. By tracking closely with the national browser market shares, the current strategy of meeting the requirements of Explorer, Safari, Firefox, and Chrome users is an appropriate strategy. Of those respondents who responded “other,” they had a chance to specify which browser they used at home and all responses were not web actual browsers, responses ranged from “Google” to “Yahoo” to “AOL.”

*Which browser do you typically use at home? (WD\_BRWSR)*

|                | Percent<br>(n = 292) |
|----------------|----------------------|
| Explorer       | 53.4%                |
| Firefox        | 15.1%                |
| Safari         | 7.8%                 |
| Chrome         | 18.2%                |
| Other, specify | 5.5%                 |

### Operating System

In addition, respondents were asked what operating system their home computer runs on. Overwhelmingly, respondents reported using Windows based operating systems which mirrors national averages during the same

<sup>9</sup> <http://www.netmarketshare.com/browser-market-share.aspx?qprid=1>

time period<sup>10</sup>. By tracking closely with the national operating system market shares, the current strategy of meeting the requirements of Windows OS and Mac OS users is an appropriate strategy.

*Recommendations.* Future research will need to be completed to determine what operating systems were categorized as “Other.” Given the timing of data collection, it is most likely that the “Other” field captures mobile OS, Windows 8, and Linux users.

What operating system does your home computer run on? (WD\_OPSYS)

|               | Percent<br>(n = 283) |
|---------------|----------------------|
| Mac OS        | 14.1%                |
| Windows 7     | 11.3%                |
| Windows XP    | 33.2%                |
| Windows Vista | 18.7%                |
| Other         | 22.7%                |

Subsequently, optimization for these browsers and operating systems should remain a priority, but be flexible enough to adapt an ever-changing environment. Because, as the avenues people use to access the internet evolves, additional research into how mobile browsers and operating systems impact respondent experience is paramount to understanding how the user defines “user-friendly.”

### c. Field Representative Experience

#### Cooperation

An FR debriefing question collected information from FRs on how cooperative web diary respondents were in terms of keeping a diary when compared to than paper diary respondents. Over 40 percent of FRs reported that the web diary respondents were more cooperative with only 13 percent noting that the web diary respondents were less cooperative.

---

<sup>10</sup> <http://www.netmarketshare.com/operating-system-market-share.aspx?qprid=10&qpcustomd=0&qpsp=168&qpnp=2&qptimeframe=M>

*Was this web diary respondent more cooperative, about the same, or less cooperative than paper diary respondent in terms of keeping the diary? (FR\_COOP)*

|                  | Percent<br>(n = 327) |
|------------------|----------------------|
| More cooperative | 40.7%                |
| About the same   | 46.5%                |
| Less cooperative | 12.8%                |

On average, the age of the more cooperative respondent was older (49 years) than that of the less cooperative respondents (45 years). These findings are counterintuitive to the impression that younger respondents are more open to completing surveys on the web; however, too much weight should not be assigned to these results since four years is not much of an age difference and those who were less cooperative may have had an objection to participating in the survey in general rather than participating in an online survey. Yet, these findings are promising such that multimodal survey offerings should allow for higher placement rates by offering the mode that is the best fit for the respondent.

*Recommendations.* Future research should include multiple survey modes in order to determine which modal offering are optimal and the most effective way to design each survey mode.

### **CAPI**

FR experience with the CAPI and web diary instruments varied across the FRs that participated in the web diary survey. Although the overall experiences were positive, there were minor issues regarding the FR's use of the CAPI instrument that could be improved.

*Recommendations.* FRs noted that if going through the instrument quickly, it was easy to miss the password screen and then difficult to return to it. If implemented, before leaving the password screen in the CAPI instrument, a dialog box should open prompting the FR that they are about to leave the password assignment screen.

### **Receipts and Recall**

FRs also reported that the receipt and recall procedures, collected through a separate tab in the CAPI instrument and collated with the web diary data at the National Processing Center (NPC), was not intuitive. Specifically, it was difficult to prompt the respondent to recall expenditures without having access to the respondent's web diary because, often, the respondent could not remember what they entered.

*Recommendations.* Future feasibility tests should assess different protocols for allowing the FR access to a summary of the respondent's web diary entries. These protocols could include allowing the FR to ask the

respondent to log into the instrument on the respondent's computer or have a special tab in the CAPI instrument that showed a summary of the respondent's entries.

### Follow-up Contacts

FRs also suggested sending reminder emails and/or text messages to respondents on regular intervals as well as when a designated time interval has passed without the respondent logging into the instrument. Reminder messages are important as they could help build rapport between the FR and the respondent and because there were a sizable number of cases, at least 21 percent of all web diary cases, that were successfully placed, but in which no expenditures were reported.

*Recommendations.* One way to accomplish this is by running daily analysis on the paradata to determine which respondents have not logged into the instrument as well as keeping a log of how many days have passed since the last successful logon by the respondent. After a specific time period has lapsed, FRs will be notified to contact the respondent to remind them to enter any expenditures.

### Training

Finally, for the web diary test, FRs were trained using a self-study training package. There were a myriad of issues with the self-study including the dissemination of an incorrect link to the test instrument, insufficient time allotted towards completing the self-study, difficulty accessing the materials, and the content inadequately preparing the FRs for placing the web diary. Compounding the training issues, the web diary placement and pick-up protocols were a wide departure from the paper diary placement and pick-up procedures. This led to specific issues regarding the dissemination of factual information regarding the web diary during placement and made receipt and recall procedures more difficult.

*Recommendations.* Subsequently, classroom training is necessary for the successful implementation of any feasibility test that features enormous departures from past protocols.

## 8. Limitations

The successful implementation and analysis of the test protocol was limited by a number of intervening factors:

1. ***Diary Review for Missed Expenditures.*** A limitation of this feasibility test is that the team was unable to add a question about missing days to the ANYRECAL screen in the CAPI instrument. Instead of a question (Did you or someone in your household make any purchases that you did not enter into the diary; maybe because you were too busy, you forgot, or because you didn't think it should be included?), an FR reminder was placed on the ANYRECAL screen due to programming time constraints. Throughout the feasibility test there were a considerable number of successfully placed web diaries with no expenditures reported through the web diary instrument. In addition, over a fifth of the completed web diaries contained

only recalled expenditures. Subsequently, relying merely on an FR reminder and not explicitly asking the respondent for missed expenditures, some expenditure data may not have been collected.

2. ***Monitoring of Log-in Information for Tailored Follow-up.*** Part of the web diary protocol required the daily monitoring of web diary paradata by CE in order to determine if there were respondents who went more than three days without logging into the web diary instrument. Any respondents flagged for failing to log into the instrument would receive a telephone call from their FR asking them if they needed any assistance with the web diary. However, Census was unable to provide CE with the necessary paradata due to security concerns until two-and-half months into the test and, subsequently, targeted follow-up calls were not placed. Targeted follow-up calls to the respondent may have led to an increase in the number of expenditures reported and a decrease in the number of total recall cases.
3. ***Operational Training.*** Furthermore, the rollout of the self-study training package had myriad of issues that hampered the success of the web diary test. In addition to Census field staff indicating that the training materials did not adequately prepare FRs for web diary data collection, some FRs did not receive the training materials prior to the start of data collection. Additionally, aspects of the training materials, such as the URL to the web diary instrument, did not function as intended. Issues such as these led to field staff being inadequately prepared for the nuances of placing the web diary and may have directly led to reduced response rates.
4. ***Training to Address and Mitigate Respondent Concerns.*** Additionally, during the mid-test debriefing, FRs indicated that some respondents refused to participate because of online privacy concerns. In addition, evidence in the CAPI case notes suggests that some respondents refused due to an aversion to the Internet and/or computers. Although this issue factored into higher-than-normal screen out rates, respondents' opposition to participating in an online survey should be mitigated by offering multiple survey modes, including more traditional modes such as a paper diary.
5. ***Resource Staffing.*** Furthermore, Census regional offices were not able to take advantage of a full allotment CE Diary field staff. Logistical issues such as the regional office realignment and web diary data collection running simultaneously with the training and data collection efforts of the March Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC). Subsequently, resources and time that would normally have been dedicated to web diary interviewing and follow-up may have been shifted to CPS ASEC.
6. ***Paradata Limitations.*** There were a number of limitations inherent in the suite of paradata that the Centurion development software provides. The Application Services Division (ASD) did not provide a



data dictionary to be used in understanding each column in the paradata. Subsequently, time slotted for analysis of the paradata was used in defining and reformatting the columns within the paradata. For example, timestamps associated with respondent actions within the instrument were sent in Epoch time. A SAS program had to be coded to change these timestamps into SAS time and then into astronomical time within the Gregorian calendar. General readability issues such as these as well as time allotment at Census have led to a delay in the release of the paradata analysis.

7. ***Web Diary and Production Diary Data Reprocessing.*** Additionally, over the course of the production cycle, web diary data and the corresponding production diary data had to be resent and reprocessed due to errors found during the review of the data. For example, the March web diary data had to be resent because the datasets contained duplicate CENSIDs. In addition, production diary Month 3 datasets contained data from Month 1 leading to another resend of the data and more reprocessing. These issues combined with the furlough of Federal employees during part of the month of October 2013, led to delays in analyzing and comparing the web diary data, specifically comparing expenditures from the web diary data to a restricted production sample.

## 9. Recommendations

The Web Diary Analysis Team will schedule meetings with the Individual Diary Feasibility Team and CE managers to discuss the following action items, based on the web diary test:

### *Low Response Rates*

- a) [Create flexible, but secure, username and password requirements](#)
- b) [Restrict default usernames and passwords to specific characters](#)
- c) [Print User Guide on non-glossy paper to facilitate username/password transcription](#)
- d) [Change the test protocol to allow for FR collection and input of records, particularly receipts with numerous items](#)

### *Lower Expenditure Amounts*

- e) [Increase FR follow-up contacts and record information, through the CHI instrument, to assess disposition/outcome of follow-up contact](#)

### *Higher Rate of Total Recall*

- f) [Allow entry, into the instrument, of in-scope expenditures after the final day of collection](#)
- g) [Assess different protocols for allowing the FR access to a summary of the respondent's web diary entries](#)
- h) [Set Date field to default to current day](#)
- i) [Run daily analysis on the paradata to determine which respondents have not logged into the instrument as well as keeping a log of how many days have passed since the last successful logon by the respondent and,](#)

after a specific time period has lapsed, FRs will be notified to contact the respondent to remind them to enter any expenditures

*Low Placement Rates*

- j) Create classroom training for any feasibility test that features enormous departures from past protocols
- k) Implement the most recent security protocols, displaying security assurances in a prominent place within the instrument, and educate FRs on the most relevant data security concerns

*Higher Ineligible Rates*

- l) Complete future research to determine what operating systems were categorized as “Other”
- m) Include multiple survey modes in order to determine which modal offering is optimal and the most effective

*Other*

- n) Employ a dialog box that opens prompting the FR that they are about to leave the password assignment screen
- o) Paradata should be formatted in a manner that is easily analyzed and readable

## 10. References

1. Couper, M. (2008). *Designing Effective Web Surveys*. New York: Cambridge University Press.