

Large Scale Feasibility Test Preliminary Report

April 20, 2021

Daniel Dorfman, Laura Erhard, Graham Jones, Brandon Kopp, Parvati
Krishnamurty, Brett McBride

Consumer Expenditure Surveys Program Report Series

1 EXECUTIVE SUMMARY

The Large Scale Feasibility Test of the Online Diary (LSF) was administered between October 2019 and March 2020 with a close out of March cases in April 2020. The LSF included both online and paper diaries. The data used in this report are unprocessed and have not gone through post-processing at BLS (including editing, allocation, imputation, and creation of weights). Therefore the expenditure findings presented in this report are preliminary, and more detailed results will be provided in the forthcoming final report. The preliminary report covers sample performance, data collection issues, debriefing results, preliminary findings for expenditure estimates and demographics based on unprocessed data, and recommendations for implementation based on these results. The final report will cover final response rates, demographics, and expenditure counts and amounts. Overall, the test was fielded successfully and led to actionable results; however, a detailed recommendation for implementation into the Consumer Expenditure Survey Production Survey cannot be provided until the analysis of the processed data is completed. The executive summary highlights the major findings based on the preliminary analysis of the LSF data. Further detail on the findings and recommendations can be found in the report.

Conditional on the final report findings, the team recommends implementation after the following items are addressed:

Figure 1.1. Preliminary recommendations

1. Do not use incentives or pre-notification postcards
2. Provide additional training for FRs on how to use the online diary
3. Improve access to the help desk and training for help desk staff
4. If planning to continue screening respondents to place online diaries, conduct additional research on wording of internet access question and other ways to screen ideal online respondents
5. Develop explicit placement procedures for paper diaries for online eligible respondents
6. Work with Census to identify a way to better capture recalled entries, particularly from receipts saved by the respondent for online diary respondents
7. Emphasize contacts with respondents via midweek calls and investigate options to text respondents
8. Test new online diary features to improve data quality such as prompts and autosuggest features
9. Address respondent's privacy concerns in advance letters and FR training
10. Improve access and usability of YouTube help videos and user guide
11. Increase use of expenditure summary report by FRs and have a clear approach for dealing with respondent privacy concerns
12. Investigate how to better elicit expenditures for other CU members in the diary

Response rates

Using the AAPOR RR2 definition, the overall LSF collection response rate at the case level¹ for October 2019 through February 2020 was 47.3 percent which was slightly lower than Production (51.7 percent) over the same five months. There were differences between the LSF and Production in the proportion of refusals and noncontacts (LSF was lower than Production) as well as in other Type A nonresponse and incompletes (LSF was lower than Production). Type A incompletes are defined as having a successful placement, but did not provide any diaries. Of the LSF sample units initially placed with an online diary, over half did not complete the survey.

Online diary placement

- Overall, of all the completed cases, 43.8 percent of respondents were eligible and placed with an online diary, while 8.6 percent of all respondents were eligible but chose to use a paper diary. The remaining 47.5 percent of respondents were not eligible for an online diary based on the recorded internet access and ability answers and were placed with a paper diary.
- Reference persons who completed the online diary were generally younger, had higher levels of education, were from a larger CU size, and were more likely to be homeowners compared to those who were placed with a paper diary.

Sample differences

- There were very slight differences in the demographic characteristics of respondents between the LSF and Production, with a slightly higher percentages of Hispanic respondents and homeowners in LSF.

Incentives and postcard experiment

- The LSF included an incentive experiment with half of the sample receiving an advance postcard and a \$5 cash incentive with the advance letter, a quarter of the sample receiving a postcard and no cash incentive in the advance letter, and a quarter of the sample only receiving an advance letter. All advance letters were sent by Priority Mail. Participation outcomes were found to not vary significantly between the test and control groups.
- Further evidence for the lack of impact of the cash incentive comes from the respondent debriefing question on incentives. About 32 percent of the 216 people who responded to the question said the incentive made them more willing to complete the diary, 1.4 percent that said it made them less willing to complete the diary and over 65 percent of respondents said the \$5 bill had no effect on their decision to participate.

¹ The collection response rates are calculated at the level of each consumer unit or case. Final CE response rates are calculated considering each week of the diary as a separate, independent interview.

Contact attempts

- The average number of contact attempts for the LSF was higher compared to Production for complete interviews (5.9 contact attempts per case vs. 5.0 attempts, respectively). A possible explanation is that LSF respondents may have required more communication with FRs due to the introduction of the online diary and any complications that may have arisen as a result. LSF FRs required more total contact attempts on average in order to successfully place a diary (2.9 contact attempts vs. 2.0 Production contact attempts, respectively). Both LSF and Production required the same average number of in-person contact attempts overall (3.6 contact attempts). Incentives appear to have had a minimal and non-significant impact on the number of contact attempts required to reach respondents.
- There was a negative relationship between total number of contact attempts and completion. A respondent receiving more total contact attempts overall was 6.3 percent less likely to finish the survey.

Expenditure counts

- Overall, the median weekly counts of diary entries for Production consumer units (CUs) were higher than for LSF CUs (27 median entries for Production CUs and 20 median entries for LSF CUs), suggesting higher quality for production CUs. The difference was largely comprised of a higher number of entries in the ‘Food and Drinks for Home Consumption’ (FDB) section (15 median entries for Production CUs and 9 median entries for LSF CUs). This could potentially be explained by a decreased collection of items via grocery receipts in the LSF, where only about 6 percent of CUs had any recalled CAPI entries. For grocery expenses, respondents sometimes give the FR a receipt to enter expenses at pickup. For online diary cases, FRs had to close out the pickup interview instrument and could not enter expenses from receipts afterwards as they could do in the paper diaries and in production. Production data do not identify recalled entries.
- Within the LSF group, online diarists provided an average of 23.9 entries per diary week, compared to 28.9 entries on average for LSF paper diarists. These numbers exclude diary weeks without expenditures². The differences were statistically significant.

Expenditure amounts

- Production diarists recorded a higher total amount of expenditures each week than did LSF diarists. The \$438 median Production amount was \$45 higher than for LSF, and the

² We will reanalyze the data after the minimal expenditure edit process is completed to see if diaries are excluded because of the edit.

average totals were \$139.34 higher, representing a significant difference. Differences in median expenditures ranged from only \$9 in November to \$86 in January.

- Within the LSF group, average amounts for online diarists were higher than those for paper diarists, with median amounts of \$469.30 and \$331.05 respectively. Further analysis will be conducted after data are processed to control for the response behavior and sample characteristics that underlie these mode differences.
- The drop-off in entries (number of expenditure entries in week 1 over number of expenditure entries reported in both weeks) for LSF CUs and Production CUs was similar – 53.3 percent and 53.6 percent, respectively.

Recall

- In the LSF, recalled entries were entered by the FR in an additional tab in the CAPI instrument in addition to answering question on whether there was recall during the pickup interview. There were inconsistencies in information on recall from these two sources.
- We classified “total recall” diaries as diaries with no expenditure entries at pickup, but with entries added during pickup with or without a majority of items entered using receipts.
- The rate of total recall with or without receipts in the LSF were similar to Production with approximately 9 percent of diaries being classified as having total recall and 1 percent of diaries being classified as having total recall without receipts.
- Looking at recall by mode, 21.4 percent of LSF cases had FRs reporting any recall, which was evenly distributed between online and paper diary modes. About 7 percent cases had at least one CAPI recall entry, which were disproportionately online. This supports the idea that FRs did not use the recall tab as much for paper diaries and may have written recalled expenses into the paper diaries.
- For respondents who had any CAPI entries, 71 percent of entries, on average, were provided by recall. Comparing diaries with or without recall, the average number of entries are comparable, ranging between 45 and 57. However, those respondents that only provided total recall via CAPI had only 24 entries on average.

Item nonresponse

- Differences in item nonresponse between LSF and Production were minimal for the cost and item description fields, suggesting LSF diarists were providing comprehensive data about their expenditures.
- For the cost field, LSF diaries had a smaller rate of records with missing cost (0.42 percent) compared to Production (0.51 percent)– a significant, though minor, difference.

- For item description, about 0.08 percent of descriptions contained item nonresponse for Production diary entries, compared to a rate of 0.32 percent for LSF which is also a significant difference.
- For the outlet name (business) field, excluding the clothing category for which there was no column in the paper diary, there was 7.2 percent item nonresponse (of all outlet entries) in the LSF with the majority of the item nonresponse being blank entries. The item nonresponse in outlet name for Production diaries measured in a similar way was slightly higher, at 9.1 percent.

Respondent debriefing

During the pickup interview respondents answered some debriefing questions about their experience.

- A little over half of the respondents (55 percent) that answered the question reported at least some burden, with no significant mode differences.
- About 36 percent reported at least a little difficulty with the diary-- paper diary were more likely to report difficulty than those that kept an online diary (37 percent versus 28 percent) and mode differences were significant;
- Forty one percent of respondents reported that the information reported was at least a little sensitive, with no mode differences.
- About 71 percent of respondents reported entering expenses on behalf of the other CU members, but there were differences in responses by mode of diary. Online diary respondents were more likely to enter expense on behalf of other CU members compared to paper diary respondents (79 percent compared to 63 percent). Paper diary respondents were more likely to have other CU members enter their own expenses than online diary respondents (12 percent compared to 4 percent). Of concern, 8 percent overall (43 respondents) reported that expenses were not entered for other CU members, and the result was higher for paper diary respondents (11 percent) compared to online diary respondents (6 percent). This finding points to underreporting of expenditures in diaries, the extent of which would depend on how much other CU members are spending.
- Respondents were asked if spending behaviors changed as a result of keeping the diary. The majority of respondents (74-77 percent) said they did not change their behavior. Of those that did change behavior, most indicated that this was due to different CU needs or other reasons and only about 10 percent of responses noted that their habits changed because the diary made them more aware of spending and just under 3 percent of responses changed because of the process required for entering items.

Online diary usability, technical issues, and help desk use

- Based on respondent debriefing questions, between 3 and 9 percent of online diary users found online diary keeping tasks at least somewhat difficult.
- Some respondents had difficulty logging in at some point during the survey and a few never managed to log in successfully. Of the 384 CUs with paradata from the online diary, 12 CUs attempted to login but were never successful during the diary period. The remaining 372 CUs successfully logged in at least once. The first attempt at logging in was mostly successful (78.6 percent of the time) and was supposed to happen at the placement interview according to interview protocols.
- The average number of successful logins per respondent was 6.3.
- Despite having the ability to customize their username and password in the online diary, very few respondents (about 1 percent and 5 percent respectively) attempted to do so. This could be potentially explained by many browsers allowing respondents to save their login information.
- Only 120 contacts were received by the help desk during the survey period. It was unknown how many calls did not go through to the help desk, which was a problem reported in FR debriefing sessions.
- The top reason for calling the help desk was for login issues which were overwhelmingly related to respondents either getting locked out of the online diary or having issues entering passwords.

Operational issues including midweek contacts, pickup and drop off protocols

- LSF interview protocols included midweek calls on day 3 and 8 of the diary keeping period to remind respondents to enter expenses in the diary. There is no precise way to measure midweek calls, however, based on outgoing calls from the FR during this period, about 33 percent of LSF completes received successful midweek calls.
- Midweek calls did not significantly impact the login behavior of online respondents. The average number of logins for those who received a contact attempt was almost the same as those who did not. Respondents who received a contact attempts had slightly more diary entries on average than those who did not, but this difference was not statistically significant.
- Based on responses to the FR debriefing questions, FRs reported logging in with the respondent as instructed during the placement interview in about 67 percent of cases.
- Generally, FRs reported very positive responses on placement and pickup procedures as well as for diary quality. However, about 28 percent reported somewhat low or low data quality which is of concern. The main reasons for low data quality reported were that the respondent was disinclined, disinterested, busy, had privacy concerns, or had physical or mental limitations.
- The expenditure summary report, which was a tool that allowed FRs to monitor respondent logins and expenditure entries in the online diary was only used in about 72 percent of online diary cases based a FR debriefing question.

Materials:

- Based on information from respondents, the user guide was found to be effective as delivered. According to a respondent debriefing question, 62 percent of online diary respondents found it useful or very useful, most of the remaining respondents reported not needing to use it.
- There were three YouTube videos which were provided as an additional resource for respondents and FRs. There were a total of 2,255 views for all three YouTube videos as of the end of March 2020, 444 views were during the diary period and 1,811 views were before the diary period. Most of the views appear to have taken place during training by Census/BLS staff, so this is a resource which was useful for training but underutilized by respondents.

2 OVERVIEW

2.1 BACKGROUND

In 2013, the Bureau of Labor Statistics (BLS) Consumer Expenditure Survey (CE) approved a redesign plan with the objective to improve the quality of the survey estimates through a verifiable reduction in measurement error. BLS will be realizing the CE survey redesign through a phased implementation of key design elements. This involves implementing an online, CU-level diary into CE Diary Survey (CED) production in January 2022 (pending results and recommendations from the Large Scale Feasibility test described here and in a forthcoming report) and implementing a streamlined questionnaire and redesigned interviewing instrument into CE Interview Survey production April 2023.

In order to assess potential fielding issues and evaluate data quality of online diaries, the CE program completed the Large Scale Feasibility Test of the Online Diary (LSF). The LSF was built on information learned from prior online diary tests and was planned to have sufficient sample to make statistical inferences. An online diary was developed by Centurion staff at Census based on requirements from an online diary designed by Westat³.

³ The Westat report is available at https://www.bls.gov/cex/research_papers/pdf/ce-online-diary-usability-testing.pdf

2.2 OBJECTIVES OF THE LSF

The target number of completed interviews for the LSF was 1,200, assuming a response rate similar to current production at the time of planning for the test. This required a starting sample size of 2,500 cases to obtain the desired 1,200 complete cases.

The main objectives for the LSF were to:

- a. Develop an online diary instrument that can be accessed on a computer and mobile device and is responsive to screen size.
- b. Develop systems, field procedures, and training materials for the LSF. The test will involve a small pilot test of 9 interviews, followed by a six-month field test with a starting sample of 2,500 cases. FR debriefing after data collection will be conducted to get feedback on procedures, instruments, and materials. Respondent feedback on the online diary procedures, instrument, materials and overall experience will be obtained through a self-administered respondent debriefing.
- c. Analyze the data and paradata from the feasibility test to make decisions on the implementation of online diaries in the CED in 2022. Research questions include identifying operational and technical issues with online diary implementation, impact of cash incentives and pre-notification postcards on response rates and contact attempts, efficacy of placement and pickup procedures, usefulness of respondent materials, and differences in LSF data quality, recalled expenses, data entry patterns, and diary reporting drop-offs compared to production.

3 TEST DESIGN AND PROTOCOLS

1. **Study design:** The study sample served as the main test group. The control group came from the production sample for the same time period. As in production, LSF FRs made two visits

to the CU, one to place the diaries and the other to pick up the diaries and conduct the respondent debriefing. In the LSF, one respondent in each Consumer Unit (CU) completed a two week online diary or two one-week paper diaries. The control group from the production sample completed two one-week paper diaries. The paper diaries in the test group and control were identical.

2. **FR training:** FRs received a half day self-study module that included practice assignments prior to a one day in-person training. After training, a short test was completed.
3. **Pilot test:** A pilot test was conducted in July 2019 to ensure that there were no issues with placement, pickup and functioning of the online diary instrument.
4. **Placement:** During the placement visit, the FR collected respondent demographic information, screened and placed the diaries. For online diaries, they provided login credentials and information about the user guide and you tube tutorials to the respondent. They were asked to help the respondent login to the online diary if the respondent was willing.
5. **Online Placement:** Respondents were provided an online diary if they qualified based on responses to screening questions. The screening questions identified respondents that had internet access at home (INT_ACC) and accessed the internet (INT_ABL) at least a few times a week. If they did not qualify based on the screening questions or were eligible but unwilling or unable to use the online diary, then they were provided with two one-week paper diaries. Spanish language paper diaries were provided to non-English speaking CUs who were Spanish speaking.
6. **Advance Materials and Incentives:** The LSF had an embedded postcard and incentive experiment. Advance letters were sent to all sample addresses by priority mail. A randomly selected half of the CUs had a \$5 bill included as an incentive. The advance letter referenced the incentive for the incentive group. A pre-notification postcard was sent to the entire incentive group and a random half of the control group. The postcard which went to the incentive group referenced the incentive.
7. **Midweek Contacts:** FRs were instructed to call the respondent on day 3 and day 8 of the diary keeping period to help with any issues and remind them to fill in the diaries.
8. **Automated Emails:** Automated emails were sent to all CUs placed with online diaries who had provided their email address within 24 hours of placement providing the respondent with

links to the diary and other resources. An additional reminder email⁴ was sent on day 8, if the respondent had not logged in.

9. **FR reports:** FRs could monitor respondent activity in online diary, including expenditures reports and logins through reports they could access through Mobile Case Management (MCM). These were updated daily during the workweek but not updated on weekends and holidays.
10. **Help desk:** Respondents could call a help desk located in one of the Census telephone centers for issues with the online diary: help desk hours were 9 am to 10 pm EST on weekdays and 11 am to 9:15pm EST on weekends.
11. **Pickup:** After the diary keeping week, the FR had a follow-up interview with the respondent to review and “pickup” the diaries, if possible, and complete the respondent debriefing. A recall tab was added to the CED CAPI instrument for FRs to enter expenses based on information given by the respondent at the time of diary pickup that had not been entered in the online or paper diary.
12. **FR debriefing:** FRs completed an FR debriefing in the CAPI instrument after each visit. The FR pick up debriefing questions in the CAPI instrument were for both online and paper diary cases. The FR placement debriefing questions were for online mode only.
13. **Respondent debriefing:** Respondent debriefing at pickup was designed to be self-administered. If the respondent was unwilling or unable to complete the debriefing via self-administration or if the interview was held over the phone, the FR administered it.
14. **Field procedures changes starting in mid-March:** In mid-March 2020, due to the Covid-19 pandemic, personal visits were discontinued and online and paper diary placement and pickup was done by phone. Paper diaries already placed were transcribed by the FR during the pickup call. For new paper placements, a link to a website was sent where the respondent could view the diary form, and the respondent was asked to keep receipts, which the FR could transcribe via phone call with the respondent at the end of the diary period. If the respondent was not willing to provide an email or did not have the technology to view the paper form online, they mailed paper diaries to the respondent.

⁴ For the complete text of both emails, see Appendix E: LSF Automated Emails.

4 SAMPLE DESCRIPTION

4.1 SAMPLE DESIGN

The LSF sample consisted of 2,502 housing units selected from the CE Diary reserve sample. The sample included 2,478 unit frame sample units and 24 group quarters frame units. The sample covered the continental United States, excluding Hawaii and Alaska. The sample was fielded with placements occurring over 6 months (October 2019-March 2020). The final close out of the test was the end of April. For this report, data from October 2019 – February 2020 will be used due to delayed delivery and potential noise in the March data due to the COVID-19 pandemic.

4.2 SAMPLE DESCRIPTION

The production sample, which makes up the control group for this project, included a total of 2,614 CUs that provided complete diaries (defined by the Census outcome code definition⁵) between October of 2019 and February of 2020 (Figure 4.2.1). In the same time frame, 838 CUs in the treatment group (LSF) had complete outcome codes (Figure 4.2.2)⁶.

Figure 4.2.1. Production case- and diary-level sample totals⁷

<i>Month</i>	<i>Cases (CUs) (complete)</i>	<i>Both-week diary CUs</i>	<i>One-week diary CUs</i>	<i>No-entry CUs</i>
October	476	441 (92.6%)	20 (4.2%)	15 (3.2%)
November	442	409 (92.5%)	20 (4.5%)	13 (2.9%)
December	425	389 (91.5%)	16 (3.8%)	20 (4.7%)
January	651	583 (89.6%)	41 (6.3%)	27 (4.1%)
February	620	560 (90.3%)	34 (5.5%)	26 (4.2%)
Total	2,614	2,382 (91.9%)	131 (5.0%)	101 (3.9%)

<i>Month</i>	<i>Diary totals</i>	<i>Non-blank diaries</i>
October	917	902 (98.4%)
November	851	838 (98.5%)
December	814	794 (97.5%)

⁵ Census includes ‘temporarily absent’ CUs (OUTCOME=217) with refusals and non-contacts as non-complete cases.

⁶ Note that sample sizes increased in January 2020 (12,000 annual production addresses to 17,800 addresses).

⁷ To derive diary counts, we retained only CUs with complete disposition codes (i.e., outcome=‘201’ or ‘206’) and merged with expenditure data, creating a separate diary week for each diary containing at least 1 entry. Consequently, if a retained CU had a weekly disposition code indicating a complete (e.g., pick_up#=‘201’) in one week but did not have any entries that week, that diary week was not included in the totals here.

January	1,234	1,207 (97.8%)
February	1,180	1,154 (97.8%)
Total	4,996	4,895 (98.0%)

Figure 4.2.2. LSF case- and diary-level sample totals

<i>Month</i>	<i>Cases (CUs) (complete)</i>	<i>Both-week CUs</i>	<i>One-week diary CUs</i>	<i>No-entry CUs</i>
October	169	150 (88.8%)	12 (7.1%)	7 (4.1%)
November	166	141 (84.9%)	8 (4.8%)	17 (10.2%)
December	166	151 (91.0%)	7 (4.2%)	8 (4.8%)
January	181	158 (87.3%)	7 (3.9%)	16 (8.8%)
February	156	138 (88.5%)	8 (5.1%)	10 (6.4%)
Total	838	738 (88.1%)	42 (5.0%)	58 (6.9%)

<i>Month</i>	<i>Diary totals</i>	<i>Non-blank diaries</i>
October	319	312 (97.8%)
November	307	290 (94.5%)
December	317	309 (97.5%)
January	339	323 (95.3%)
February	294	284 (96.6%)
Total	1,576	1,518 (96.3%)

For production, 2,382 of the 2,614 CUs completing diaries completed them in both weeks, a rate slightly below that of the LSF sample – 91 percent and 88 percent, respectively. Much of this difference could be accounted for by a large share of production CUs not providing any entries in February. When examining the number of diaries among complete CUs in the samples, there were 4,996 diaries (98.0 percent of which had any entries) from production, and 1,576 diaries (96.3 percent with entries) from LSF.

4.3 SAMPLE CHARACTERISTICS

The LSF demographic composition was very similar to that of the CED Production sample (Figure 4.3.1)⁸. The most noticeable differences between both samples were the Hispanic origin

⁸ Analyses involving demographics are based on preliminary unprocessed data and are subject to change.

of respondents and housing tenure. LSF cases had a slightly higher proportion of Hispanic respondents compared to the CED Production (15.2 percent vs. 12.9 percent). Additionally, LSF respondents were more likely to be homeowners compared to CED Production (69.2 percent vs. 65.4 percent). However, these differences in sample composition were negligible due to the relatively small sample size of the test. While introducing a new mode of collection may encourage participation from different groups, there is little evidence to support this. Ultimately, there was little to no difference in the demographic characteristics of respondents between the LSF and Production.

Figure 4.3.1 Sample characteristics

<i>Demographic category</i>	<i>LSF Complete cases</i>	<i>CED Production Complete cases</i>	<i>Percent Difference (LSF minus Production)</i>
No. of Consumer Units	838	2,614	
Race of Respondent*			
White	77.8	79.6	-1.8
Black	10.8	9.1	1.7
Other (incl. Asian, multi, other)	11.5	11.3	0.2
Hispanic Origin of Respondent*	15.2	12.9	2.3
Gender of Respondent*			
Female	54.8	52.4	2.4
Male	45.2	47.6	-2.4
Age of Respondent			
Under 25 years	6.0	5.8	0.2
25-34 years	14.3	14.7	-0.4
35-49 years	22.6	23.7	-1.1
50-64 years	29.2	27.5	1.7
65 years and older	27.9	28.3	-0.4
Education*			
Less than high school	7.7	7.6	0.1
High school graduate	22.3	20.4	1.9
Some college	32.1	33.0	-0.9
College graduate	37.9	39.0	-1.1
CU Size			
Single person	31.5	28.0	3.5
2-3 persons	47.6	49.6	-2.0
4+ persons	20.9	22.4	-1.5
Housing Tenure*			
Renter	30.8	34.6	-3.8

<i>Demographic category</i>	<i>LSF Complete cases</i>	<i>CED Production Complete cases</i>	<i>Percent Difference (LSF minus Production)</i>
Owner	69.2	65.4	3.8

*Race is unknown for 2 LSF cases and 9 production cases; Hispanic origin is unknown for 5 Production cases; Gender is unknown for 4 LSF cases and 34 Production cases; Education is unknown for 9 LSF cases and 25 Production cases; and Housing tenure is unknown for 3 LSF cases and 32 Production cases.

5 PARTICIPATION

5.1 ONLINE PLACEMENT AND DEMOGRAPHICS OF MODE CHOICE

5.1.1 Diary placement process

The diary placement protocol was similar to current CED placement procedures, except that every CU respondent was screened to see if they had home internet access (INT_ACC) and were able to access the internet either daily or at least a few times a week (INT_ABL) (by self-reporting). As a result, respondents were provided an online diary based on satisfying these screening questions. If they did not qualify based on the screening questions or were otherwise unwilling or unable to use the online diary, then they were provided with the standard paper diary used in the CED. For respondents completing the online diary, the FR was instructed to assist the diarist with setting up access to the diary and reviewing the diary-keeping procedures. The FR was also tasked with providing materials and information that the respondent would need during the diary-keeping weeks.

Among complete CUs (n=838), 67.5 percent of respondents were recorded as having internet access, while 30.2 percent did not have access and 1.4 percent reported they did not know (Figure 5.1.1.1). Of the 566 cases that had internet access, 51.1 percent reported having the ability to access the internet on a daily basis with 1.7 percent saying they have access a few times per week. Ultimately, the sample was reported to have 52.4 percent of CUs (439 CUs) who were eligible for the online diary. Respondents who were online eligible but were non-English Spanish speakers were then assigned Spanish paper diaries.

The internet access rate of 67.5 percent observed in the LSF is lower than the national estimate of 86.4 percent broadband access from the 2019 American Community Survey.

Figure 5.1.1.1: Diary placement process

<i>Screeners questions and eligibility</i>	<i>Percent</i>
Internet access* (asked of all CUs, N=838)	
Yes	67.5
No	30.2
Don't know	1.4
Internet Ability (asked if Internet access = yes, N=566)	
Daily	51.1
A few times per week	1.7
A few times per month (ineligible for online diary)	5.2
Less than a few times per month (ineligible for online diary)	6.0
Don't know	4.1
Eligible for online diary (of all eligible members, N=838)	52.4

*Internet access is unknown for 7 cases.

However, in light of the online diary placement outcomes displayed in Figure 5.1.1, there were several potential discrepancies during the data collection process that may challenge the reliability of these variables. There is evidence that FRs changed responses to these screening questions in order to place a paper diary with an online eligible respondent instead of the expected protocol, which was to “refuse” the online diary and then place a paper diary. Therefore, we may have found lower numbers for online diary eligible paper diary placements and higher numbers for non-eligible paper diaries.

We further investigated these issues by examining the LSF audit trail data. Audit trail data log all navigational movements and data transactions (e.g., data entry, data storage) within the instrument, noting when an FR changes values for a particular field. Figure 5.1.1.1 is recreated to reflect initial responses before the FR changed the answer as seen in the audit trails (Figure 5.1.1.2). It appears that internet access, ability and online eligibility was higher than recorded in Figure 5.1.1. Further information on the mode changes can be found in Appendix A.

Figure 5.1.1.2. Diary placement process (initial response based on audit trail)

<i>Screeners questions and eligibility</i>	<i>Percent</i>
Internet access* (asked of eligible members, N=838)	
Yes	68.2
No	30.5
Don't know	1.4

<i>Screening questions and eligibility</i>	<i>Percent</i>
Internet Ability (asked if Internet access = yes, N=566)	
Daily	53.4
A few times per week	2.4
A few times per month (ineligible for online diary)	4.0
Less than a few times per month (ineligible for online diary)	4.9
Don't know	3.0
Eligible for online diary (of all eligible members, N=838)	55.8

*Internet access is unknown for 7 cases.

LSF protocols required that FR's encourage respondents to complete the online diary, and only provide the option for a paper diary if the respondent does not have sufficient internet or computer access or as a refusal avoidance tactic. Figure 5.1.1.3 displays the mode distribution of eligible member's internet access for completed cases. Overall 43.8 percent of the 838 respondents were eligible and were placed with an online diary, while 8.6 percent were eligible, but placed with a paper diary. Meanwhile, 47.5 percent of respondents were not eligible for an online diary based on the recorded internet access and ability answers and were placed with a paper diary.

Figure 5.1.1.3 Mode placement distribution of eligible members (N=838)

<i>Mode of placement</i>	<i>N</i>	<i>Percent</i>
Online	367	43.8
Paper (eligible for online)	72	8.6
Paper (ineligible for online)	398	47.5

Comparing the placement of the diaries to the mode of completion (Figure 5.1.1.4) for all 1772 CUs, we find that, of the 445 respondents placed with online diaries, 367 were completed online, and the remaining were Type A or temporarily absent at pickup. Of the 530 respondents placed with paper diaries in English or Spanish, one completed the diary online and 470 respondents completed paper diaries. The remaining 59 were Type A or temporarily absent. There was negligible switching of modes once the diary was placed.

Figure 5.1.1.4: Comparing mode of placement and mode of completion

<i>Mode of Placement</i>	<i>Mode of completion</i>			<i>Total</i>
	Online	Paper	Type A or Temporarily absent	

Online	367	0	78	445
Spanish paper	0	55	8	63
Paper	1	415	51	467
Type A at placement	0	0	797	797
Total	368	470	934	1772

5.1.2 Demographics of Mode Choice

Figure 5.1.2.1 presents the demographic composition of the online and paper reference persons, respectively⁹. A Reference person is defined as the first member mentioned by the respondent when asked to "Start with the name of the person or one of the persons who owns or rents the home." Reference persons in the online diary group were generally younger, had higher levels of education, a larger CU size and were more likely to be homeowners. Over 53 percent of the online group had a reference person who was under the age of 49 compared to 34 percent of the paper group. Additionally, 46.9 percent of the online group had a college graduate for a reference person compared to only 30.7 percent of the paper group. CUs more than one person were more likely to be in the online diary group (75 percent compared to 64 percent). Lastly, almost 71 percent of online respondents were homeowners compared to 68.2 percent of paper respondents¹⁰.

Figure 5.1.2.1 Demographics of reference person by mode choice (N=838)⁵

<i>Demographic category</i>	<i>LSF Online diary</i>	<i>LSF Paper diary</i>	<i>Percent difference (online-paper)</i>
No. of Consumer Units	368	470	
Race			
White	78.0	77.6	0.4
Black	10.3	11.1	-0.8
Other (incl. Asian, multi, other)	11.7	11.3	0.4
Hispanic Origin of Respondent	12.2	17.4	-5.2
Gender*			
Female	56.0	53.8	2.2
Male	44.0	46.2	-2.2
Age			
Under 25 years	7.9	4.5	3.4

⁹ Analyses involving demographics are based on preliminary unprocessed data and are subject to change.

¹⁰Differences were not significant.

<i>Demographic category</i>	<i>LSF Online diary</i>	<i>LSF Paper diary</i>	<i>Percent difference (online-paper)</i>
25-34 years	17.7	11.7	6.0
35-49 years	28.0	18.3	9.7
50-64 years	29.1	29.4	-0.3
65 years and older	17.4	36.2	-18.8
Education*			
Less than high school	2.7	11.7	-9
High school graduate	13.9	29.0	-15.1
Some college	36.5	28.6	7.9
College graduate	46.9	30.7	16.2
CU Size			
Single person	25.3	36.4	-11.1
2-3 persons	51.1	44.9	6.2
4+ persons	23.6	18.7	4.9
Housing Tenure*			
Renter	29.4	31.8	-2.4
Owner	70.6	68.2	2.4

*Race is unknown for 2 cases; Gender is unknown for 4 cases; Education is unknown for 9 cases; and Housing tenure is unknown for 3 cases.

The option to complete the diary online presents both a unique opportunity to reach certain respondents while risking the cooperation of other respondents. Completing an online diary requires a basic technological understanding along with adequate internet access and appropriate technology, all of which can vary greatly across socioeconomic background. We modelled this with a multivariate logistic regression, predicting whether a diary was completed online based on seven demographic characteristics (Figure 5.1.2.2). As we used unprocessed survey data which have not had any editing, allocation of combined expenses, or imputation of missing expenses done to them, we decided to omit CU income from our model. Income collected in the survey is subject to post-processing edits that account for missing income values.

Figure 5.1.2.2. Model of diary mode (Online mode = 1) for completed interviews by demographic characteristics

<i>Regressors</i>	<i>Coefficient</i>	<i>SE</i>
Intercept	1.3216***	(0.3528)
Age	-0.0397***	(0.00517)
Hispanic	-0.6085**	(0.2215)

CU size	-0.0252	(0.0565)
Homeowner	0.4557*	(0.1777)
Male	-0.1289	(0.1507)
White	0.2211	(0.1844)
College degree	0.6245***	(0.1537)
R ² = 0.1088	N = 828	
*p<.05; **p<.01; ***p<0.001		

In line with Figure 5.1.2.1, we find that several of these demographic variables are significantly associated with choosing to complete the online diary. The significant negative associations were for age, homeowner, and Hispanic origin. This largely follows the demographic composition shown in Figure II.5. In Figure II.5, most reference persons 65 years and older were likely to be in the paper diary group (36.2 percent vs. 17.4 percent, respectively), with Hispanic reference persons likely to be in the paper diary by greater margins (17.4 percent vs. 12.2 percent, respectively). It is worth noting that non-English speaking Spanish speakers were automatically given the paper diary, so this result is not surprising. On the other hand, a white reference person and those with a college degree were positively associated with accepting the online diary; however, only having a college degree was statistically significant, which was consistent with the demographic composition of the sample (Figure 4.3.1).

5.2 OVERALL COLLECTION RESPONSE RATES

For the LSF test, we calculated the overall collection response rates using the AAPOR RR2 definition¹¹: the total number of complete and partial interviews divided by the total number of eligible interviews (the sum of complete interviews, partial interviews, plus Type A non-interviews). Complete and partial interviews (cases) were comprised of respondents who completed at least one full week of the diary-keeping period. Type A non-interviews are interviews that were not completed by respondents due to either respondent refusal or the inability to reach respondents, including cases that were identified as temporarily absent by the

¹¹ AAPOR RR2 definition can be found at “https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf”

FR. Figure 5.2.1 presents the overall LSF response rates compared to similarly calculated CED Production response rates for the same months of October through February. The overall LSF response rate was slightly lower than Production (47.3 percent vs. 51.7 percent, respectively). In further sections we investigate the effects of incentives and the introduction of the online diary mode on LSF response rates.

Figure 5.2.1 Overall Response Rates Compared to CED Production¹²

<i>Counts/Rates</i>	<i>Definition</i>	<i>LSF</i>	<i>CED Production</i>
Starting Sample	(Type A + Type B/C + Complete ¹³)	2,099	6,030
Type B/C Cases		327	972 ¹⁴
Type B/C Rate	$\left(\frac{\text{Type B/C}}{\text{Type A + Type B/C + Complete}} \right)$	15.6%	16.1%
Eligible Cases	(Type A + Complete)	1,772	5,058
Type A Cases		934	2,444
Type A Rate	$\left(\frac{\text{Type A}}{\text{Type A + Complete}} \right)$	52.7%	48.3%
Complete Cases/Diaries		838	2,614
Response Rate	$\left(\frac{\text{Complete}}{\text{Type A + Complete}} \right)$	47.3%	51.7%

The LSF sample had 934 eligible respondents (52.7 percent) marked as a Type A refusal or noncontact (Figure 5.2.1). Of these eligible non-respondents, over 55 percent refused to participate in the survey, while 17 percent were noncontacts. Additionally, over 27 percent of eligible respondents who started the survey did not finish due to other reasons (e.g. diaries placed too late, majority recall). These results are compared to Production data over the same five months. The most discernable difference is that the rate of refusals is lower in the LSF (55.4

¹² Response rates calculated are considered Interim case level response rates, which are subject to revisions after processing.

¹³ Type A refers to non-responding cases, and Type B/C refers to various types of ineligible cases (e.g., vacant or abandoned housing units, businesses)

¹⁴ Type B/C counts may differ from official production counts due to differences in methodology of calculation.

percent) compared to Production (59.4 percent). Rates of noncontacts (including outcomes of temporarily absent and no one home) was also slightly lower in the LSF (17.2 percent) compared to Production (18.3 percent). The rate of Other Type A¹⁵ (27.4 vs. 22.3 percent) and the rate of Type A incompletes (14.1 vs 12.2 percent) were higher in the LSF compared to Production. Type A incompletes were respondents who started the survey but did not finish providing at least a week’s worth of diary keeping via diary entries, records or recall with the FR and did not have their diaries picked up.

Figure 5.2.2 Overall LSF Dispositions compared to CED Production

<i>Disposition</i>	<i>LSF</i>		<i>CED Production</i>	
	<i>N</i>	<i>Percent</i>	<i>N</i>	<i>Percent</i>
Eligible Sample	1,772	100	5,058	100
Type A Refusal or Noncontact	934	52.7	2,444	48.3
<i>Refusal</i>	517	55.4	1,452	59.4
<i>Noncontact</i>	161	17.2	446	18.3
<i>Other</i>	256	27.4	546	22.3
Type A Incomplete (placed but not picked up)*	137	14.1	299	12.2
Complete Cases	838	47.3	2,614	51.7

*See Appendix A for assumptions on Type A Incompletes.

To further explore the reasons for the higher percentage of Type A incompletes, we looked at the mode breakdown at placement for these cases. According to Figure 5.2.3, over half of the sample units who did not complete the diary were initially assigned an online diary.

Figure 5.2.3. LSF Incompletes by mode at placement (N=137)

<i>Initial placement by mode</i>	<i>N</i>	<i>Percent</i>
Online diary	78	56.9
Spanish paper diary	8	5.8
Paper diary	51	37.2

¹⁵ Other Type A includes silent refusals and outcomes when the diary was disqualified for various reasons.

5.3 THE EFFECTS OF LSF ADVANCE MATERIALS ON SURVEY PARTICIPATION

5.3.1 Postcard and Incentive experiment design

As described above, the LSF had an embedded postcard and incentive experiment. Prior to each month of LSF data collection, a varying combination of pre-notification postcards, advance letters, and monetary incentives were sent to sample addresses during the phase of respondent recruitment. Half of the sample received an advance postcard and a \$5 monetary incentive included with the advance letter sent via Priority Mail; a quarter of the sample received an advance postcard and an advance letter via Priority Mail; and a quarter of the sample received only an advance letter via Priority Mail.

To isolate the impact of pre-notification postcards, advance letters, and monetary incentives on participation outcomes, sample CUs were analyzed in several treatment and control groups based on the following criteria; their assignment to receive postcards or incentives, and if they reported receiving postcards or advance letters during the interview¹⁶. These test groups were captured by a set of variables that provided initial group assignment information as well as whether the Consumer Unit (CU) reported receiving each of the individual LSF advance materials. Note that only CUs that started the interview questions had the opportunity to report receiving the advance materials.

5.3.2 Pre-notification Postcards Overview

The first respondent recruitment method for the LSF was the pre-notification postcard. The postcard was sent through the United States Postal Service (USPS) prior to the advance letters and monetary incentives¹⁷ to all of the incentives test group and a randomly selected half of the remaining CUs – overall making up roughly 75 percent of the total sample. The total breakdown of each postcard group is detailed below in Figure 5.3.2.1.

Figure 5.3.2.1 Postcard Sample Group Assignment

<i>Postcard Groups</i>	<i>Group Totals</i>	<i>Group Percent</i>
Postcard Treatment Group	1,332	75.2%
Control Group	439	24.8%

¹⁶ All sample CUs were sent an advance letter, which contained a monetary incentive for those in the assigned incentive group. For this reason, advance letter test groups were determined solely on whether the CU reported receiving it.

¹⁷ 1,332 sample addresses were sent postcards out of the total sample of 1,771, leaving a control group of 439.

Total Sample ¹⁸	1,771	100%
----------------------------	-------	------

Among eligible CUs assigned to the postcard group, 466 were not asked the question about whether they received the postcard (all but two of these cases were ultimately coded Type A). The remaining 867 participated in the placement at least to the point of providing an answer to the postcard receipt question, and of that group 506 (i.e. roughly 58 percent who answered the question) reported receiving the postcard.

Figure 5.3.2.2 Postcard Reported Received by those in Treatment Group

<i>Postcard Received Groups</i>	<i>Group Totals</i>	<i>Group Percent</i>
Postcard Reported Received	506	38%
Postcard Not Reported Received	360	27%
Question Not Asked	466	35%
Total Postcard Treatment Group	1,332	100%

As for the 360 sample CUs that were assigned to the postcard treatment group, but did not report receiving it when questioned, there are some plausible explanations. For example, the reference person answering the question may not have been the CU member who retrieved the mail on the day that the postcard arrived, the postcard could have been lost in the shuffle of bills and advertisements and was simply forgotten or even accidentally disposed of.

5.3.3 Pre-notification Postcards Analysis Results

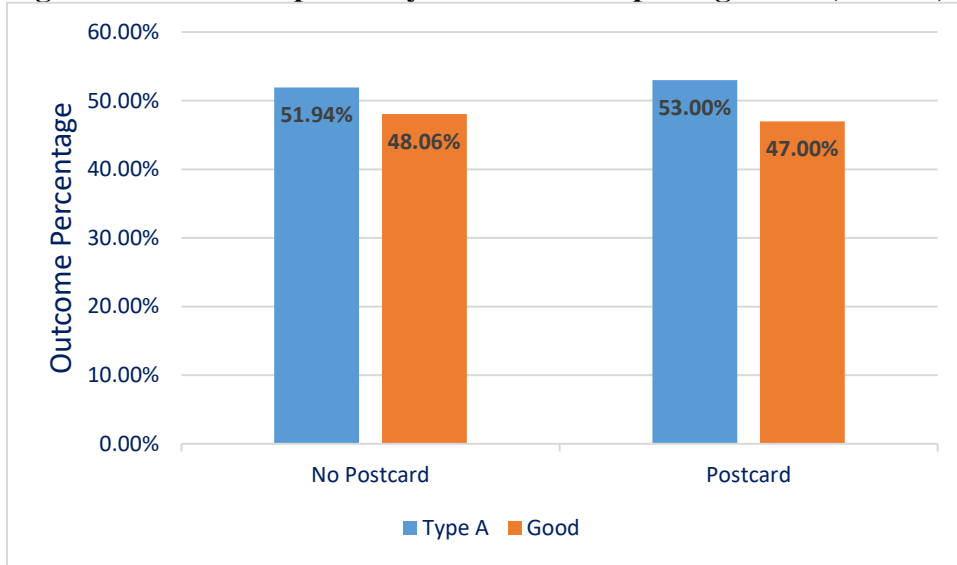
In an effort to isolate the effects of postcard group assignment from reported postcard receipt, two separate analyses were conducted: (1) looking at participation based on assigned postcard group among all eligible CUs and (2) looking at participation based on reported receipt of postcard among CUs that completed the placement interview. The results of the analyses showed that survey participation outcomes did not vary significantly between assigned postcard groups, but did vary significantly between CUs who reported differing postcard receipt statuses.

Among all eligible CUs, Figure 5.3.3.1 below shows that there was little difference in participation outcomes between groups. For those CUs who were assigned to the postcard group,

¹⁸ Total sample in Figure 5.3.2.1 is 1,771 rather than 1,772 due to a spawned case that was not initially assigned to a postcard group.

47 percent were associated with “Good” outcomes, while their counterparts in the control group were associated with “Good” outcomes 48 percent of the time. These differences are not statistically different.¹⁹

Figure 5.3.3.1 Participation by Postcard Group Assignment (n=1771)



Looking at the group that completed the placement interview by whether they reported receiving a pre-notification postcard or not (Figures 5.3.3.2 and 5.3.3.3), we find that those who reported receiving a postcard were associated with “Good” final outcome codes 75 percent of the time. For those CUs who answered the postcard receipt question, but did not report having received a pre-notification postcard, the within group percentage of “Good” outcome codes was approximately 68 percent. These differences were found to be statistically significant²⁰.

¹⁹ For Figure 5.3.3.1, we fail to reject the null of no significant relationship.
 Pearson $\chi^2(1) = 0.1508$ Pr = 0.698 Likelihood-ratio $\chi^2(1) = 0.1507$ Pr = 0.698
²⁰ For Figure V.3.3.2, we may reject the null of no significant relationship.
 Pearson $\chi^2(1) = 4.8955$ Pr = 0.027 Likelihood-ratio $\chi^2(1) = 4.8640$ Pr = 0.027

Figure 5.3.3.2 Participation by Postcard Receipt Status (n=866)

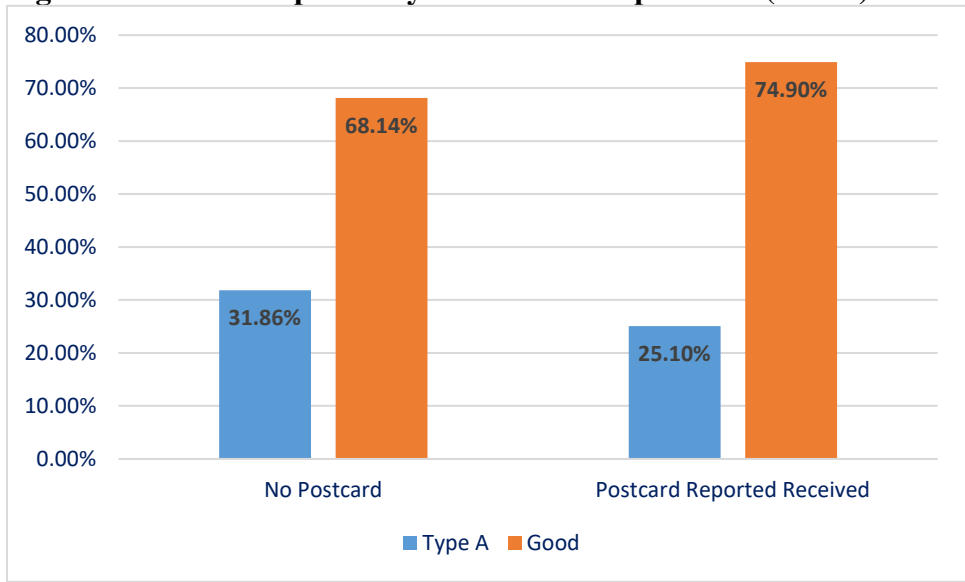


Figure 5.3.3.3 Participation by Postcard Receipt Status

<i>Postcard Receipt Groups</i>	<i>Type A</i>	<i>Good</i>	<i>Total</i>
Postcard Reported Received	127 (25.1%)	379 (74.9%)	506 (100%)
Postcard Not Reported Received	115 (31.9%)	245 (68.1%)	360 (100%)
Total	242 (28%)	624 (72%)	866 (100%)

A sample unit’s inclusion in this part of the analyses was predicated on the CU providing an answer to the postcard receipt question posed by the FR. The fact that there was a successful contact to the CU and that the reference person did not refuse to answer, generally bodes well for survey participation when compared to the alternative.

5.3.4 Advance Letter and Incentives Overview

Advance letters were sent out to all sample addresses using USPS Priority Mail, subsequent to the pre-notification postcard. Similar to the postcard, CUs were asked whether they recalled receiving the advance letter during the placement interview. A tabular breakdown of advance

letters receipt status captured during the placement interview questioning can be found in Figure 5.3.4.1 below.

Figure 5.3.4.1 Advance Letter Receipt Status –eligible sample

<i>Advance Letter Receipt Status</i>	<i>Group Totals</i>	<i>Group Percent</i>
Letter Reported Received	1,017	57.4%
Letter Not Reported Received	193	10.9%
Question Not Asked	561	31.7%
Total Advance Letter Group	1,771	100%

The resulting data revealed that among all eligible CUs, 561 CUs were not asked the question about whether they received the letter and were coded as “Type A” at placement of the diary. The remaining CUs participated in the placement interview, at least to the point of the advance letter receipt question being analyzed, and of this group 1,017 out 1,210 (i.e. roughly 84 percent) reported receiving the advance letter. Roughly half of these sample addresses were randomly selected to receive a \$5 incentive contained within the advance letter. A breakdown of the incentive groups can be found in Figure 5.3.4.2 below.

Figure 5.3.4.2 Incentive Group Assignment

<i>Incentive Groups</i>	<i>Group Totals</i>	<i>Group Percent</i>
Incentive Group	879	49.7%
Non-Incentive Group	889	50.3%
Total Sample ²¹	1,768	100%

Of the eligible sample CU’s assigned to the monetary incentive group, 285 were not asked the question about whether they received the postcard (coded as Type A at the placement interview). The data show that the remaining 594 CU’s from the incentive group participated in the placement interview, at least long enough to be asked the advance letter receipt question. Of these CUs, 501 (i.e. roughly 84 percent) reported that they had received the advance letter, which

²¹ Total sample in Figure 5.3.4.1 is 1,768 rather than 1,772 due to four spinoff cases not being initially assigned to an incentive group.

contained the monetary incentive. A further breakdown of the letter receipt status by incentive group can be found in Figure 5.3.4.3 below

Figure 5.3.4.3 Advance Letter Receipt Status – Incentive Group

<i>Incentive Group</i>	<i>Group Totals</i>	<i>Group Percent</i>
Letter Reported Received	501	57%
Letter Not Reported Received	93	10.6%
Question Not Asked	285	32.4%
Total Incentive Treatment Group	879	100%

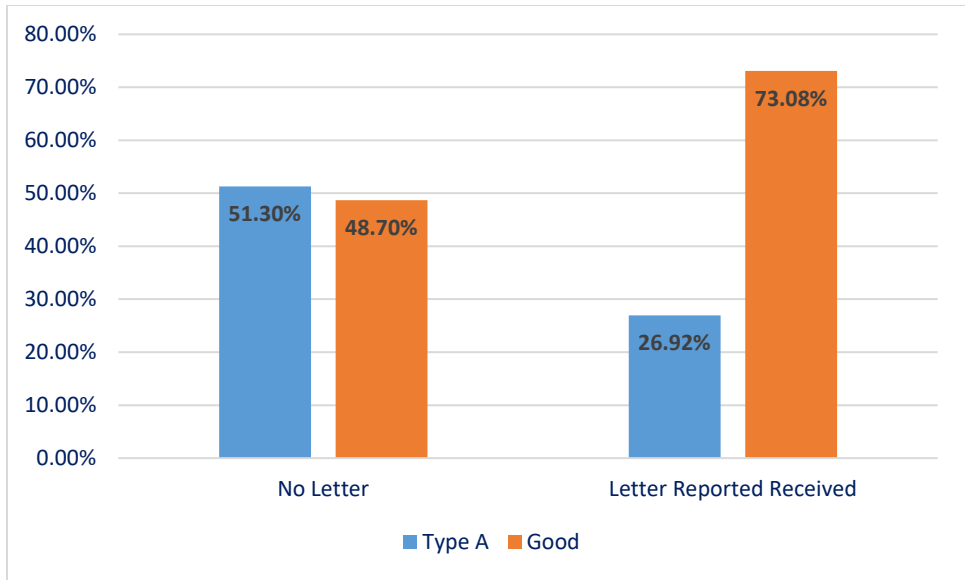
For those CUs that provided an answer, but did not recall receiving an advance letter, this could have been the result of a lost or misplaced letter. When comparing the breakdowns in Figures 5.3.4.1 and 5.3.4.3 to 5.3.2.2, respondents were proportionally more likely to report receiving the advance letter (57 percent) than the pre-notification postcard (38 percent). It is not totally clear why this was the case, but it could be due to the fact that postcards by nature are smaller and likely easier to be lost in the shuffle of normal mail or mistaken for a junk-mail advertisement. It is also worth noting that the postcard arrived at eligible sample addresses at an earlier date than the advance letter, and that advance letters were sent by USPS Priority Mail while postcards were sent by regular U.S. Mail, making them more likely to be misplaced or forgotten.

5.3.5 Advance Letter and Incentive Analysis Results

The initial analysis showed that participation of CUs that were asked the placement interview questions varied significantly between CUs who reported receiving advance letters compared to those who did not.

Figure 5.3.5.1 illustrates participation outcomes by whether CUs that took part in the placement interview reported receiving an advance letter and shows that those who reported receiving an advance letter were associated with “Good” outcome codes 73 percent of the time. For those CUs reporting not to have received an advance letter, the within group percentage of “Good” outcome codes was approximately 49 percent.

Figure 5.3.5.1 Participation by Advance Letter Received Status (n=1,210)

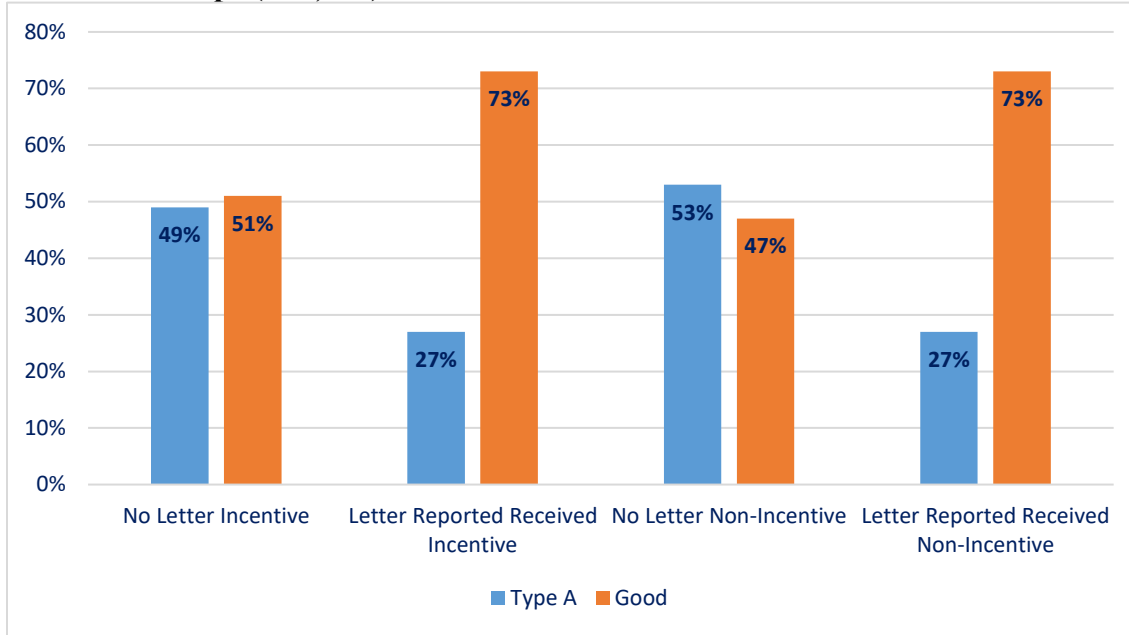


The difference in participation outcomes between those who **did** reported receiving an advance letter and those who **did not** is fairly wide. The data do show, and significance tests confirm, that those CUs who reported receiving an advance letter were more likely to be associated with “good” participation outcomes than those the in the control group.²²

After determining an initial positive statistical relationship existed between advance letter receipt status and participation outcomes, the effect of advance letter receipt status was further analyzed by assigned incentive group. Figure 5.3.5.2.below shows participation outcomes by reported advance letter receipt status, as well as by incentive group assignment.

²² For Figure 5.3.5.1., we may reject the null hypothesis of an independent relationship at the 1% level of significance. Pearson $\chi^2(1) = 45.2436$ Pr = 0.000 Likelihood-ratio $\chi^2(1) = 42.3570$ Pr = 0.000

Figure 5.3.5.2 Participation by Advance Letter Received Status for Incentive and Non-Incentive Groups (n=1,208)



The data in Figure 5.3.5.2 show that for CUs that were in the incentive group, roughly 73 percent who reported receiving an advance letter, and by extension their monetary incentive, were associated with “Good” participation outcomes. For those who were in the incentive group, but did not report receiving an advance letter, the data show that the distribution of “Good” and “Type A” participation outcomes were roughly equal. Figure 5.3.5.2 also shows a similar trend for CUs who were not assigned to the incentive group. The data shows that about 73 percent of CUs that were not assigned to the incentive group and reported receiving an advance letter, were associated with “Good” participation outcomes. CUs in the non-incentive group that did not report receiving an advance letter were associated with “Good” participation outcomes only 47 percent of the time. Figure 5.3.5.3 below shows the total frequency of participation outcomes by receipt status for both the incentive group and the non-incentive.

Figure 5.3.5.3 Participation by Advance Letter Received Status for Incentive and Non-Incentive Groups

<i>Advance Letter Receipt Status</i>	<i>Type A</i>	<i>Good</i>	<i>Total</i>
Incentive Group Letter Reported Received	136 (27.1%)	365 (72.9%)	501 (100%)
Incentive Group Letter Not Reported Received	46 (49.5%)	47 (50.5%)	93 (100%)

<i>Advance Letter Receipt Status</i>	<i>Type A</i>	<i>Good</i>	<i>Total</i>
Non-Incentive Group Letter Reported Received	138 (26.8%)	376 (73.2%)	514 (100%)
Non-Incentive Group Letter Not Reported Received	53 (53%)	47 (47%)	100 (100%)
Total	373 (30.9%)	835 (69.1%)	1,208 (100%)

Further analysis reveals that a statistically significant relationship exists between advance letter receipt status and participation outcomes for those in the incentive group.²³ This was also found to be the case for those in the non-incentive group, in which significance testing once again confirms that a statistically significant positive relationship exists between CUs who reported receiving an advance letter, and “Good” participation outcomes.²⁴

The results above show that participation outcomes, when analyzed by advance letter receipt status, were nearly identical between the incentive and non-incentive group. This would indicate that the mere receipt of the advance letter is sufficient in effecting participation outcomes, and that the inclusion of the monetary incentives makes no significant difference.

5.3.6 Effectiveness of Unconditional Incentives on CU participation

As earlier described in section 5.3.4, the LSF test included an embedded experiment in which a \$5 cash incentive was included with the advance letter for a random half of the sample. The advance letter also referenced the \$5 bill for the incentive group. Figure 5.3.6.1 presents the distribution of CU’s that received an incentive by complete cases and Type A cases. The incentive group had a response rate of 46.9 percent compared to the control group which had a response rate of 47.6 percent, suggesting incentives had no significant impact on response rates.

²³ For Figure 5.3.5.3, we may reject the null hypothesis of an independent relationship at the 1% level of significance. Pearson chi2(1) = 18.3822 Pr = 0.000 likelihood-ratio chi2(1) = 17.2442 Pr = 0.000
²⁴ For Figure 5.3.5.3 we may reject the null hypothesis of an independent relationship at the 1% level of significance. Pearson chi2(1) = 26.7153 Pr = 0.000 likelihood-ratio chi2(1) = 25.0068 Pr = 0.000

Figure 5.3.6.1. Percent of CUs receiving unconditional incentives

<i>Incentive status*</i>	<i>Complete</i>	<i>Type A</i>	<i>All</i>	<i>Response Rates</i>
Incentive group	412	467	879	46.9%
Control group	423	466	889	47.6%
Total	835	933	1,768	47.2%

*There were 4 cases not categorized in either the incentive or control group. Including these 4 cases would result in the overall response rate of 47.3% as shown in Figure 5.2.1.

The goal of the LSF test was to evaluate the effectiveness of an online diary for potential use in production. Providing an incentive is thought to encourage participation and make respondents more amenable to taking the diary in a new technological format. Figure 5.3.6.2 breaks down the mode distribution by incentive group. Among the 412 respondents that received an incentive, only 43.7 percent used the online diary. This also holds for those in the control group, with most respondents choosing to complete the paper diary (e.g. internet access problems, respondent preferences etc.). There was no significant difference in mode choice between the incentive group and control group. In both groups, a higher percentage chose paper diaries.

Figure 5.3.6.2. Percent distribution of mode by Incentive group

<i>Group*</i>	<i>N</i>	<i>Online diary</i>	<i>Paper diary</i>
Incentives	412	43.7	56.3
Control group	423	44.2	55.8
Total	835	43.9	56.1

*There were 3 cases not categorized in either the incentive or control group

The following figures provide responses from the respondent debriefing questions. The respondent debriefing at pickup was a self-administered survey, although respondents could choose to have the FR enter answers for them. In section 6, we further analyze the remainder of the respondent debriefing questions.

Among respondents who received the incentive, we asked if they received the \$5 bill and if receiving the money affected their decision to participate the survey. According to Figure 5.3.6.3, 65.1 percent of respondents reported receiving the \$5 cash incentive while 11.1 percent did not. However, 21.4 percent of respondents reported that they did not know if they received the \$5 bill with their advance letter.

Figure 5.3.6.3. Percent receiving \$5 incentive among CUs completing the debriefing (N=332)

<i>Question</i>	<i>Response options</i>	<i>% of CUs</i>
In the first letter we sent you, explaining the survey and asking for you to participate, we included a five-dollar bill. Did your CU receive this money?	Yes	65.1%
	No	11.1%
	Don't know	21.4%
	Refused	2.4%

From the respondents who reported receiving the incentive, approximately 32 percent said that the \$5 made them more willing to complete the diary compared to only 1.4 percent that said it made them less willing to complete the diary (Figure 5.3.6.4). Meanwhile, over 65 percent of respondents said the \$5 had no effect on their decision to participate in the survey. These conclusions should take into account that only 216 people (52 percent) responded from the entire incentive group and were respondents that completed the interview and debriefing. Therefore, the respondent debriefing questions only provide a partial insight into the effectiveness of incentives on participation.

Figure 5.3.6.4. Effect of incentives on CU participation (N=216)

<i>Question:</i>	<i>Response options</i>	<i>% of CUs</i>
How did the money your CU received affect your decision to complete the diary?	It made me more willing to complete the diary	31.9%
	It made me less willing to complete the diary	1.4%
	It had no effect on my decision to complete the diary	65.3%
	Don't know	1.4%

Lastly, we empirically tested for the impact of \$5 incentives on CU completion of the LSF, dependent on successful placement. We were interested if incentives and contact attempts encouraged survey participants to complete the diary regardless of mode. To model this, we used data comprising complete cases and Type A incompletes (N=975). This allowed us to develop a multivariate logistic regression model, predicting completion in the LSF based on receiving the cash incentive, total no. contact attempts and other demographic correlates.

Figure 5.3.6.5. Model of completion rates in LSF based on interview & demographic characteristics

<i>Regressors*</i>	<i>Coefficient</i>	<i>SE</i>
Intercept	2.3868***	(0.5174)
Total no. contact attempts	-0.0611*	(0.0283)
Incentive received	0.0138	(0.1812)
Age	0.00501	(0.00611)
CU size	-0.1562**	(0.0603)
Homeowner	0.3229	(0.2012)
Male	-0.0838	(0.1827)
White	-0.1963	(0.2227)
College degree	0.1959	(0.1935)
Internet access ²⁵	-0.4244*	(0.1945)
R ² = 0.0271	N = 975 ²⁶	
*p<.05; **p<.01; ***p<0.001		

*("Good" outcome = 1).

The results of our model (Figure 5.3.6.5) show that the total number of contact attempts, CU size, males, white reference persons and internet access are negatively associated with completing the LSF ($P < 0.05$). We defined completion of the LSF if the CU was marked as having completed at least one diary week. College degree and age were all positively associated with completion. On the other hand, holding other variables constant, a reference person was less likely to complete the diary the more contact attempts received, the larger the CU size, and if they had internet access. Also, in line with our sample characteristics (Figure 4.3.1), males and white reference persons were less likely to finish the LSF test (10.33 percent and 16.75 percent, respectively). The reason for the negative association between internet access and likelihood of completion is not clear.

²⁵ Internet access is defined as having internet access *and* the ability to access internet daily or at least a few times a week

²⁶ Note that the smaller number of observations here (975 compared to 1772) is due to missing demographic characteristics for a large proportion of type A respondents, who probably never started the interview.

While respondents receiving the incentive were more likely to complete the LSF (1.38 percent) this result was not statistically significant. When considered in context of the non-significant decrease in contact attempts²⁷ and the lower response rates compared to the control group²⁸, incentives do not play any role in gaining respondent cooperation or successful completion of the LSF. Meanwhile, we see a negative relationship between total number of contact attempts and completion. Despite encompassing all types of contacts throughout the collection period, the total contact attempts variable can serve as a useful proxy for gaining respondent's cooperation. A respondent receiving more total contact attempts overall were 6.1 percent less likely to finish the survey ($P < 0.05$).

5.4 RESPONDENT PARTICIPATION AND CONTACT EFFECTIVENESS

5.4.1 Effectiveness of Contact Attempts on CU participation

Survey producers often expend numerous resources in order to obtain high quality data and ensure adequate response rates. However, reaching out to respondents proves to be a costly effort, not only for survey producers but also for respondents in terms of respondent burden. In this section, we analyze FRs' reported contact attempt history information from the Contact History Instrument as a proxy for the cost associated with reaching respondents and gaining their participation.

In this section, we compare the number of contact attempts between the LSF and production groups. This allows us to analyze how well LSF FRs were able to reach respondents compared to production (Figure 5.4.1.1). In Figure 5.4.1.2, we break down FRs' reported contact attempt information from the LSF by respondents who either received or did not receive the \$5 incentive. From this we can infer the impact, if any, of incentives on the cost associated with gaining

²⁷ Figure 5.4.1.2 displays little to no difference between incentive and control groups in the no. of contact attempts and contacts made with sample unit member (5.8 vs. 5.9 average contact attempts between incentive and control groups, respectively and 3.3 successful contacts made with sample unit member between both groups).

²⁸ According to Figure 5.3.6.1, CUs receiving the cash incentive had a response rate of 46.9 percent compared to 47.6 percent for the control group.

cooperation. To determine the average count of contact attempts per case, we use the following definition for both LSF and Production samples:

$$\text{Average counts per case} = (\text{total counts} / \text{number of cases})$$

Figure 5.4.1.1 shows the average number of contact attempts for the LSF was higher compared to Production for complete interviews (5.9 contact attempts per case vs. 5.0 attempts, respectively). A possible explanation is that LSF respondents may have required more communication with FR's due to the introduction of the online diary and any complications that may have arisen as a result (as described in Section 6).

We were also interested in seeing the average number of contact attempts necessary to securing a “successful placement” of the diary with the respondent. This is critical to reducing operational costs in the future. We defined the number of contact attempts prior to “successful placement” if the contact date was either before or on the placement date. LSF FRs' required more total contact attempts on average in order to successfully place a diary (2.9 contact attempts vs. 2.0 Production contact attempts, respectively). We were also interested in comparing the average number of LSF midweek and residual contact attempts with production. We defined midweek contact attempts if the FR attempted to make a phone call or personal visit to respondents between days 3 and 8 after initial placement. Residual contact attempts are the remaining number of contacts after placement and midweek contacts. However, there were fewer residual contact attempts made after placement and the midweek contact periods for LSF cases compared to Production (2.5 contact attempts vs. 2.9 Production contact attempts, respectively).

Lastly, to understand the costs associated with reaching respondents, we looked at the average number of in-person contact attempts for LSF. This is because in-person contact attempts are more costly than contacting respondents by phone. Both the LSF and Production fielding required the same average number of in-person contact attempts overall (3.6 contact attempts). However, among in-person attempts, LSF FRs successfully made contact with sample unit members more often, on average, compared to Production (3.3 contacts vs. 2.0 contacts, respectively).

Figure 5.4.1.1 Average counts per case from contact attempt history data (LSF compared to CED Production)

	<i>Complete</i>	<i>Type A</i>	<i>All</i>
No. of LSF cases	838	934	1,772
No. of Production cases (CED)	2,614	2,444	5,058
Total no. of contact attempts			
<i>LSF</i>	5.9	4.7	5.3
<i>Production</i>	5.0	5.0	5.0
Total no. attempts to placement outcome of the diary ²⁹			
<i>LSF</i>	2.9	4.2	3.6
<i>Production</i>	2.0	3.9	2.9
Total no. of midweek contact attempts (attempts made between days 3 and 8 after initial placement)			
<i>LSF</i>	0.5	0.3	0.4
<i>Production</i> ³⁰	0.2	0.1	0.2
Residual no. of contact attempts (after attempts to placement outcome and midweek attempts) to final disposition			
<i>LSF</i>	2.5	0.3	1.3
<i>Production</i>	2.7	1.1	1.9
Total no. of in-person attempts			
<i>LSF</i>	3.6	3.6	3.6
<i>Production</i>	3.6	3.6	3.6
Total no. contacts made with sample unit member			
<i>LSF</i>	3.3	1.3	2.3
<i>Production</i>	2.0	2.0	2.0

*See Appendix A for assumptions behind Figure 5.4.1.1.

We further were interested in determining the impact of incentives on enhancing the contact effectiveness and overall participation. We compared the average number of contact attempts between LSF respondents who received the \$5 cash incentive to those who did not. This allowed

²⁹ Placement outcome for Type A's include completed placement dates.

³⁰ Midweek contact attempts for Production are calculated the same way as LSF midweek contacts. (For more information, see Appendix D.)

us to test our hypothesis that incentives may make the respondent more receptive to participating in the survey, or at least more receptive to the initial contact.

Among total completed cases, 412 respondents received the incentive while 423 did not. Incentives appear to have had a minimal and non-significant impact on the number of contact attempts required to reach respondents. LSF cases receiving the incentive required 5.8 attempts on average compared to 5.9 attempts for the control group. The number of in-person contact attempts was also similar, with the incentive group requiring only 3.5 attempts on average compared to the control group's 3.7 attempts, as was the number of contact attempts required to place the diary (3.0 attempts vs. 2.9 attempts, respectively). Furthermore, we looked at midweek contact attempts, which were contacts that FRs were supposed to make on day 3 and 8 of the diary keeping period. There was no discernable difference in the number of midweek contact attempts or successful contacts made with the sample unit for either incentive or control groups. Ultimately, the incentive had no impact on gaining respondents participation in the LSF.

5.4.1.2. LSF incentive group contact attempts vs. LSF control group*

	<i>Complete</i>	<i>Type A</i>	<i>All</i>
No. of LSF cases	838	934	1,772
No. of LSF cases receiving \$5 cash incentive	412	467	879
No. of LSF cases not receiving \$5 cash incentive (Control group)	423	466	889
Total no. of contact attempts			
<i>LSF cases receiving incentive</i>	5.8	4.7	5.2
<i>LSF cases not receiving incentive</i>	5.9	4.8	5.3
Total no. attempts to placement outcome of the diary			
<i>LSF cases receiving incentive</i>	3.0	4.2	3.6
<i>LSF cases not receiving incentive</i>	2.9	4.2	3.6
Total no. midweek contact attempts			
<i>LSF cases receiving incentive</i>	0.5	0.3	0.4
<i>LSF cases not receiving incentive</i>	0.5	0.3	0.4

Residual no. contact attempts (after Successful placement and midweek attempts) to final disposition			
<i>LSF cases receiving incentive</i>	2.4	0.2	1.2
<i>LSF cases not receiving incentive</i>	2.5	0.3	1.4
Total no. of in-person attempts			
<i>LSF cases receiving incentive</i>	3.5	3.6	3.5
<i>LSF cases not receiving incentive</i>	3.7	3.5	3.6
Total no. contacts made with sample unit member			
<i>LSF cases receiving incentive</i>	3.3	1.3	2.2
<i>LSF cases not receiving incentive</i>	3.3	1.3	2.3

*Appendix A for assumptions behind Figure 5.4.1.2

5.4.2 CU doorstep concerns

In Figure 5.4.1.1, we found that the LSF sample required more contact attempts on average compared to regular Production. In addition to identifying this outcome, we referred to doorstep concerns to help explain why LSF field staff may have had more difficulty in obtaining cooperation from respondents regardless whether cooperation was obtained. The Contact History Instrument (CHI) provides CE field staff with the opportunity to record any doorstep concerns that may be expressed by respondents recruited to participate in the CED. The list of possible doorstep concerns displayed in the instrument are derived from the Census Bureau’s predefined list of statements most commonly made by respondents during survey recruitment that explain why they do not participate in the survey. In total, there are seventeen indicators field staff can use to provide an explanation for what respondents are thinking when they agree (or refuse) to participate in the CED (although not mutually exclusive). In line with prior analyses³¹, we grouped these indicators into seven similar categories. For instance, if CE field staff recorded “6-survey voluntary”, “7-privacy concerns”, “8-anti-government” or “9-does not understand survey” then these would be grouped into “Survey voluntary/privacy”.

³¹ We used the same grouping method used by Kopp and colleagues (2013 "An Exploratory Study on the Association of Doorstep Concerns with Three Survey Quality Measures for the CE Interview Survey")

Figure 5.4.2.1 Comparison of CU doorstep concern themes (not mutually exclusive) by LSF incentive group and CED Production*

<i>Doorstep theme**</i>	<i>LSF Incentive vs. LSF Non-Incentive</i>		<i>LSF Total vs. CED Production</i>	
	<i>LSF Incentive (N=1,969)</i>	<i>LSF Control (N=2,010)</i>	<i>LSF Total (N=3,988)</i>	<i>CED Production (N=10,254)</i>
Intends to quit survey	2%	2%	2%	2%
Not interested/hostile	11%	9%	10%	12%
Time	20%	23%	21%	20%
Survey voluntary/privacy	14%	16%	15%	16%
Gatekeeping	4%	3%	4%	3%
Other	10%	9%	9%	6%
No concerns***	55%	55%	55%	57%

*Among those contacted by an FR regardless of interview completion status

** Please refer to Appendix B for the full grouping of the doorstep concerns

***Coded to indicate no doorstep concern indicated across all CU contacts

We examined the distribution of doorstep concerns by grouped themes for LSF (total), LSF incentives and control, and CED Production. According to Figure 5.4.2.1, there was little difference in the reported doorstep concerns between all four test groups. Between the LSF incentive and control groups, LSF respondents receiving the incentive were less likely to exhibit concerns about “Time” and “Survey voluntary/privacy” (20 vs. 23 percent, respectively and 14 vs. 16 percent, respectively). On the other hand, respondents receiving the \$5 cash incentive were more likely to express the concern of “Not interested/hostile” compared to the control group (11 vs. 9 percent, respectively). Meanwhile, when comparing the LSF respondents to overall CED production, FRs reported “time” and “gatekeeping” concerns at a higher rate than in regular production (1 percent higher, respectively). However, the most noticeable difference arose from “other” doorstep concerns. FRs reported that 9 percent of LSF respondents exhibited other doorstep concerns compared to only 6 percent of regular CED respondents. “Other”

concerns include doorstep concerns that do not fall into any other categories. FRs are able to write-in these concerns in the CHI instrument, which arise from unique circumstances.³²

5.4.3 Effectiveness of Midweek Calls on Entries, Logins, and Participation

Field representatives were tasked with calling respondents throughout the middle of the collection period to remind them to complete the survey, check in to see how the diary process was going, and to see if the respondent had any questions. FRs were instructed to make midweek calls on day 3 and day 8. For respondents taking the online diary, FRs could also monitor the respondent’s expenditures and logins in a report accessible through MCM.

Respondents were sent midweek email reminders to complete the survey as well. All online respondents that provided an email during placement were sent an automated welcome email within 24 hours of the diary placement and on day 8, if there had been no successful logins. A total of 259 online diary users (69.6%) received a welcome email shortly after diary placement and 97 (26.1%) received a Welcome and a Reminder email on or around Day 8.

In this section, we focus on midweek calls and analyze the effect of the midweek call protocol on diary entries, online logins, and participation.

Figure 5.4.3.1. Distribution of overall contact attempts per CU by contact method (N=838 “Complete” cases)

<i>Contact Type</i>	<i>N</i>	<i>Percent</i>
Personal Visit	3,018	61.4
Telephone (Outgoing)	1,569	31.9
Telephone (Incoming)	327	6.7
Total	4,914	100.0

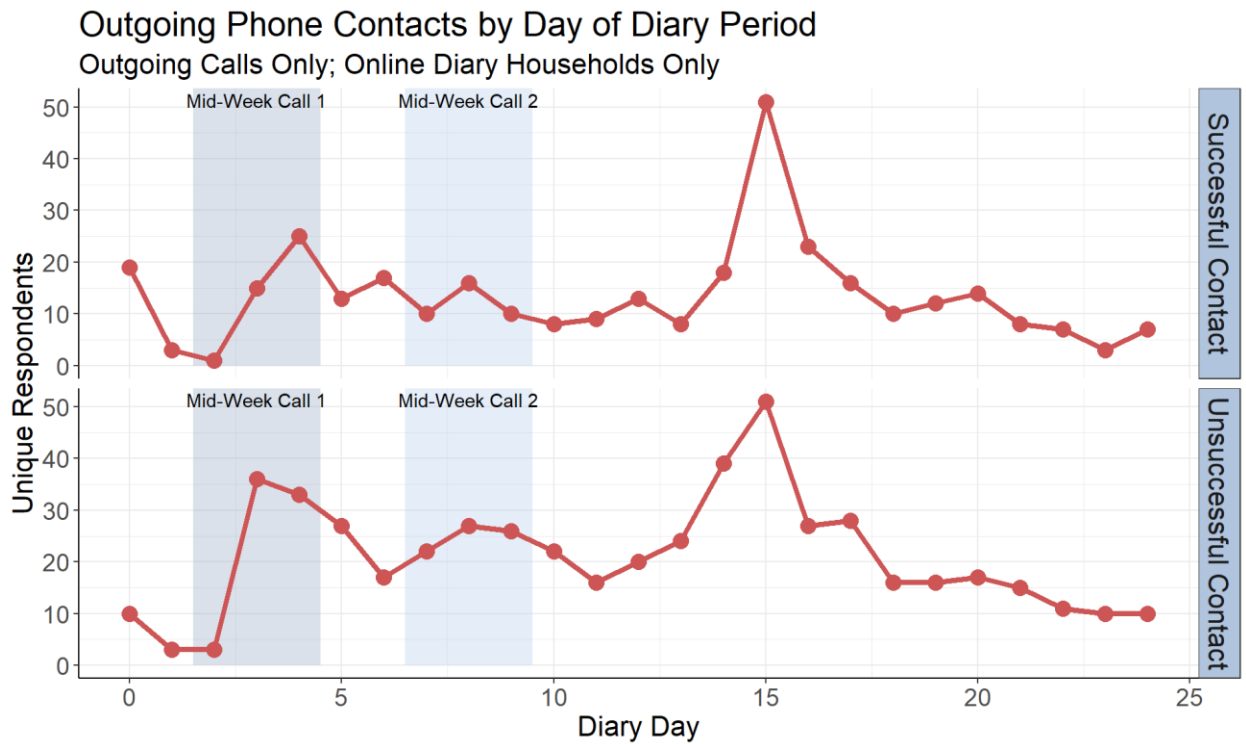
Figure 5.4.3.1 presents the overall distribution of contact attempts per CU by contact type, regardless of mode. Following the LSF protocol, we analyzed data from the 1,569 outgoing telephone calls made. From these calls, 320 were made during the midweek contact period (3rd and 8th day of the collection period) amongst 272 CUs.

Using the Contact History Instrument (CHI), we can determine if and when these contacts occurred and we can align those contacts with the paradata for online diary respondents to see if

³² BLS does not currently receive “other” concerns that are written-in by Census FRs during data collection.

there were logins within close proximity of those contacts. In Figure 5.4.3.2, which shows outgoing phone contacts, there is a spike within the window indicating the first midweek call and a smaller spike in the window for the second midweek call for successful contacts. The pattern is similar for unsuccessful contacts, with the spike for the first midweek call appearing earlier. The spikes at day 0 and around day 15 are related to placement and pickup visits. Note pickup visits could occur up until 10 days after day 15 and diary respondents could continue entering data for the diary keeping period during that window, which explains why logins continue past day 15.

Figure 5.4.3.2. Outgoing Phone Contacts from FR by Diary Day



We further examined the effect of midweek call attempts on the login behavior of online respondents (Figure 5.4.3.4). In this analysis, we used the survey paradata to infer the login activity of online respondents before and after receiving a midweek call.

Figure 5.4.3.4. Effectiveness of midweek call attempts in obtaining additional logins among online respondents* (N=384)

<i>Contact status</i>	<i>N (Respondents)</i>	<i>Mean Logins</i>	<i>Mean Number of Diary Days with Activity</i>	<i>Number of Rs with 0 Successful Logins</i>	<i>Mean Time in Diary (in mins)</i>
With successful contact attempt	138	7.3	5.2	3	37.6
No successful contact attempt	246	7.4	4.7	9	33.6

*For more information about the metrics in Figure 5.4.3.4, see Section 7

There was no difference in the average number of logins between those who received a midweek call and those who did not. Those who were successfully contacted interacted with the diary on more diary days, were able to successfully log in at least once at a higher rate, and spent more total time in the diary. Midweek calls may have served to remind respondents to take their time with entering subsequent expenditures, entering them as they occur, and to use the proper reference materials that improve overall data quality (user guide, YouTube videos, diary examples, receipts etc.).

We computed the number of successful midweek calls for the LSF at the CU-level. Although the protocol was to contact respondents on days 3 and 8 of the collection period, we included the days in-between to provide a more detailed account of outgoing phone calls. Cases with multiple contact attempts would be counted as 1 attempt. Approximately 272 of the 838 participating CUs from the October through February collection period, or 32.5 percent, received a successful midweek call. While evaluating these results, it's important to consider that there is no precise variable for midweek contact attempts. While FRs can write the purpose of the contact attempt within the CHI instrument, it still may not provide an exact explanation for their call or visit. For the purpose of the analysis, we assume that all contacts made within the timeframe described were for the goal of conducting the midweek calls.

Figure 5.4.3.3. Effectiveness of midweek call attempts on number of diarist entries (N=838)

<i>Contact Status</i>	<i>N</i>	<i>Mean Entries</i>	<i>Median Entries</i>	<i>Min</i>	<i>Max</i>
With contact attempt	272	49.5	39.0	0	251
No contact attempt	566	48.0	39.0	0	273

To examine the effects of midweek call attempts on respondent’s behavior, we compare the average number of entries between respondents who were contacted (regardless of whether they picked-up) and those who were not (Figure 5.4.3.3). Those who were contacted with a midweek call were found to have recorded approximately 1.5 more entries on average than those who were not contacted with a midweek contact, though the result was not statistically significant.³³

6 RESPONDENT EXPERIENCE

6.1 DIARY BURDEN

Respondents were asked a series of questions about burden of the Diary that were modeled after the questions asked in the CE Interview Survey. These questions attempt to draw different dimensions of burden: subjective burden, sensitivity, difficulty, and length. These responses only reflect the most cooperative respondents that completed the diary and final interview, so their interpretation is limited to that group. It’s reasonable to assume that respondents who were more burdened would be more likely to drop out of the survey and therefore excluded from being asked the burden question. Table 6.1.1 shows the indication of the first three different dimensions of burden for respondents (burden, difficulty, and sensitivity). Overall, a little over half of the respondents (55 percent) that answered the question reported at least some burden, with online diary respondents reporting 58 percent and paper diary respondents reporting 52 percent, a difference that was not statistically significant^{34,35}.

A lower percentage of respondents reported at least a little difficulty when asked how difficult was to complete the diary. Respondents that kept a paper diary were more likely to report

³³ Wilcoxon Two-Sample test shows no significant differences between midweek call group and the non-midweek call group at the .05 level of significance. Pr>Z 0.4164

³⁴ All tests of significance were conducted excluding missing values because of the strong association between mode of diary and mode of debriefing section, which were also strongly associated with the rate of missingness for the debriefing questions. FR-administered debriefings generally had a higher rate of missingness, likely because they are trained to use “CTRL-D” and “CTRL-R” when a respondent doesn’t know or isn’t willing to provide the answer. Self-administered participants were allowed to skip questions; however, were not explicitly given the option.

³⁵ Fishers exact test (online vs paper), Two-sided Pr <= P: 0.6467

difficulty than those that kept an online diary 42 percent versus 30 percent, Figure 6.1.1); however this result could be driven by differing characteristics of online versus paper respondents since online diary respondents were more likely to be younger, completed higher levels of education, and have larger CUs. Respondents that answered that keeping a diary was at least somewhat difficult were asked to elaborate what was difficult. These responses were coded into categories and are provided in Figure 6.1.2. Most of these responses related back to the amount of time that was required or the respondent’s busyness, which was fairly evenly split between online (n=13) and paper respondents (n=19). The amount of burden and detail required was the next most frequent reason, similarly split between online (n=14) and paper respondents (n=13). Reasons for difficulty cited exclusively for paper respondents were related to issues specific to the respondent that were not related to the task at hand , such as identifying themselves as too old, having health issues, or lacking the ability to write (n=22). Reasons for difficulty cited that were primarily specific to online diaries that made respondents report difficulty were related to technical issues, such as login problems and access issues (n=11).

When asked to think about how sensitive the items recorded in the diary were, 41 percent of respondents reported that they were at least a little sensitive. There was no statistical difference between the online and paper diary respondents on this topic reporting 40 percent and 41 percent, respectively.

Figure 6.1.1 Respondent reported burden measures - overall and by mode

<i>Response</i>	<i>Overall</i>		<i>Online Diary</i>		<i>Paper Diary</i>	
	<i>n</i>	<i>Percent</i>	<i>n</i>	<i>Percent</i>	<i>n</i>	<i>Percent</i>
<i>At least a little burdensome</i> ³⁵	716	60	334	61	382	60
<i>At least a little difficult</i> ³⁶	714	36	335	30*	379	42*
<i>At least a little sensitive</i> ³⁷	707	41	334	40	373	41

* Indicates a significant difference in the distribution of online and paper respondents at the 99% level

³⁶ Fishers exact test (online vs paper), Two-sided Pr <= P: 0.0005

³⁷ Fishers exact test (online vs paper), Two-sided Pr <= P: 0.9390

Figure 6.1.2 Reasons for rating at least a little difficult

<i>Coded Response</i>	<i>Overall</i>	<i>Online Diary</i>	<i>Paper Diary</i>
	<i>n</i>	<i>n</i>	<i>n</i>
Busy/Time	32	13	19
Burden/Detail	27	14	13
Respondent issues	22	0	22
Other	18	4	14
Tech issue	12	11	1
Language Issue	8	1	7

The last burden dimension measured was the respondent’s perception of time required to complete the diary. The respondents were asked to estimate on average how much time was spent per day recording entries into the diary, shown in Figure 6.1.3. On average, respondents reported about 10 minutes per day recording entries into their diaries. However, the median time was lower, 5 for paper diaries and 7.5 for online diaries. As a follow up to the question asking about the amount of time spent, respondents were asked whether they thought the amount of time was short or long. Overall, less than 10% of respondents rated the amount of time as somewhat long or very long with no statistical difference between respondents keeping an online diary or a paper diary (Figure 6.1.4).

Table 6.1.3. Reported average daily time spent - overall and by mode

<i>Mode Type</i>	<i>N</i>	<i>Mean</i>	<i>95% Confidence interval</i>	<i>Median</i>
Overall	670	9.9	[9.0, 10.8]	7
Online Diary	324	10.0	[8.8, 11.3]	7.5
Paper Diary	346	9.8	[8.6, 11.1]	5

Figure 6.1.4. Respondent reported time burden measures - overall and by mode

<i>Response</i>	<i>Overall</i>		<i>Online Diary</i>		<i>Paper Diary</i>	
	<i>n</i>	<i>Percent</i>	<i>n</i>	<i>Percent</i>	<i>n</i>	<i>Percent</i>
Somewhat long/very long ³⁸	695	9%	331	9%	364	10%

³⁸ Fishers exact test (online vs paper), Two-sided Pr <= P: 1.0000

6.2 RESPONDENT BEHAVIORS

As part of the respondent debriefing, all respondents regardless of diary mode were asked questions about their behaviors during the diary keeping week. The first question was asked to CUs that had more than one member and related to proxy reporting in the diary. There were 538 CUs that were asked the question, with 499 providing an answer to one of the response options, which explicitly included an option for “Don’t know”. Figure 6.2.1 shows the responses for how expenses were entered for other CU members. The majority of responses overall indicated that the respondent entered expenses on behalf of the other CU members (71 percent); however there were differences in responses by mode of diary. Online diary respondents were more likely to enter expense on behalf of other CU members with 79 percent compared to 63 percent for paper diary respondents. In contrast paper diary respondents were more likely to have other CU members enter their own expenses with 12 percent compared to 4 percent of online diary respondents. This is a notable finding and could be the result of having easy access to a paper diary compared to an online credentialing process. One matter of concern is that 8 percent overall (43 respondents) reported that expenses were not entered for other CU members. This result was higher for paper diary respondents (11 percent) compared to online diary respondents (6 percent), though this result was not significant. This finding suggest that there is underreporting of expenditures in diaries, the extent of which would depend on how much other CU members are spending. Based on this finding, CE should investigate how to best elicit expenditures for other CU members and may consider specific prompts or reminders multi-member CUs.

Figure 6.2.1. How expenses of other CU members were entered

<i>Method of entry</i>	<i>Overall (n=499)</i>	<i>Online Diary (n=248)</i>	<i>Paper Diary (n=251)</i>
	<i>Percent</i>	<i>Percent</i>	<i>Percent</i>
Respondent entered ³⁹	71	79*	63*
CU member(s) entered ⁴⁰	8	4*	12*
Expenses not entered ⁴¹	9	6	11

³⁹ Fishers exact test (online vs paper), Two-sided Pr < P: 0.0001

⁴⁰ Fishers exact test (online vs paper), Two-sided Pr = P: 0.0029

⁴¹ Fishers exact test (online vs paper), Two-sided Pr = P: 0.1101

<i>Method of entry</i>	<i>Overall (n=499)</i>	<i>Online Diary (n=248)</i>	<i>Paper Diary (n=251)</i>
	<i>Percent</i>	<i>Percent</i>	<i>Percent</i>
No expense from other CU members ⁴²	8	7	10
Don't Know ⁴³	9	7	11

*Indicates a significant difference in the distribution of online and paper respondents

The other set of behavior questions were asked of all respondents and related to whether the respondent's spending behaviors had changed as a result of keeping the diary (Figure 6.2.2.). One drawback of using a diary to collecting spending data is that the act of keeping a diary may influence to variable of interest that is spending. Findings suggest that the majority of respondents (74-77 percent) did not change their behavior. Respondents that did change their behavior were asked the follow up question of why they changed their shopping or dining habits during the two weeks the diary was maintained (Figure 6.2.3). Of those that were asked and provided an answer (295 respondents), there were several cases that indicated that their behavior changed directly as a result of keeping a diary. Namely, just under 10 percent of responses noted that their habits changed because the diary made them more aware of spending and just under 3 percent of responses changed because of the process required for entering items. Encouragingly, most indicated that their changed spending behavior was due to either different CU needs during the two weeks or some other reason⁴⁴.

Figure 6.2.2. Spending behavior while diary keeping

<i>Question:</i>	<i>Response options</i>	<i>Overall</i>	<i>Online</i>	<i>Paper</i>
	N	712	335	377
During the two weeks of diary keeping, did you - eat out... ⁴⁵	Less often than usual	15%	15%	16%
	About the same amount	76%	75%	77%
	More often than usual	8%	10%	7%
	N	711	333	378
	Less often than usual	17%	15%	19%
	About the same amount	77%	78%	77%

⁴² Fishers exact test (online vs paper), Two-sided Pr = P: 0.2592

⁴³ Fishers exact test (online vs paper), Two-sided Pr = P: 0.4294

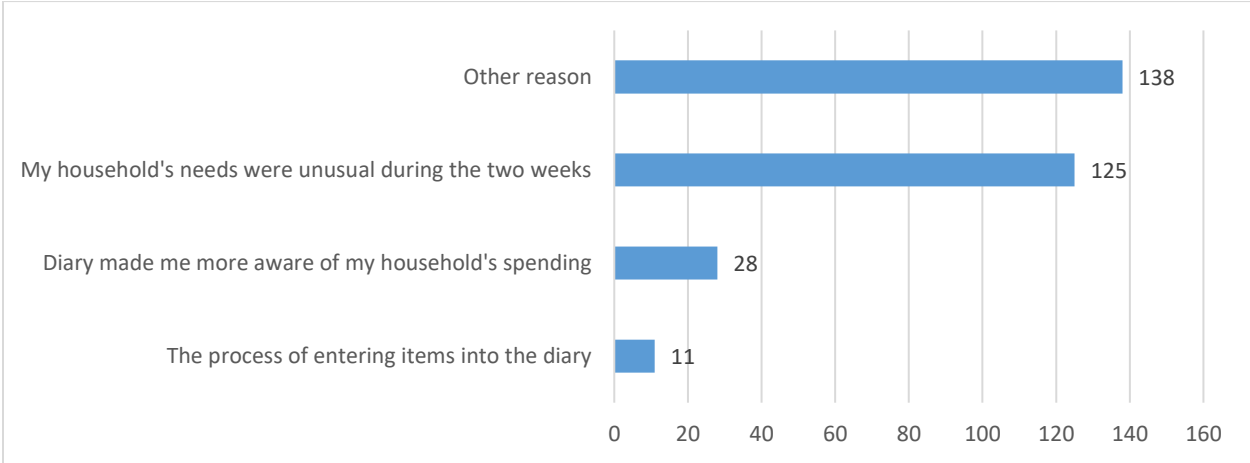
⁴⁴ Other reasons cited were mostly exogenous to the Diary keeping process, such as holidays or vacation.

⁴⁵ Online vs Paper Pearson $X^2(2) = 2.4848, p = .2887$

<i>Question:</i>	<i>Response options</i>	<i>Overall</i>	<i>Online</i>	<i>Paper</i>
During the two weeks of diary keeping, did you – shop for food... ⁴⁶	More often than usual	5%	7%	4%
	N	709	331	378
During the two weeks of diary keeping, did you – shop for nonfood items... ^{47*}	Less often than usual	16%	12%	20%
	About the same amount	74%	74%	75%
	More often than usual	9%	14%	6%

*Indicates a significant difference in the distribution of online and paper respondents at the 99% level

Figure 6.2.3. Reasons for change in spending (n=295, not mutually exclusive)



⁴⁶ Online vs Paper Pearson $X^2(2) = 4.0320, p = .1332$

⁴⁷ Online vs Paper Pearson $X^2(2) = 19.1722, p < .0001$

7 ONLINE DIARY USE

7.1 LOGINS AND LOGIN ISSUES:

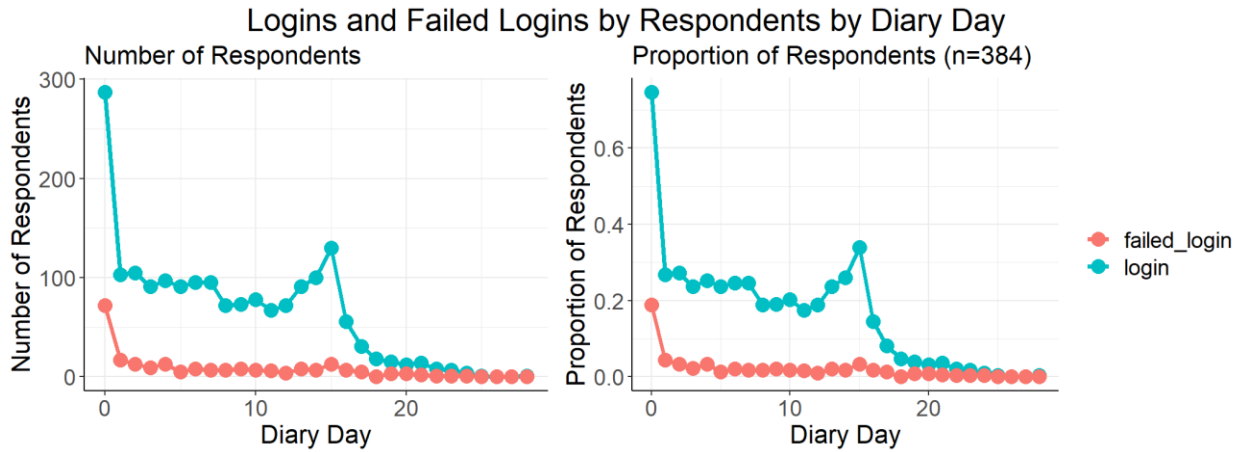
In order for respondents to report in the online diary, they have to log in. As opposed to the paper diary, entering data into the online diary requires (1) opening a computer or mobile device, (2) navigating to the Census Centurion website, and (3) correctly entering one's credentials.

Combining the October through February data, a total of 445 CUs were placed with an online diary. The paradata from the online diary for the same time period has information on 384 CUs, therefore 61 CUs are assumed to have never logged in or attempted to login even though they were placed with an online diary. Of these 384 CUs with paradata from the online diary, 12 CUs attempted to login but were never successful during the diary period. The remaining 372 CUs successfully logged in at least once.

There were 372 CUs that successfully logged in. Overall, there were a total of 2,825 successful logins which were from those 372 CUs. There were also 478 failed logins, including the 12 CUs who only failed to login. The average number of successful logins per respondent was 6.3 (including users that never attempted to login), the maximum number of successful logins for an individual respondent was 50 and the minimum was 0.

Figure 7.1.1 shows the number and proportion of users with successful and unsuccessful logins by Diary Day (i.e., number of days since diary placement). The online diary protocol stressed that the FR attempt to log in during the placement interview with the respondent to show them the diary. In the figure, these appear as logins on Diary Day 0. There were 287 respondents (74.7 percent) who successfully logged in on Diary Day 0, suggesting that the protocol was followed well. Following Diary Day 0, 321 (83.6 percent) respondents logged into the online diary while 63 respondents (16.4 percent) never successfully logged in after Diary Day 0. The number of failed logins was high on day 0, but subsequently remained very low through the rest of the diary period. The figure also shows logins after the diary keeping period which could be due to data being entered later (by recall) by the respondent or data entered during the pickup interview.

Figure 7.1.1 Logins during the diary keeping period



Looking at the first attempt to login, 302 respondents (78.6 percent) were successful and about 70.2 percent of those respondents had no subsequent failed logins. Eighty-two respondents (21.4 percent) had a failed first attempt at logging in, of these, the majority were able to login later, but 12 respondents (14.6 percent) of those 82 respondents never managed to login successfully.

Also, despite having the ability to customize their username and password in the online diary, very few respondents attempted to update/customize their username or password. A total of 19 users updated their password and 4 updated their username in the October through February data. The diary was not designed to allow passwords to be saved; however, it is possible that they could have done so based on the system or browser they were using and this could be a factor in why there were so few changes to passwords.

7.2 DEVICE USE

The paradata also provides some information on which devices were used by respondents⁴⁸. The online diary was device optimized, meaning it was designed to render on different screen sizes to provide flexibility for respondents and enhance contemporaneous reporting. Of the respondents who were able to login successfully, 162 respondents (43.5 percent) used desktop only, 134 (36

⁴⁸ The mobile version can be presented on a desktop computer (and vice versa). The key determinant is the pixel width of the screen. Devices with a screen width less than 576px are shown the mobile version and those with a screen width greater than or equal to 576px are shown the desktop version.

percent) used mobile only and 76 (20.4 percent) used both desktop and mobile devices. There were a higher proportion of failed first login attempts for desktop only users (21 percent) and mixed users (19.7 percent), compared to mobile only users (15.7 percent)⁴⁹. One possible explanation could be that respondents were able to save passwords more easily on their mobile phones. The average number of logins for desktop only and mobile only users were similar at around 7. The average number of logins for those who used both types of devices was slightly higher. This is consistent with our expectation that those who used both types of devices would report more contemporaneously on their expenses.

Figure 7.2.1 Device used by LSF online diary users (n=372)

<i>Device type</i>	<i>Respondents</i>	<i>Average number of logins</i>	<i>Failed first logins</i>
Desktop only	162	7.19	34
Mobile only	134	7.09	15
Desktop and mobile	76	9.05	21
Total logged in	372	-	-
Never logged in	12	-	12

The most common operating systems used were Windows and iPhone. The main operating systems used are shown in Figure 7.2.2.

Figure 7.2.2. Operating systems used by LSF online diary users (n=372)

<i>Operating system</i>	<i>Frequency</i>
Windows	92
iPhone	80
Android	59
Macintosh	41
Android and Windows	24
iPhone and Windows	23
iPhone and Macintosh	12
Other combinations	41

⁴⁹ For failed logins, we can only classify what device they used if they had a successful login at some point, in which case the device used for the successful login is what was counted.

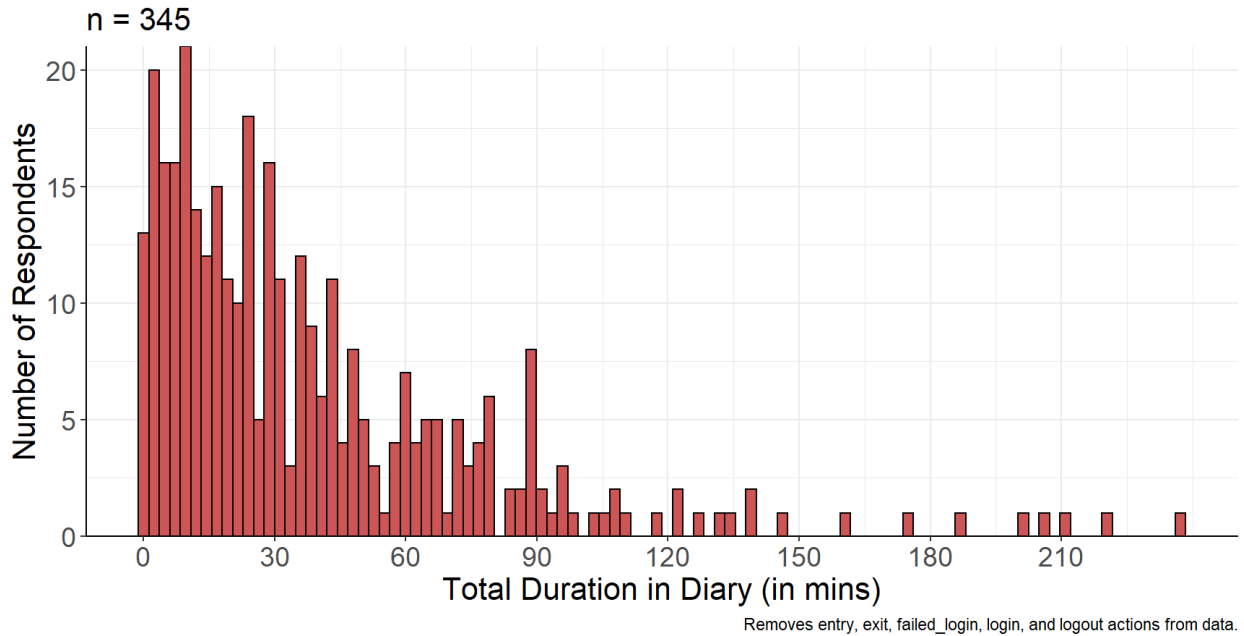
The online diary was only available in English. Only a small number of users had the language set in their browser to anything other than English. These were Spanish, Korean and Chinese.

7.3 TIME SPENT IN THE DIARY:

Measuring time spent in the diary required making certain assumptions and dropping some events in order to get a good estimate. Prior to calculating total time spent in the diary, certain events such as entry, exit, failed login, login, and logout were dropped and anything that happened on certain pages such as the login or post login form were dropped. These changes include trying to account for people who closed out of their browser without logging out (they may have had 15 extra minutes of time before the system logged them out) by not counting the amount of time between the last event and their logout/exit from the instrument. Time spent on the post login screen where respondents select their start date was not counted. Time spent on the login screen was also not included because there is no preceding event so we don't know when they landed on the login screen, only when they clicked the login button. The resulting time estimate should capture time spent entering expenses, rather than time spent logging in and out, which is of interest here.

As shown in Figure 7.3.1, the average time spent entering expenses in the online diary was 39.0 mins and the median was 27.6 mins. The maximum time spent was about 4 hours.

Figure 7.3.1. Time spent in the Online Diary (n=345)



7.4 YOUTUBE VIDEOS

To supplement the online diary instructions, there were three YouTube videos produced and hosted on the Census Bureau’s YouTube page:

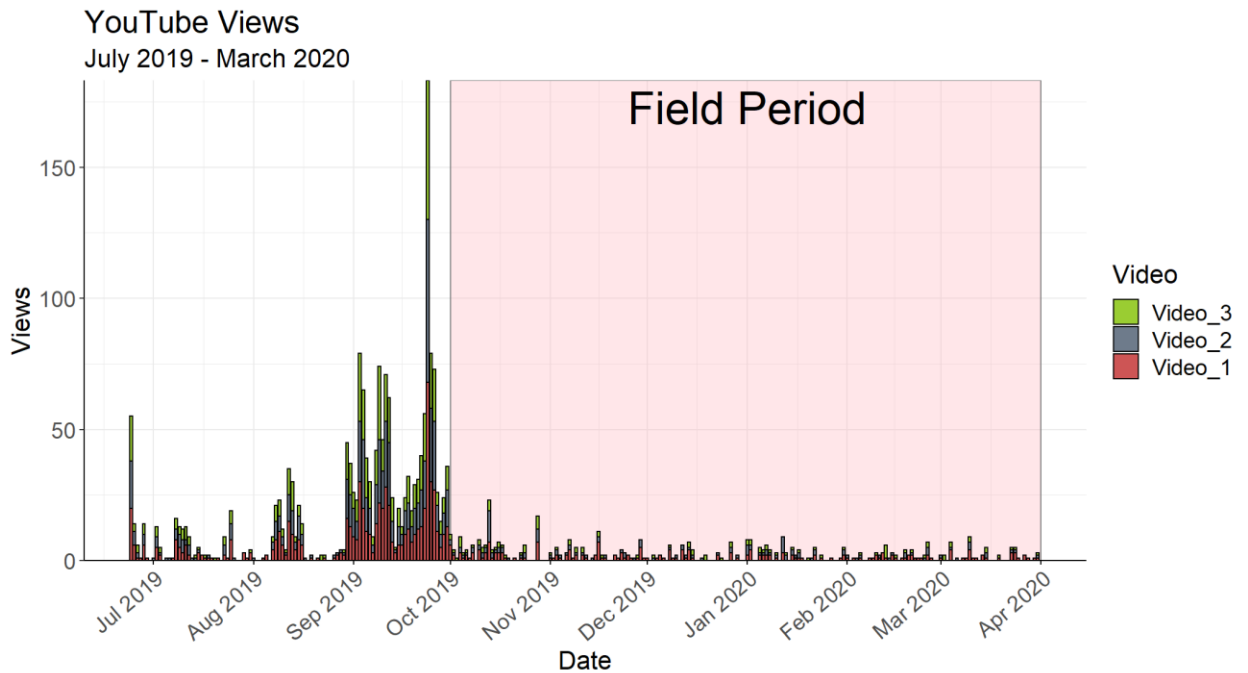
- Getting Started (Video 1)
- Entering and Editing Expenses (Video 2)
- What to Enter in the CE Diary (Video 3)

These videos were linked in the instructions and FRs were instructed to reference them during placement. In order to better understand whether these videos were used, we requested that Census provide us with YouTube analytics data. These data cannot be linked with either expenditures or paradata from the LSF, so we can only look at aggregates.

There were a total of 2,255 views for all three videos as of the end of March 2020. There were 889 views for *Getting Started*, 737 views for *Entering and Editing Expenses*, and 629 views for *What to Enter in the CE Diary*. Of those views, 444 views were during the diary period and 1,811 views were before the diary period. Views before the diary period were likely Census staff (including FRs) or BLS staff reviewing the videos.

The plot (Figure 7.4.1) below shows daily views for all three videos across the full field period for which there are analytics data. The highlighted area indicates the field period for the LSF.

Figure 7.4.1. Number of YouTube Views of LSF video July 2019-March 2020

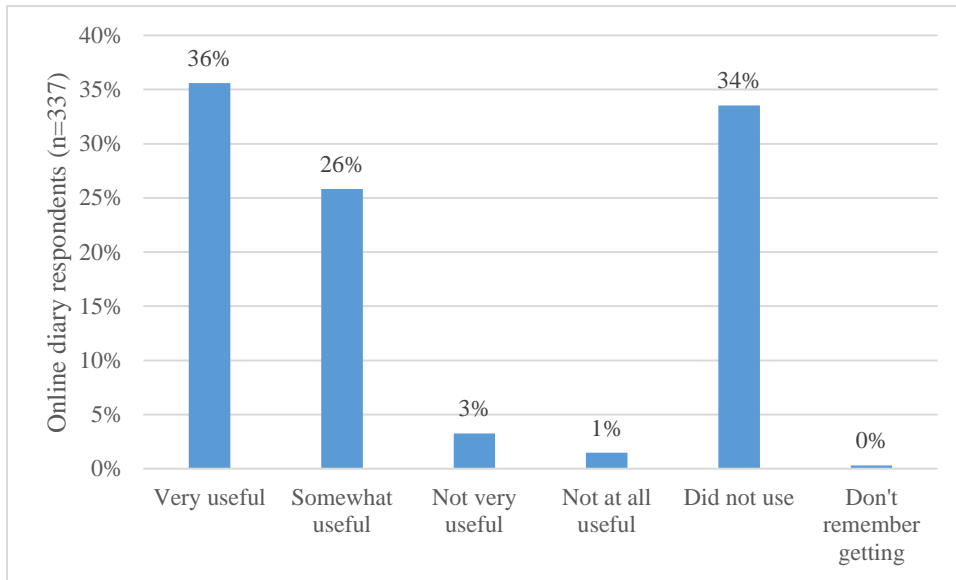


7.5 ONLINE DIARY RESPONDENT EXPERIENCE

In the respondent debriefing, respondents who were placed with an online diary were asked a set of questions designed to capture feedback on specific features of the online diary such as the respondent materials and the user help desk.

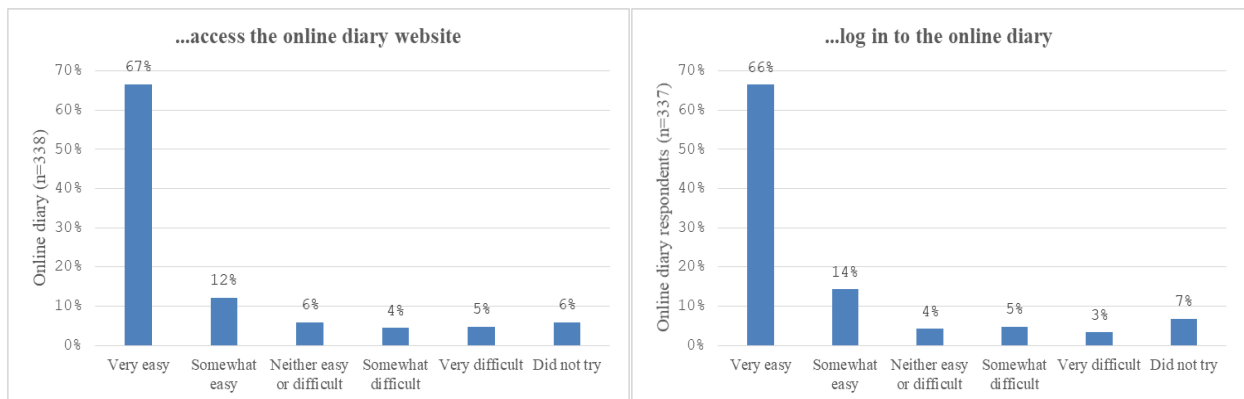
The first of these questions asked the respondents about how useful the user guide that was provided at the start of the diary keeping week was. The majority of users found it useful or very useful (62 percent). Most of the remaining users did not use the guide. When asked the reason, they cited reasons such as “self-explanatory” and “did not need it” suggesting that the online diary was programmed in a way that didn’t require the user guide. Among the 4 percent of respondents that found the user guide not very helpful or not at all helpful, there were only a handful of comments, including “It seemed a little redundant (sic)” and “not clear” and “too complicated”. Based on the very limited negative feedback, it seems that the user guide was effective as was delivered and no changes are needed.

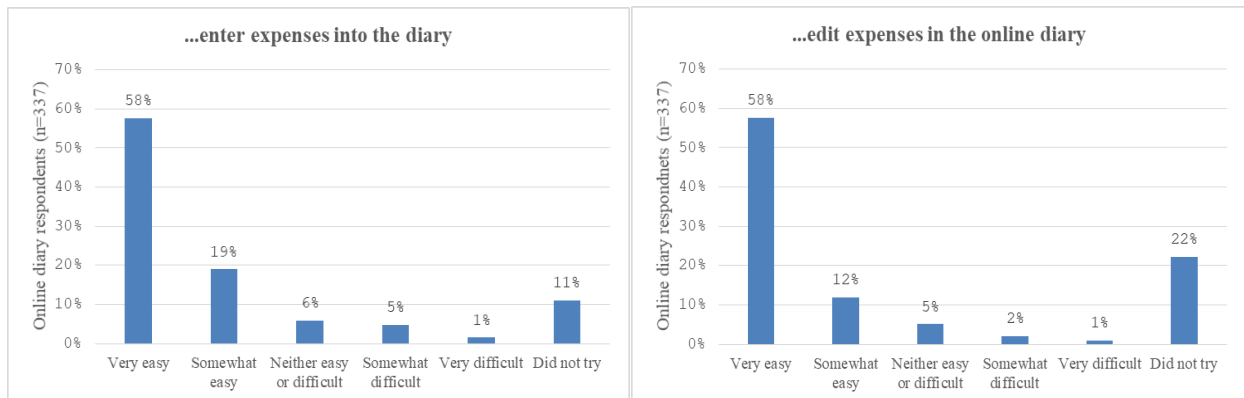
Figure 7.5.1. Usefulness of user guide as reported by online diary respondents (n=337)



The next series of questions asked online diary respondents how easy or difficult it was to complete different tasks required for online diary keeping, including accessing, logging in, and entering and editing expenses. Between 3 and 9 percent found these tasks at least somewhat difficult.

Figure 7.5.2. Ease of completing online diary tasks, as reported by online diary respondents (n=338)





Outside of those indicating difficulty for these tasks, there were 41 diary respondents (12 percent) that indicated that they had other problems not listed above. When looking at specific responses, these issues revolved around various sporadic access issues or other types of issues. For a full list, see Appendix E.

7.6 HELP DESK DATA

During the fielding period, Census operated a help desk dedicated to resolving technical issues – including how to access the online diary and password problems – that would inhibit respondent participation in the diary task. The help desk did not address non-technical questions about completing the CE survey, and in some cases referred issues to Census headquarters for more in-depth troubleshooting. The help desk accepted calls or emails from either respondents or FRs from Monday through Friday (9am-10pm), Saturday (9am-7:15pm), and Sunday (11am-9:15pm, Eastern time). Contacts could be routed to Census headquarters if they were received during normal business hours. Overall, there were not many contacts received during the survey period; 120 contacts were logged into the help desk database. It was unknown how many calls did not go through to the help desk (a problem some reported in FR debriefing sessions). Additionally, not all contacts could be associated with their corresponding LSF identifier. Despite those limitations, we analyzed who contacted the helpdesk, the types of problems reported, whether they were resolved, and situations preventing problem resolution.

There were a total of 105 logged contacts from FRs and respondents when excluding 15 involving non-substantive contacts (e.g., caller immediately hung up, people calling the wrong number, FR testing to see if the help desk number worked). Of these, about three-quarters (77) came from respondents with the rest coming from Census staff. Figure 7.7.1 lists the frequency of the types of problems logged by who contacted the help desk.

Figure 7.7.1. Help Desk Problem Categories – Overall and by Origin

<i>Problem type</i>	<i>Overall # instances</i>	<i>% of total</i>	<i>Respondent # instances</i>	<i>% of total</i>	<i>Field Rep # instances</i>	<i>% of total</i>
Login ⁵⁰	69	66%	55	71%	14	50%
Other	22	21%	13	17%	9	32%
Instrument	14	13%	9	12%	5	18%
Total	105	100%	77	100%	28	100%

As indicated in Figure 7.7.1, there was not a large difference in the types of problems by origin, although respondents were slightly more likely to contact the help desk with problems logging in to the diary (71 percent of their contacts).

- Login problems were overwhelmingly related to respondents either getting locked out of the online diary or having issues entering passwords, although one mentioned not having a password written down and one had trouble accessing the correct diary site.
- Problems categorized as ‘other’ were more varied. Aside from some reclassified to login, others involved questions more relevant to CE (e.g., how to record credit card purchases, how to create a shortcut to the diary), and a few were calls that ended prematurely.
- For the instrument problem type, 5 involved logistical questions about completing the diary – how to complete an online diary for the respondent, and an issue with selecting the start date.

Encouragingly, most of the problems directed to the help desk could be resolved (Figure 7.7.2).

Figure 7.7.2. Problem Resolution Status and Origin (N=105)

<i>Resolved?</i>	<i># instances</i>	<i>% of total</i>	<i>Origin</i>	<i># instances</i>	<i>% of subtotal</i>
Yes	83	79%	Respondent	59	71%
			FR	24	29%

⁵⁰ Includes 2 with problem types reclassified from ‘Password’ and ‘Refusal’, as well as others reclassified from ‘Other.’

<i>Resolved?</i>	<i># instances</i>	<i>% of total</i>	<i>Origin</i>	<i># instances</i>	<i>% of subtotal</i>
No	22	21%	Respondent	19	86%
			FR	3	14%

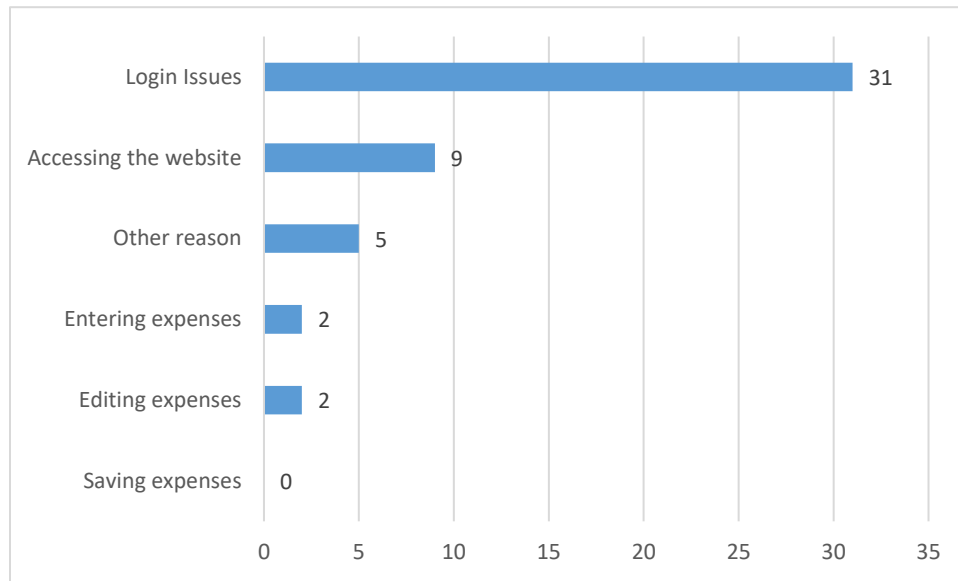
A full 79 percent of contacts logged by the help desk involved issues that were classified as resolved. Those involving respondent contacts were more predominate among unresolved contacts (86 percent compared to 71 percent for resolved issues). Examining the unresolved issues, the 3 involving FRs were redirected to Census headquarters as out-of-scope for the help desk. Among the 19 respondent issues that were unresolved: 7 were redirected to Census/FRs, 5 were disconnected calls, and the remaining 7 had a variety of issues.⁵¹

Although, as noted earlier, most logged issues could not be associated with LSF identifiers, there were 6 among those with unresolved issues that were associated with case information, and all but one ended up completing the diary (the one who didn't complete the diary logged in 2 weeks after the start date).

In the respondent debriefing, respondents were asked about their experience with calling the help desk to resolve technical issues. Forty online diary respondents reported calling the help desk. The top reason for calling the helpdesk was for login issues (Figure 7.7.3). Respondents that reported calling the help desk were asked whether the helpdesk assistant had resolved the issue and only 53 percent reported that their issue was resolved with the remaining 47 percent of issues going unresolved. As we move towards a production setting with an online diary, further training of help desk staff will be needed to ensure that they are able to resolve respondent issues.

⁵¹ Two weren't able to get their password to work, one had credentials that never worked so was told to fill a paper form, 1 forgot password and decided to refuse when the help desk couldn't resolve issue, one 'site is down,' one couldn't assist respondent who wanted to 'submit the completed diary.'

Figure 7.7.3. Reasons for calling the helpdesk (n=40, not mutually exclusive)



8 DATA QUALITY

This section investigates the comparative quality of the online diaries in terms of the number of entries recorded, preliminary examination of expenditure totals, the drop-off in weekly reporting behavior, and rates of item nonresponse. Other factors such as the respondent technological problems mentioned in the prior section, diary placement issues, and uncertainty about the diary task could account for poorer data quality among groups. A subsequent report will include analysis of quality measures such as post-processing expenditure totals, allocation and imputation rates, legitimate blank diaries, and comparisons with demographic controls.

8.1 OVERALL DIARY ENTRY COUNTS, AND COUNTS BY SECTION, BY GROUP

We compared groups by the count of entries recorded in non-blank diaries (Figure 8.1.1).

Figure 8.1.1. Diary entry count statistics, by group⁵²

Month	Production			LSF			Median Difference (prod-LSF)
	Diaries	Total (Avg.)	Total (Median)	Diaries	Total (Avg.)	Total (Median)	
October	902	32.3	27	312	24.6	20	7

⁵² The production numbers in the report may be different from what managers normally see.

<i>Month</i>	<i>Production</i>			<i>LSF</i>			<i>Median Difference (prod-LSF)</i>
	<i>Diaries</i>	<i>Total (Avg.)</i>	<i>Total (Median)</i>	<i>Diaries</i>	<i>Total (Avg.)</i>	<i>Total (Median)</i>	
November	838	30.6	26	290	26.2	18	8
December	794	33.9	29	309	28.3	21	8
January	1,207	31.4	27	323	27.5	22	5
February	1,154	33.6	28	284	27.3	20	8
Total	4,895	32.4	27	1,518	26.8	20	7

The counts per diary suggest that control group CUs may have provided higher quality data, as they had a median of 27 entries in a diary week, compared to only 20 for the LSF group, a significant difference of 7 entries⁵³. This difference was consistent across months. Examining this in more depth, we can see the differences by diary section for the two groups (Figure 8.1.2 & 8.1.3).

Figure 8.1.2. Production entry count statistics, by section

<i>Month</i>	<i>Diaries</i>	<i>FDB (Avg)</i>	<i>FDB (Med.)</i>	<i>MLS (Avg)</i>	<i>MLS (Med.)</i>	<i>CLO (Avg)</i>	<i>CLO (Med.)</i>	<i>OTH (Avg)</i>	<i>OTH (Med.)</i>
October	902	19.1	15	3.5	2	1.2	0	8.5	7
November	838	18.0	14	3.4	2	1.2	0	8.0	6
December	794	18.9	14	3.6	3	1.4	0	10.0	7
January	1,207	18.3	15	3.9	2	0.8	0	8.5	6
February	1,154	19.7	16	4.0	3	1.0	0	8.9	7
Total	4,895	18.8	15	3.7	2	1.1	0	8.7	7

Figure 8.1.3. LSF entry count statistics, by section

<i>Month</i>	<i>Diaries</i>	<i>FDB (Avg)</i>	<i>FDB (Med.)</i>	<i>MLS (Avg)</i>	<i>MLS (Med.)</i>	<i>CLO (Avg)</i>	<i>CLO (Med.)</i>	<i>OTH (Avg)</i>	<i>OTH (Med.)</i>
October	312	12.9	6.5	3.7	2	0.9	0	7.2	6
November	290	14.5	8	2.7	2	1.1	0	7.8	5
December	309	15.7	10	3.2	2	1.4	0	8.1	6
January	323	15.9	11	3.1	2	0.7	0	7.7	5
February	284	15.8	9	3.2	2	0.8	0	7.5	6
Total	1,518	15.0	9	3.2	2	1.0	0	7.7	6

⁵³ Wilcoxon Two Sample Rank-Sum Test; Z=-8.40, two-sided (p<0.0001).

Most of the difference in diary counts between the groups can be attributable to the ‘Food and Drinks for Home Consumption’ (FDB) section, in which production diarists entered a median of 15 entries compared to 9 for LSF diarists. This difference was particularly pronounced in the first month of the test, with production CUs entering a median of only 6.5 diary items. Given that the difference occurred in this section, it may have been due to less collection of items via grocery receipts in the LSF, where only about 6 percent of CUs had any CAPI entries. Although respondents are instructed to record all of their expenses in the diary during the two week period, FRs often suggest that they hold onto their receipts and will offer to record the entries for the respondent at the time of pickup. During the debriefings after the LSF was complete, we learned that not all FRs were clear on the process of entering recalled items via receipts, or had trouble entering them after ending the interview. The difference in FDB in particular, suggests that grocery receipts collected at recall may not have been entered for the LSF consistently. In other sections, the differences in number of reports were less noticeable, with one greater median entry among production diarists in the ‘Other’ (OTH) section and the same counts in the remaining ‘Meals and Drinks Away from Home’ (MLS) and ‘Clothing and Accessories’ (CLO) sections.

8.2 OVERALL DIARY EXPENDITURE TOTALS BY GROUP

We also carried out a preliminary examination of the expenditure totals recorded in diaries using unprocessed data (Figure 8.2.1).

Figure 8.2.1. Diary expenditure totals, by group

<i>Expenditure Totals</i>	<i>Production</i>			<i>LSF</i>		
	<i>Diaries</i>	<i>Average</i>	<i>Median</i>	<i>Diaries</i>	<i>Average</i>	<i>Median</i>
October	902	\$950.47	\$466	312	\$826.26	\$416
November	838	\$817.19	\$404	290	\$664.67	\$395
December	794	\$897.51	\$466	309	\$803.53	\$404
January	1,207	\$828.52	\$450	323	\$683.68	\$364
February	1,154	\$906.78	\$418	284	\$713.65	\$357
Total	4,895	\$878.69	\$438	1,518	\$739.35	\$393

Figure 8.2.1 shows that production diarists recorded a higher total amount of expenditures each week than did LSF diarists as is expected with a lower number of overall reports. The \$438 median production amount was \$45 higher than for LSF, and the average totals were \$139.34

higher, representing a significant difference⁵⁴. Differences in median expenditures ranged from only \$9 in November to \$86 in January. It is important to note that, once processed data become available, it will be possible to adjust for missing cost information that needs imputing and to control for sample demographics that may have caused the differences that were observed.

8.3 COUNTS AND EXPENDITURE TOTALS BY MODE FOR LSF SAMPLE

Participation in the LSF occurred through online or paper diaries, depending on eligibility. The statistics of the count of entries by placement mode suggested the possibility that the mode, or the participants selected into the online mode, affected diary participation (Figure 8.3.1).

Figure 8.3.1. LSF entry counts by mode (N=1,518)

<i>Mode</i>	<i>Diaries</i>	<i>Average</i>	<i>Median</i>
Paper	872	28.9	24
Online	646	23.9	16

Figure 8.3.1 indicates online diarists recorded an average of 23.9 entries per diary week, compared to 28.9 entries on average for LSF paper diarists. These differences were statistically significant⁵⁵. For reference, production (paper) diarists recorded average and median counts of 32.2 and 27 entries, respectively. The expenditure totals recorded in each mode are shown below (Figure 8.3.2).

Figure 8.3.2. LSF expenditure totals by mode (N=1,518)

<i>Mode</i>	<i>Diaries</i>	<i>Average</i>	<i>Median</i>
Paper	872	\$659.57	\$331.05
Online	646	\$847.04	\$469.30

As with counts, there was a significant difference in expenditure totals associated with the mode of diary completion⁵⁶. Average amounts for online diarists were higher than those for paper diarists, with median amounts of \$469.30 and \$331.05 respectively. Further analysis will be needed to control for the response behavior and sample characteristics that underlie these mode differences.

⁵⁴ Wilcoxon Two Sample Rank-Sum Test; Z=-3.12, two-sided (p<0.0018).

⁵⁵ Wilcoxon Two Sample Rank-Sum Test; Z=-5.52, two-sided (p<0.0001).

⁵⁶ Wilcoxon Two Sample Rank-Sum Test; Z=5.77, two-sided (p<0.0001).

8.4 REPORTING DROP-OFF OVER THE DIARY KEEPING PERIOD

Respondents may experience fatigue from the diary expenditure reporting process and report fewer expenditure entries over time. This ‘drop-off’ was measured by calculating the proportion of two-week diary expenditure entries reported in week 1 (i.e., number of expenditure entries in week 1 over number of expenditure entries reported in both weeks) for each CU that had at least 1 entry over the two-week period. Thus, it takes into account CUs that did not provide any entries in the second diary week (as well as the smaller number of CUs only providing entries in the second week). Proportions greater than 50 percent reflect a drop-off in entries across the 2 weeks, potentially signaling a reduction in data quality (Figure 8.4.1).

Figure 8.4.1. CU entry drop-off by placement mode⁵⁷ and group (N=3,112)

	<i>Online LSF</i> (N=334)	<i>Paper LSF</i> (N=446)	<i>Overall LSF</i> (N=780)	<i>Production</i> (N=2,332)
Week 1 mean entries	24.6	30.4	27.9	33.1
Week 2 mean entries	21.6	26.2	24.2	29.9
Mean entries (total)	46.2	56.6	52.1	63.0
Mean (CU-level) drop-off	54.4%	52.5%	53.3%	53.6%

As seen in Figure 8.4.1, CUs reported fewer entries in the second week than in the first, on average. For online CUs, the drop-off involved 54.4 percent of entries coming in the first week, a share that was not significantly different than the 52.5 percent among CUs placed with paper diaries⁵⁸. Regardless of placement mode, the drop-off in entries for LSF CUs and production CUs was similar – 53.3 percent and 53.6 percent, respectively⁵⁹.

8.5 ITEM NONRESPONSE

Another aspect of data quality involves the extent of diarist item nonresponse. LSF diarists were required to provide certain fields, such as date and type (i.e. ‘Food and Drinks for Home

⁵⁷ Examining by completion mode does not affect the findings.

⁵⁸ Wilcoxon Two Sample Rank-Sum Test; Z=0.73, two-sided (p=0.46).

⁵⁹ Wilcoxon Two Sample Rank-Sum Test; Z=-0.64, two-sided (p=0.52).

Consumption’ or ‘Other’), but other fields, such as cost and item description, could be left blank. Using unprocessed data, item nonresponse rates for cost were compared by group (Figure 8.5.1).

Figure 8.5.1 Item nonresponse for the cost field, by group

<i>Item Non-Response</i>	<i>Records</i>	<i>Don't Know</i>	<i>Illegal*</i>	<i>Refusal</i>	<i>Total</i>
Production	157,606	0.05%	0.46%	Negligible	0.51%
LSF	40,662	0.01%	0.41%	0%	0.42%

*Illegal Response set in a Phase 1 or Phase 2 edit

The comparison indicated that, among diaries with one or more entries, there was a significant difference in nonresponse for the cost field by group⁶⁰. In total, there were 0.51 percent of records with missing cost for production diary entries, compared to a rate of 0.42 percent for LSF.

For item description – a character field – item nonresponse was composed of a count of entries that were blank, or for which Census had supplied ‘blank.’⁶¹ Unlike for cost, when excluding observations in the “meals” section (for which respondents do not supply an item description), the rate of missing was 0.32% among LSF entries, a significantly larger rate than the 0.08% among production diary entries⁶².

Since 2019, the name of the outlet where each item entered in the diary was purchased has been collected in the CED paper diary for all expense categories except clothing. This is also true of the LSF paper diary since it was the same as the CED paper diary. The LSF online diary was designed to collect outlet name in the “business” field for all four categories of expenses. In the paper diary form, outlet fields can be left blank, but the online diary was designed to not save an expense entry unless something was entered in the outlet, date, type of expenditure, and one additional field. The paper diary, used in Production and LSF, does not have an outlet name field for the clothing category, while the LSF online diary has an outlet name field for clothing.

Looking at LSF item nonresponse for all categories except clothing, including blank, don’t knows and other invalid entries for the outlet name in the data from October 2019 to February

⁶⁰ Wilcoxon Two Sample Rank-Sum Test; Z=-2.21, two-sided (p=0.0270).

⁶¹ Census’ National Processing Center (NPC) identified expenditure with some information provided, but not the description field entered. For these, the data transmitted to BLS had the word ‘blank’ entered.

⁶² Wilcoxon Two Sample Rank-Sum Test; Z=11.69, two-sided (p<0.0001).

2020, there were 39,864 entries in the business/outlet field. Item nonresponse for the outlet field was about 7.2 percent of all entries. This is slightly lower than the 9.1 percent item nonresponse rate in Production (Figure 8.5.2).

Figure 8.5.2. Item nonresponse in the LSF and Production outlet name field*

<i>Response in Outlet Field</i>	<i>LSF (paper and online)</i>	<i>Production (all paper)</i>
Total No. of Entries	39,864	153,292
Invalid entries	2,860 (100%)	13,918 (100%)
Blank	2,631 (92.0%)	12,738 (91.5%)
DK/don't know/?	226 (7.9%)	1,174 (8.4%)
N/A or None	3 (0.1%)	6 (<0.1%)

*Clothing is excluded for both LSF and Production

As shown in Figure 8.5.2, the largest proportion of invalid entries for outlet name in the LSF and Production were blank (92 percent in the LSF compared to 91.5 percent in Production). About 8 percent of invalid responses in the LSF were don't know, DK, or a question mark similar to Production. In the LSF, most of the item nonresponse for outlet name was in paper diaries. In addition to the nonresponse in the LSF, there were many types of generic entries where the actual outlet name was not provided, including super market, store, grocery or groceries, food, gas, lunch, dinner, deli, home, hospital, taxi, web, online, doctor, hair salon, etc.

8.6 RECALLED EXPENSES

Recall was captured in two ways in the LSF. The first was in a recall tab added to the LSF CAPI instrument where FRs were supposed to enter expenses based on information given by the respondent at the time of diary pickup which had not been entered in the online or paper diary. The second protocol, which is similar to how recall is captured in CED production, involves FRs entering responses during the pickup interview to questions about diary entries, recall and the use of receipts for each week (n) of the diary keeping period.

(1) ANYENTRn: whether there were any entries in the Diary

(2) ANYRECA_n: whether any expenditures were added through recall at pickup

- If the answer is yes, the next question is ANYRECP_n: whether receipts were used for the majority of the recalled items
- If the answer is no and ANYENTR_n is no, the next questions is ANYBUY_n: whether the HH had any expenses last week

In an ideal setting, there would be no recall required and the respondents would report all of their expenditures as they make them into the diary. However, since this is an unrealistic expectation for respondents, the goal is to have minimal amounts of recalled expenditures and to encourage recall with the aid of receipts. For this analysis, we define three types of recall diaries for analysis:

- (1) Partial recall: ANYENT_n=1 and ANYRECA_n=1
- (2) Total recall with receipts: ANYENT_n=2, ANYRECA_n=1, and ANYRECP_n=1
- (3) Total recall without receipts: ANYENT_n=2, ANYRECA_n=1, and ANYRECP_n=2

Figures 8.6.1.1 and 8.6.1.2 compare the incidence of these types of recall diaries in the LSF and CED production by diary week, also including information on diaries with no purchases or missing information.

Figure 8.6.1.1. LSF recall vs. CED production recall – week 1

	<i>LSF (N=838)</i>		<i>CED Production (N=2,607)</i>	
	N	Percent	N	Percent
Purchases with no recall	640	76.4	1,978	75.7
Partial recall	72	8.6	264	10.1
Total recall w/receipts	68	8.1	213	8.2
Total recall w/o receipts	9	1.1	24	0.9
No purchases	49	5.9	129	4.9
Missing/Other	-	-	6	0.2

Figure 8.6.1.2. LSF recall vs. CED production recall – week 2

	<i>LSF (N=838)</i>		<i>CED Production (N=2,607)</i>	
	N	Percent	N	Percent
Purchases with no recall	616	73.5	1,935	74.0
Partial recall	83	9.8	257	9.8
Total recall w/receipts	70	8.4	217	8.3

	<i>LSF (N=838)</i>		<i>CED Production (N=2,607)</i>	
	N	Percent	N	Percent
Total recall w/o receipts	6	0.7	27	1.0
No purchases	57	6.8	168	6.4
Missing/Other	7	0.9	12	0.4

The results suggest that the levels of recall were fairly similar between production and the LSF; however, based on discussions with the FRs during the debriefing and inconsistencies with expenditure entries that should have been made in the recall tab in the LSF, it is unclear whether the LSF and CED production results were actually the same. For example for the LSF, there were some instances where there were expenditure entries in the recall tab despite what FRs reported at ANYRECA_n. Conversely there were even more cases where FRs indicated recall at ANYRECA_n but there were no recorded entries in the CAPI recall tab. The extent of these discrepancies in the LSF is shown in Figure 8.6.2. While there was no definitive explanation for these results, Census provided some insight on potential reasons:

- Expenses may not have been entered in the recall tab.
- The case was closed out and the FR was not able to get back into the case in order to enter saved receipts (this was mentioned during one of the FR debriefing calls).
- Expenses were not initially entered into the recall/receipts tab before the respondent found more receipts and ANYRECA_n was not subsequently updated.
- Respondents may have been asked to enter recalled expenses into the online diary.

Figure 8.6.2. Any recall reported (ANYRECA1=1 or ANYRECA2=1) by CAPI Recall Entries

	<i>CAPI Recall Entries</i>		<i>No CAPI Recall Entries</i>	
	n	Percent	n	Percent
Any recall reported (ANYRECA1=1 or ANYRECA2=1) (n=179)	52	29.0	127	71.0
No recall reported (n=659)	5	0.8	654	99.2

As a result, we cannot know the true number of recall diaries in the LSF and how it compares to production. The declared presence of recall by the FRs and lack of CAPI entries further brings

into question whether recall was not entered for some respondents. For respondents that completed paper diaries, an FR could potentially complete recall in the paper diary as is done in the current production protocol and this would not be distinguishable from any other diary entries. For online participants, we suspect that these recall entries may not have been entered consistently and would have been omitted from the data.

We further examined the FR indications of recall matched with the presence of CAPI entries in the LSF by mode. Figure 8.6.3 shows 21.4% percent of LSF cases had FRs that reported that there were *any* recalled expenses over the two week period. These cases were fairly evenly distributed between online and paper diary respondents. Meanwhile, 6.8 percent of cases had at least one entry made in the CAPI, with this disproportionately occurring for online diaries. Figure 8.6.4 shows the breakdown of cases with no expenditure entries and only recalled expenses. FRs reported no entries made in the diary over two weeks and recalled expenses for one or both of the weeks in 6.5 percent of online cases and in 9.2 percent of paper cases. Based on entries made in the CAPI recall tab, 7.1 percent of online cases only had CAPI entries, while only 1.3% of paper cases only had CAPI entries.

Figure 8.6.3 Extent of recall diary CUs by mode

	<i>N</i>	<i>ANYRECA_n</i>		<i>Any CAPI entries</i>	
		<i>Any Recall diary</i>	<i>Percent</i>	<i>CAPI recall</i>	<i>Percent</i>
Online	368	74	20.1%	41	11.1%
Paper	470	105	22.3%	16	3.4%
Total	838	179	21.4%	57	6.8%

Figure 8.6.4 Extent of total recall diary CUs by mode

	<i>N</i>	<i>ANYRECA_n</i>		<i>Any CAPI entries</i>	
		<i>Only Recall diary</i>	<i>Percent</i>	<i>CAPI recall</i>	<i>Percent</i>
Online	368	24	6.5%	26	7.1%
Paper	470	43	9.2%	6	1.3%
Total	838	67	8.0%	32	3.8%

Proportion of Recall Amongst Respondents

Next, we looked at the proportion of CAPI recall entries out of total expenditure entries based on the diary mode for the LSF. In Figure 8.6.5, we can see that the proportion of CAPI entries provided to FRs was, on average, 0.71 (or 71 percent) of the total number of entries entered in the LSF diary. In other words, among the 57 cases with any CAPI recall entries, an average of 71 percent of their total entries were provided via CAPI recall. For paper respondents that provided any CAPI recall entries, 55 percent of their entries were comprised of recalled CAPI expenses, compared to 77 percent for online respondents.

Figure 8.6.5. Proportion of total entries from CAPI recall by mode assigned.

	<i>N</i>	<i>Mean</i>	<i>Median</i>	<i>Std. Dev</i>	<i>Min</i>	<i>Max</i>
Online	41	0.77	1.0	0.35	0.02	1.0
Paper	16	0.55	0.56	0.41	0.02	1.0
Total	57	0.71	1.0	0.38	0.02	1.0

*(Proportion of total entries entered from recall *on average*)

In Figure 8.6.6, we compare the average entry counts between modes in the LSF with and without recall. Overall, respondents that provided entries and recall were similar to those that had no recall ranging between 45 and 57. However, those respondents that only provided total recall via CAPI had only 24 entries on average. In line with the debriefing responses seen in Figure 8.6.7, many CAPI recall respondents used the recall process as a result of overall disengagement or other obligations that prevented them from recording a large number of expenses.

Figure 8.6.6. Total entry counts by entry mode

	<i>N</i>	<i>Mean</i>	<i>Median</i>	<i>Std</i>
Online diary w/out CAPI recall	285	48.16	34.00	43.75
Paper diary w/out CAPI recall	439	56.76	50.00	39.45
Online diary w/ CAPI recall	15	45.07	35.00	31.47

Paper diary w/ CAPI recall	10	57.60	41.50	59.89
CAPI recall only diary	32	24.00	15.00	25.41

As Figure 8.6.7 confirms, technological issues using the online diary led to some cases of total recall in the LSF. Further examination of the most commonly listed reasons for total recall among respondents is presented below.

Figure 8.6.7. Most frequently listed reasons for total recall in the LSF (not mutually exclusive)

<i>Online diary pickup FR debriefing questions- problems and notes*</i>	<i>Selected examples:</i>
Preferred recall/unreceptive	<ul style="list-style-type: none"> • CU was very willing to save receipts and let me do the work of entering • she went thru receipts on her phone and read them off to me, didn't hand me any • they did not fill out the diary and would only provide a few expenses • vacation/left receipts @ home, entered thru recall • RSPDNT WAS VERY NICE BUT DID NOT PUT FORTH MUCH EFFORT
Tech issues/helpline	<ul style="list-style-type: none"> • the respondent was willing to get help from the helpline but had trouble getting through them, i recorded expenditures from the exit interview • respondent could not log in and he was put on hold when he called the help line and gave up • they didn't make any entries, she said she kept putting it off because she didn't feel she could do it right • LOGIN ISSUES - RSPDNT WAITED UNTIL THE LAST DAY OF DIARY PERIOD TO LOGIN, SO THE HELP DESK DID NOT HAVE SUFFICIENT TIME TO HELP • too much data for them enter, during placement apple created a strong password which then locked them out of online diary
Time/busy concerns	<ul style="list-style-type: none"> • never had the time to log in and figure it out • she is busy, wanted me to do for her

<i>Online diary pickup FR debriefing questions- problems and notes*</i>	<i>Selected examples:</i>
	<ul style="list-style-type: none"> • they were just so overwhelmed with their business, they did not have any time to participate other than recall • she thought it was too much work signing onto site and entering all expenses

*There were 32 total responses entered by FRs for DIPKNOTE and DIPKPROB for the Total recall diaries.

The main reasons respondents' expenditures were collected by recall were respondent preference, technology issues that prohibited their use of the online diary, and other respondent obligations. There were also a few cases of respondent illness and privacy concerns that prevented traditional recording of expenditures in the diary. Ultimately, some respondents who had disengaged from the survey ended up providing receipts to their FR or recalling their expenditures.

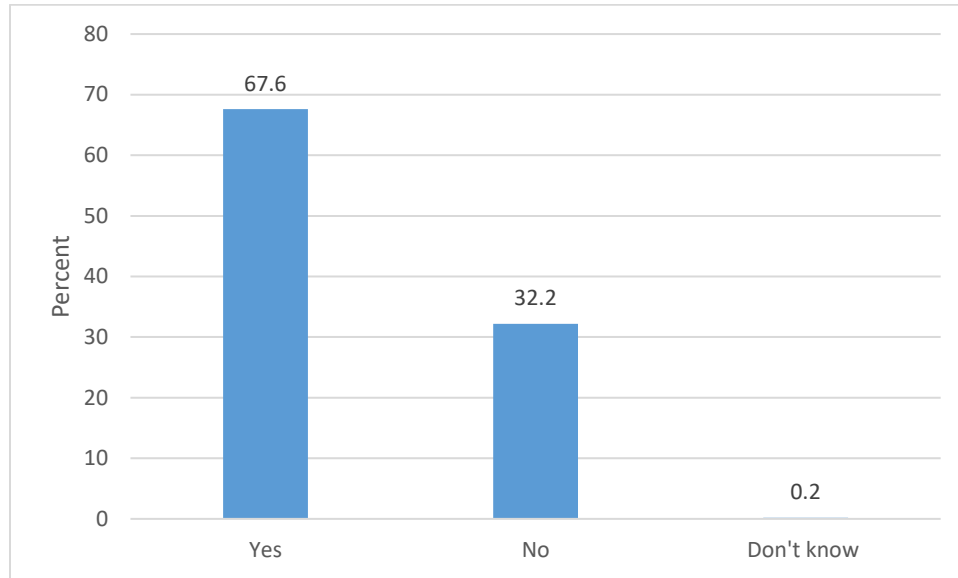
9 FR DEBRIEFING:

9.1 FR DEBRIEFING QUESTIONS

FRs answered debriefing questions after the placement visit and the pickup visit which provided information on various procedures like logins, placement, and pickup procedures and included an assessment of the diary quality from the FR perspective. FRs only answered the debriefing questions for successful placements and pickups, so nonrespondents were not included. The FR pickup debriefing questions in the CAPI instrument were for both online and paper diary cases. The FR placement debriefing questions were for online mode only.

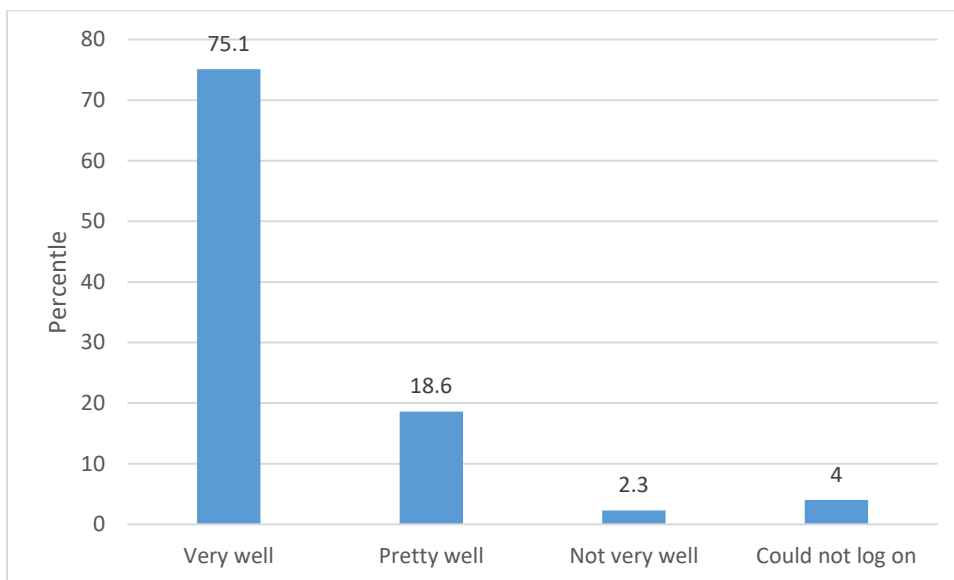
FRs were instructed to login with the respondent during the placement interview whenever possible. FRs reported attempting to login with the respondent during placement in 67.6% of the interviews while in 32.2% of interviews they were unable to do so (see Figure 9.1.1), a high rate that is substantiated by the paradata findings in Figure 7.1.1.

Figure 9.1.1. Percentage of FRs who attempted to login during the placement interview (n=445)



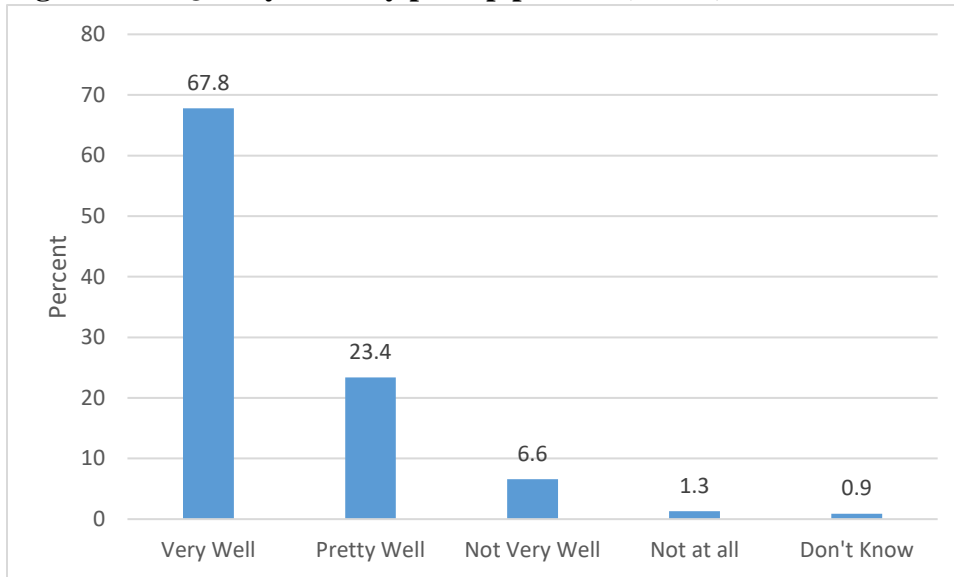
For the large majority of cases, the login process seems to have gone well as shown in Figure 9.1.2. Of those that logged in, 75.1% said the login process went very well and 18.6% said it went quite well, however there were 12 respondents (4%) who could not logon successfully with the FR. In cases where the login process did not go well, the major reasons provided were user error, problems accessing the website, internet access issues, or getting locked out of the diary.

Figure 9.1.2 Quality of the login process (n=301)



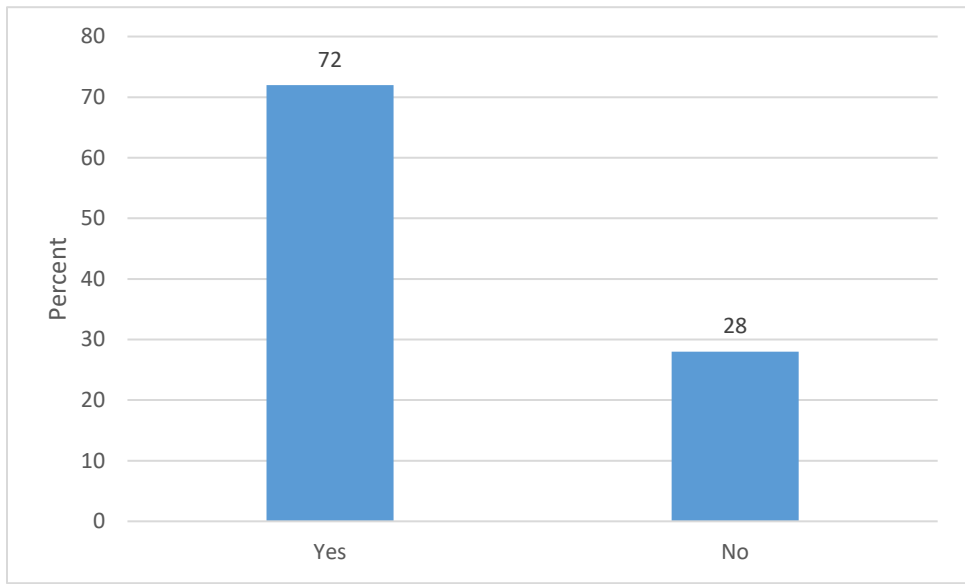
FR reports on the diary pickup process were very positive as indicated in Figure 9.1.3. The diary pickup process went very smoothly or pretty smoothly in 71.2% of cases according to the FR. For cases where the pickup did not go smoothly, the reasons included language barriers, respondent impairment, technical problems, concerns about privacy or scams, business during the diary period, respondent burden, concerns about itemizing groceries, and refusals.

Figure 9.1.3 Quality of Diary pickup process (n=770)



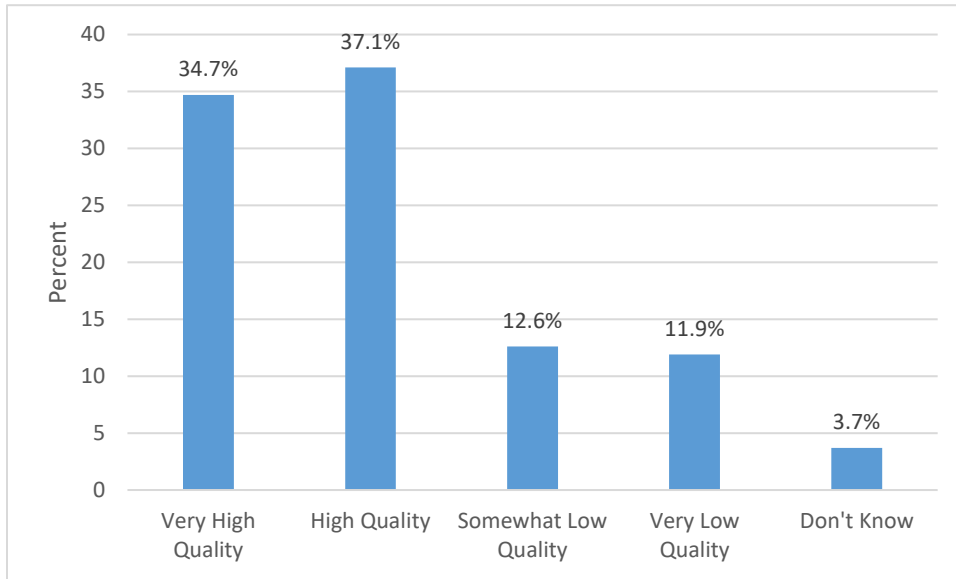
The expenditure summary report was a tool used by FRs to track whether the respondent was making expenditure entries in the online diary and what expenditures they had reported. FRs were instructed to look at this before phone or in-person contacts with the respondent. FRs reported using the expenditure summary report in 72% of cases, indicating that the reports were widely used (Figure 9.1.4).

Figure 9.1.4. FR use of expenditure summary report (n=347)



Additionally, FRs were asked to assess the data quality of the online diaries. The data quality assessment is shown in Figure 9.1.5. The FRs reported high or very high data quality in 71.8% of cases where they placed online diaries. However, about 28% reported that the diaries were not high quality. The most frequent explanations provided were respondent disinclination to fill the diary including being busy, disinterested or concerns with privacy. The second most frequent explanation was respondent limitations such as mental, physical or language barriers. There were also some cases of the respondent not following instructions (such as itemizing groceries) and technical limitations (such as technical abilities, computer, or login issues).

Figure 9.1.5. FR perception of data quality of the online diaries (n=763)



9.2 FR DEBRIEFING SUMMARY OF RESULTS FROM CALLS:

Two debriefing calls were held with field FRs, one in December 2019 and the second in May 2020. During the December debriefing one of the main topics discussed was the lower than expected placement rate for online diaries. FRs mentioned that the main reasons for placing paper diaries for online eligible respondents were language issues, technical challenges, poor connectivity and respondent distrust or security concerns with using an online diary. Age and tech savviness were mentioned as the biggest factors for online diary placement. The general impression of the online diary was positive. FRs also discussed midweek calls and some mentioned that they did not always enter these into the Contact history instrument. Some FRs seemed confused about the procedures for entering recalled expenses, and a field memo was sent out clarifying the recall procedures and also reiterating the importance of midweek calls and entering them into CHI.

In the May debriefing, one of the main issues reported by FRs was lack of response from the help desk. They also mentioned several issues with training including inadequate training on entering responses in the recall tab, insufficient practice during training, and lack of emphasis on the YouTube videos. Another concern was that FRs felt awkward/dishonest about not

mentioning to respondents that FRs could see their expenditure reports for midweek calls⁶³. On a positive note, FRs received positive feedback from online diary respondents about the diary and felt that the online diaries seemed suitable for implementation into production. Respondents who used online diaries, liked them and didn't report many issues. Consistent with the earlier debriefing, FRs noted that respondents who didn't get placed with online diaries despite being eligible were largely older, had cognitive or language issues, or lacked tech savviness or connectivity. Respondents mostly seemed willing to provide their emails, though some needed to be reassured about privacy. Some FRs felt that the postcard was an effective way of getting the attention of the respondent, though the data suggests otherwise. They also reported that after COVID guidance went out, pickup visits for online diaries were easy to do by phone. Their main suggestions for improving the implementation of online diaries were:

- Providing FR reminders as the FRs have many tasks to perform.
- It would be useful for the FR to be able to enter expenses into the online diary, especially during midweek calls.
- Sending the username or password to respondents in advance
- Additional training in the online diary and pickup procedures
- Giving the FR access to the online diary form from the CAPI instrument rather than the current method of accessing expenditure reports.

10 RECOMMENDATIONS FOR IMPLEMENTATION AND LESSONS LEARNED

We have several recommendations for implementation into production based on results from the preliminary analysis of LSF data from October 2019 to February 2020. These should be considered preliminary as they are based on unprocessed data and did not include March LSF data in the analysis.

1. Since prepaid incentives and pre-notification postcards had no significant impact on participation, we recommend not continuing with them for implementation.

⁶³ During training FRs were told not to mention to respondents that they could see their expenses in their expenditure report as it was felt that respondents may feel that was a violation of their privacy.

2. Additional training is needed for FRs in how to use the online diary
3. Better access to the help desk is required to reduce issues reported by FRs and respondents during debriefing. Online diary paradata also indicates a large proportion of failed logins to the online diary, which could be reduced with better training for help desk and FRs.
4. While we saw lower than expected rates on online diary placement, the reasons for that are related to non-tech savvy respondents, poor connectivity, and language issues which are not under our control. As technical abilities and connectivity improves across the nation we will likely see higher rates on placement of the online diary over time. However, we observed lower rates of internet access in the LSF compared to the US population, and there is a possibility that the question is being misunderstood and we need to clarify that internet access includes access via data plans. If we decide to continue to screen respondents, some research and cognitive testing is needed on the internet access question and other ways to screen ideal online respondents.
5. If we plan to screen for eligibility, some modifications may be needed to the protocol to make it very clear how to place a paper diary when a respondent is otherwise eligible for online. This will allow us to get correct data for the screener questions and monitor eligibility and placement over time.
6. The observed rates of CAPI recall in the data, results from FR debriefing, and input from Census indicate that entering recalled expenses at pickup was not done consistently in the recall tab, especially for online diaries. This may have resulted in underreporting in the online diary especially for grocery expenses, where respondents sometimes give the FR a receipt to enter expenses at pickup. For online diary cases, FRs had to close out the pickup interview instrument and could not enter expenses from receipts afterwards as they could do in the paper diaries and in production. Our recommendation is to work with Census to identify a way to better capture recalled entries, particularly from receipts saved by the respondent for online diary respondents. Additional training will be needed to reinforce these changes.

7. Even though we did not see significant differences in contact attempts or logins from midweek calls, contact with respondents is still viewed to be an important way to remind respondents to login and answer their questions during the diary period. However, only 33 percent of LSF completes received successful midweek calls, so additional emphasis on making these calls is needed during training. Additional research is planned on trying to improve the survey's ability to reach respondents for a midweek reminder via texting.
8. While item nonresponse for cost, item description, and outlet name are quite low and comparable to production for cost and item description, we do see a lot of generic entries for outlets which should be addressed since these are not useful for CPI's outlet data collection. In a test version of the online diary currently being developed at Census, we hope to test some new online diary features including auto suggest and additional instructions for the item description field which may help reduce the entry of generic responses in outlet and item description fields in the future.
9. For about 28 percent of cases in the FR debriefing the quality of diary data was reported as being low, and there was no particular difference by diary mode. One reason was due to the respondent's privacy concerns, and this issue needs to be addressed in advance letters and FR training.
10. The YouTube videos were not used much by respondents. These and the user guides need to be highlighted as a resource for respondents in understanding how to use the online diary. Our recommendation is to link YouTube videos to the online diary instrument as part of the help menu⁶⁴. In a 2020 FR focus group, FRs suggested providing a printed user guide to online diary respondents with all the information respondents would require and an additional flap for collecting receipts. This is another suggestion that is worth following up as it also provides a way to collect receipts for online diary users.
11. The expenditure summary report was developed as a way for FRs to monitor data entry in the online diaries so they could provide more targeted help during

⁶⁴ This has been implemented in the online diary after the LSF.

midweek calls and at the pickup interview. This report was only used for about 72% of cases. Our recommendation is to expand the use of this report and for Census to have a clear approach to dealing with issues raised during FR debriefing calls about how FRs should handle this issue keeping in mind respondent privacy concerns.

12. Based on responses to questions on proxy reporting, there seems to be underreporting of expenditures in diaries, the extent of which would depend on how much other CU members are spending. Based on this finding, CE should investigate how to better elicit expenditures for other CU members and may consider specific prompts or reminders multi-member CUs.

Note that further analysis of expenditure counts and amounts as well as other data quality issues will be the focus of the final report. We will also revise response rates and some other analyses, once the data is processed.

APPENDIX A: METHODOLOGY AND ASSUMPTIONS BEHIND FIGURES V.1.1.1 AND V.1.1.2 AND TYPE A INCOMPLETES

Assumptions for Figures V.1.1.1 and V.1.1.2

We look at the percentage frequency distribution of those paper respondents that switched their answers regarding their access to the internet (Figure II.2). Over 85 percent of the 28 respondents switched their answers from having internet access to not having internet access. Including those who initially reported having internet access but later changed their responses, would have increased internet access to approximately 71 percent of eligible members.

Figure II.2 Changes in online screening questions (Internet access) among paper respondents

	<i>Yes to No</i>	<i>Yes to blank (no answer)</i>	<i>No to Yes</i>	<i>No to Don't know</i>	<i>Total</i>
Internet access (N=28)	85.7%	3.6%	7.1%	3.6%	100%

Similarly, a majority of changes to the Internet ability responses reflected a higher number of non-eligible paper diaries. Approximately two-thirds of the respondents changed their answers to the Internet ability question that ultimately made them ineligible for the online diary. Therefore, the final amount of online eligible respondents would have been higher if not for the changes in respondent answers as seen in the audit trail data.

Figure II.3. Changes in online screening questions (Internet ability) among paper respondents

<i>Change</i>	<i>Internet ability (N=42)</i>
Sufficient Internet ability -> Insufficient Internet ability	
Daily -> A few times per month (ineligible for online diary)	7.1%
Daily -> Less than a few times per month (ineligible for online diary)	14.3%
Daily -> Don't know	19.1%
A few times per week -> A few times per month (ineligible for online diary)	19.1%

<i>Change</i>	<i>Internet ability (N=42)</i>
A few times per week -> Less than a few times per month (ineligible for online diary)	7.1%
Subtotal	66.7%
Insufficient internet ability -> Sufficient internet ability or no change	
A few times per week -> Don't know	7.1%
Daily -> A few times per week or vice versa	11.9%
Online diary ineligible Internet ability -> Online diary eligible internet ability	14.3%
Subtotal	33.3%
Total	100%

Assumptions for Type A Incompletes

- CED Production:
 - We determined if a case was deemed a Type A incomplete if the diary was successfully placed with a sample unit member but was a Type A at pick up (Using the variables plcecode and pickcode).
- LSF:
 - An LSF case was classified as a Type A incomplete if the diary was successfully placed with a sample unit member (online or paper diary) (onlinepl) but not picked up (No mode of completion) (mode=0).
 - i. This is viewed as a more accurate measure of Type A incompletes than what was available using the variables instat1 and pick_up1 since these variables may include pickups despite being Type A's. Using the same methodology as CED, there were 134 Type A incompletes, only 3 less than those reported in Figure III.2.

APPENDIX B: METHODOLOGY AND ASSUMPTIONS BEHIND FIGURES V.4.1.1 AND V.4.1.2.

Assumptions for LSF and CED Production contact attempts

- Only cases with at least 1 contact attempt (cntctyp= '1', '2', '3') are used in the analysis.
 - We removed cntctyp=4 from the data. These are instances of FRs “Not attempting contacts” who may have incidentally opened the Contact history instrument.
- A contact attempt for a case is identified as a unique visit date-time CHAI record.
- Contact attempts are counted as those attempts made with the CU.

Contact attempts to placement outcome

- We define the total number of contact attempts to placement outcome as the following:
 - If contact day <= Placement date then CU received contact prior to placement.
 - If contact day > Placement date then CU received contact after placement.

Residual no. contact attempts (after attempts to placement outcome and midweek attempts) to final disposition

- Residual contact attempts were computed as the following:
 - LSF: Total no. of contact attempts minus the number of attempts to placement outcome and no. of midweek contacts.
 - CED Production: Total no. of contact attempts minus the number of attempts to placement outcome.

Aggregate counts for Figures V.4.1.1 and V.4.1.2.

	<i>Complete</i>	<i>Type A</i>	<i>All</i>
No. LSF cases	838	934	1,772
No. LSF cases receiving \$5 cash incentive	412	467	879
No. LSF cases not receiving \$5 cash incentive	423	466	889
No. Production cases (CED)	2,614	2,444	5,058
Total no. of contact attempts			
<i>LSF</i>	4,914	4,398	9,312
<i>LSF cases receiving \$5 cash incentive</i>	2,404	2,174	4,578
<i>LSF cases not receiving \$5 cash incentive</i>	2,498	2,224	4,722
<i>Production</i>	12,922	12,292	25,214

	<i>Complete</i>	<i>Type A</i>	<i>All</i>
Total no. attempts to placement outcome of the diary			
<i>LSF</i>	2,463	3,893	6,356
<i>LSF cases receiving \$5 cash incentive</i>	1,222	1,957	3,179
<i>LSF cases not receiving \$5 cash incentive</i>	1,236	1,936	3,172
<i>Production</i>	5,294	9,409	14,703
Total no. midweek contact attempts			
<i>LSF</i>	397	263	660
<i>LSF cases receiving \$5 cash incentive</i>	199	137	336
<i>LSF cases not receiving \$5 cash incentive</i>	197	126	323
<i>Production (not applicable)</i>	563	297	860
Residual no. contact attempts (after Visit 1 and midweek attempts) to final disposition			
<i>LSF</i>	2,054	242	2,296
<i>LSF cases receiving \$5 cash incentive</i>	983	80	1,063
<i>LSF cases not receiving \$5 cash incentive</i>	1,065	162	1,227
<i>Production (=Total no. attempts-#attempts to placement)</i>	7,065	2,586	9,651
Total no. of in-person attempts*			
<i>LSF</i>	3,018	3,324	6,342
<i>LSF cases receiving \$5 cash incentive</i>	1,434	1,670	3,104
<i>LSF cases not receiving \$5 cash incentive</i>	1,577	1,654	3,231
<i>Production</i>	9,351	8,897	18,248
Total no. contacts made with sample unit member*			
<i>LSF</i>	2,780	1,208	3,988
<i>LSF cases receiving \$5 cash incentive</i>	1,375	594	1,969
<i>LSF cases not receiving \$5 cash incentive</i>	1,396	614	2,010
<i>Production</i>	5,313	4,941	10,254

*Total no. of LSF in-person attempts is missing 7 cases which were not classified as either incentive or control group. Total no. of LSF contacts made with sample unit member is missing 9 cases which were not classified as either incentive or control group.

APPENDIX C: GROUPING OF DOORSTEP CONCERN ITEMS (RSPDNT) TO FORM DOORSTEP CONCERN THEME

(Using definition determined by Kopp, McBride, & Tan)

<i>Intends to quit survey</i>	<i>Time</i>	<i>Not interested/Hostility</i>	<i>Survey Voluntary/Privacy</i>	<i>Gatekeeping</i>	<i>Other concerns</i>	<i>No concerns</i>
17-info previously given 18 - too many questions previously 19- too many interviews 21-intends to quit survey	2-too busy 3-intv too time consuming 5-scheduling difficulty	1-not interested 11-hangs up/slams door 12-hostile /threatening	6-survey voluntary 7-privacy concerns 8-anti-govt 9-does not understand survey	15-family issues	23-other	22-no concerns

APPENDIX D: METHODOLOGY AND ASSUMPTIONS BEHIND MIDWEEK CONTACTS AND LOGINS

Midweek calls assumptions:

- Since there was no direct midweek call variable recorded by FRs, we approximated midweek calls in order to see how they impacted average entries, logins and successful completion of the survey. The LSF protocol indicated that midweek calls were supposed to be made by FRs on day 3 and day 8. However, we marked contact attempts as “Midweek” if they occurred between days 3 and 8 after the initial placement date. This was to ensure that most midweek contacts would be accounted for due to the likelihood FRs make a midweek contact in the subsequent days after due to scheduling conflicts.
- Midweek calls were limited to outgoing phone calls (in line with the LSF FR protocols).
- Midweek contacts in tables 9 and 9.1 include cntctyp= (‘1’, ‘2’, ‘3’).
- Successful midweek calls in Figures V.4.1.1 and V.4.1.2 are identified if it is a contact made with the sample-unit member (ctstatus=1).
- Production midweek contacts were calculated the same way as LSF midweek contacts.
 - According to the CED FR Manual, FRs are instructed to follow-up with diarists via phone call “two or three days after placing diaries”.

Midweek logins assumptions:

- For Mean Logins Subsequent, we make assumption that “Login after” includes logging in on either the same day as midweek call or subsequently afterwards.

“Login before” for Mean Logins Prior only includes the number of logins before the day of receiving the first midweek call. Appendix E: LSF AUTOMATED EMAILS

A. Welcome email sent by National Processing Center within 24 hours of diary placement

- Subject: Welcome to the Consumer Expenditure Survey – How to start using your diary
-
- Dear CE Diary Household,
-
- Welcome to the Consumer Expenditure Diary Survey! In this email, you will find all of the information you need to help get you started.
-
- **Step 1** Click or tap on this link to go to your diary
 <https://respond.census.gov/lsf>
- **Step 2** Enter the username and password from the user guide provided to you by your Census Field Representative
- **Step 3** Select the date you are/were scheduled to begin your diary: {Fill Diary Start Date}
-
- If you have any questions about filling out the diary, please refer to your user guide or feel free to contact your Field Representative or your Census Bureau Regional Office at <RO Name and Toll Free #>.
-
- We appreciate your time and effort, and look forward to seeing you again in about two weeks!
-
-
- Sincerely,
U.S. Census Bureau
-
-
- **For additional help on filling out your diary**, please go to the following online resources:
<Diary User Guide url>
<Video Tutorials url>
-
- **For help logging in to your diary:**
- Contact our Help Desk at 877-744-1522 or CEhelpdeskemail@census.gov
-
- Help Desk hours:
- Monday – Friday, 9am – 10pm Eastern

- Saturday, 11am – 9:15pm Eastern
- Sunday, 11am – 9:15pm Eastern
-
- *Please do not reply to this message; it was sent from an unmonitored email*

B.Reminder email sent 8 days after placement, sent to participating households who have not logged in or not entered any expenses

Subject: How to start using your Consumer Expenditure Survey diary

Dear CE Diary Household,

We noticed you might need help with your Consumer Expenditure diary. How can we help you get started?

To log into your diary, tap or click here <https://respond.census.gov/lst>

For help logging into your diary, please contact our Help Desk toll free at (877) 744-1522 or CEhelpdeskemail@census.gov

For guidance on filling out your diary, go to the following these online resources:

<Diary User Guide url>

<Video Tutorials url>

For any other questions about the diary, please contact your field representative, <FR Name> at <FR phone number>.

Thank you for participating – your cooperation makes a difference!

Help Desk hours:

Monday – Friday, 9am – 10pm Eastern

Saturday, 11am – 9:15pm Eastern

Sunday, 11am – 9:15pm Eastern