
**PROOF OF CONCEPT TEST FOR THE
CONSUMER EXPENDITURE SURVEY**

**RESULTS ON RESPONDENT COOPERATION,
SELECT INTERVIEW AND DIARY
CHARACTERISTICS, AND RESPONDENT
EXPERIENCE**

VERSION DATE: 8/31/2016

REVISED DATE: 1/17/2018 (By Laura Erhard and Yezzi Angi Lee)

**Laura Erhard
Brett McBride
Song Park
Lucilla Tan**

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

Table of Contents

I. Executive Summary.....	1
II. Report Scope.....	5
III. Overview	5
IV. Sample Description.....	7
a. Sample design	7
V. Overall Response Rates	9
VI. Respondent Cooperation and Incentives	13
a. Number/effectiveness of contact attempts.....	13
b. Individual diary placement and pickup process.....	15
d. Extent of diarist participation	19
e. Effectiveness of incentives.....	20
VII. Records and Recall Interviews	22
a. Length of the interview elements.....	22
b. Use of records.....	25
VIII. Individual Diaries	28
a. Diary completion.....	28
b. Respondent experience	30
i. Diary-level preference.....	30
ii. Online diaries	31
IX. Individual Diaries Data Quality.....	37
X. Overall Respondent Experience.....	41
a. Analysis of respondent debriefing questions.....	41
XI. Conclusion.....	46
APPENDIX A: CHI	47

I. Executive Summary

The Proof of Concept (POC) test was fielded to assess methodological, operational, and experiential issues related to the Consumer Expenditure Survey's (CE) redesign plan. This report describes aspects of each issue (e.g., methods for the three component design, operational aspects of online diaries and incentive delivery, and respondent experiences with the design). Following on a prior report that examined training and Field Representative (FR) debriefing findings, this report presents findings for the following research objectives:

1. Feasibility of completing one wave of the proposed survey redesign

- The POC test just met the a priori threshold for the number of completed cases needed for the test to be considered a success, with 520 completes for a response rate of 50 percent ([Table 5a](#)).
- Although the response rate was lower than desired given the provision of incentives, changes in diary placement protocols and efforts to address common household reasons for refusal could lead to overall improvements in response for future implementations of the design.

2. Effectiveness of incentives

- The majority (56.7 percent) of debriefing respondents reported incentives did not impact their decision to respond ([Table 6l](#)), although a large majority (84.7 percent) of Consumer Units (CUs) that responded to the debriefing reported receiving the \$2-bill incentive ([Table 6i](#)), suggesting that, at very least, the incentive registered in their memory.
- Beyond respondents' remembering receipt of incentives, data suggest that incentives may have affected their survey behavior, as 89.6 percent of complete CUs received the full \$40 incentive amount for completing the records interview with use of a record ([Table 6j](#)).
- There is the potential that the \$20 debit card provided for completing Visit 1 affected diarist expenditures; between Visit 1 and Visit 2, 43 percent of debriefing respondents used (or attempted to use) the debit card ([Table 10e](#)).
- A small share of those who did not fulfill the requirements to be a complete case fulfilled components that qualified them to receive an incentive – 27.5 percent receiving the recall incentive and 14.7 percent receiving the records incentive ([Table 6j](#)).

3. Respondent willingness to complete all components of the integrated survey

- Almost all CU members present at diary placement were willing to receive a diary ([Table 6g](#)), and the provision of a paper diary option allowed respondents who were unwilling or unable to complete an online diary to participate ([Table 6d](#)); furthermore, full within-CU compliance with completing the diary was high – 71 percent in CUs where multiple diaries were placed ([Table 6h](#)).

4. Number/effectiveness of contact attempts

- Having repercussions for the survey cost, FRs made *fewer* contact attempts for POC than “Restricted Production” (RP) cases (6.1 vs. 7.9, respectively). In addition, in-person visit attempts (instead of phone attempts) were also lower for POC than for RP (3.8 vs 4.1, respectively, [Table 6a](#)).
- In terms of effectiveness, FRs made fewer contacts to complete interviews for POC compared to RP (3.4 vs. 4.7, respectively), suggesting the POC may have been an ‘easier sell’ for respondents ([Table 6a](#)).

5. Length of the interview elements

- POC survey component lengths (averaging 117 minutes for both visits and debriefing, [Table 7a](#)) remained below the a priori thresholds, but were still longer than production counterparts, with the differences being statistically significant ([Table 7b](#)). Longer interview times may be an indication of additional expenditure reports, which would be a positive outcome in the POC.
 - The mean total time of the recall interview sections was 15.8 minutes in POC and 12.8 minutes in the comparable RP sections ([Table 7b](#)).
 - The mean total time of the records interview sections was 39.9 minutes in POC and 21.0 minutes in RP (29.2 minutes in RP cases with extensive record use, [Table 7c](#)).

6. Technical issues respondents have with completing the online diaries

- Online diarists reported technical issues involving the login process, however there was no evidence these were widespread nor that they had a large impact on data quality.

- Although three-quarters of online diarists responding to the debriefing question reported not having any problems using the online diary, of the small subset answering the question about the login process, two-thirds said they had trouble logging in to the diary ([Table 8i](#)).
- Paradata confirmed that a little over half of all online diarists had one or more login failures ([Table 8k](#)), but only 16 percent of online diarists were unable, or unwilling, to ever successfully log-in to their diaries ([Table 8k1](#)).
- Despite these limited technical issues, online diarists generally participated at high levels; 65 percent logged in within the first 2 days, something associated with a higher number of total logins ([Table 6e](#)), and only 4.7 percent supplied all diary entries via the recall and receipts process, less than the corresponding 21.9 percent in the earlier Individual Diaries Feasibility Test (IDFT) ([Table 8c](#)).
- Among multi-member CUs in which all eligible members were assigned online diaries, participation among CU members was more widespread for POC than for the IDFT – 75 percent completion and 43 percent, respectively ([Table 8e](#)).

7. Individual diary placement and pickup process and data quality

- Individual diary placement visits were relatively well-attended, although placement rules and diarist preferences reduced the number of diarists placed with online diaries.
 - While placement was well attended (i.e., in multi-member CUs, slightly over half of eligible members were present when the FR trained members, [Table 6b](#)), presence at diary placement did not appear to be a decisive factor in determining completion, as a larger share of multi-member CUs had full completion than had full member presence at placement ([Tables 8e, 6b](#)).
 - Placement of online diaries was limited: 25 percent of diarists did not meet eligibility criteria for the diaries (i.e., home or mobile internet access that they used multiple times per week), and a larger share – 39 percent – opted to use a paper diary despite being eligible ([Table 6d](#)).
- Few FRs recorded making midweek reminder contacts with CUs – 16.5 percent ([Table 6f](#)); nevertheless, upon diary pickup, over 90 percent of the 991 diarists placed with diaries had completed them, with the proportion making at least one entry directly into their diary prior to pick-up approaching 90 percent ([Table 6g](#)).

- Overall, POC diary data was of higher quality than that from RP; the mean number of entries exceeded one-week production equivalents (38 and 31 respectively, [Table 9b](#)), and missing data rates were lower ([Tables 9c, 9c1](#)).

8. Respondent experience

- Overall, respondents' experiences were positive, as those responding to debriefing questions did not report being overly burdened by the time they spent on the survey, and debriefing respondents who completed the survey reported the diary instruction materials were useful.
- Diarists from complete CUs reported an average time of 35.6 minutes recording expenditures in the diary, and respondents from complete CUs that provided records reported an average of 41.8 minutes collecting records ([Table 10a](#)).
- Almost all respondents reported that the amount of time spent on survey activities was reasonable or somewhat reasonable – 95.7 percent for recording expenditures in the diary and 97.8 percent for collecting records ([Table 10a1](#)).
- Respondents reported spending more time recording expenditures (in response to debriefing questions) than was indicated in the diary paradata; among diarists that self-reported a time in the debriefing and had device time from paradata, the mean paired difference in the two times (self-reported minus paradata) was 11.8 minutes ([Table 10b1](#)).
- Record collection was said to be generally easy; only 16.7 percent of debriefing respondents reported that there were records that were difficult to collect ([Table 10d](#)).

In conclusion, the analysis in this report found that the various aspects of the POC – the interview visits, diary placement and pickup, and the use of incentives – were feasible, being met with strong levels of respondents participation. Respondents agreed to receive diaries and completed them at high rates; they also participated in the follow-up records interview, using records at high rates. FRs needed to make fewer contact attempts to secure interviews. Although the interview durations were longer than equivalent sections in the RP sample, most respondents did not report feeling burdened by the time needed to complete the survey. These findings suggest the POC components would be feasible to implement as part of a redesigned survey. Furthermore, a preliminary analysis of data quality indicators – diary missing data rates and reporting frequencies – suggested a higher data quality in POC compared to the RP sample, with further analysis to follow in report 3.

II. Report Scope

The final results and recommendations from the POC test will be detailed in three separate reports:

Report 1: written jointly by Census and BLS and focused on training, field procedures and the Field Representative (FR) Debriefing summary. The report included recommendations for training, protocols, and design as well as preliminary response rate estimates.

Report 2 (this report): written by BLS focuses on analyzing POC data that have been processed by the BLS Initial Edit System. This report includes detailed analysis of BLS defined response rates, contact attempts, diary use, and analysis of FR and respondent debriefing questions.

Report 3: written by BLS and focusing on analyzing POC data that have been processed by the BLS Edit and Estimation System. This report will include detailed analysis of total expenditures and data quality in the POC data.

III. Overview

The Proof of Concept (POC) test was fielded to assess methodological, operational, and experiential issues related to the Consumer Expenditure Survey's (CE) redesign plan. The initial analysis focused on field operations and found no major methodological, operational, or experiential issues in the design based on the field operations. FRs provided positive feedback about the design and were generally able to follow the design protocols. The team made some recommendations on ways to improve training and protocols, but recommendations were minor and fit within the redesign plan.

Prior to making a final recommendation on whether the redesign "concept" is feasible, the team must analyze the results of the final data of the test. The major research objectives of this analysis are:

1. Feasibility of completing one wave of the proposed survey redesign
2. Effectiveness of incentives
3. Respondent willingness to complete all components of the integrated survey
4. Number/effectiveness of contact attempts
5. Length of the interview elements
6. Technical issues respondents have with completing the online diaries
7. Individual diary placement and pickup process and data quality
8. Respondent experience

The basic design of the POC consisted of 3 components: an in-person recall interview (“recall interview”) as part of Visit 1, a week of diary keeping for CU members 15 and older (“individual diaries”), and an in-person interview with a record focus (“records interview”) as part of Visit 2. These components reflect one single wave of the [CE Redesign plan](#). The POC test also had an incentive structure as described in the redesign plan consisting of a prepaid \$2-bill incentive mailed with the advance letter to sampled addresses, a \$20 debit card handed to the respondent upon completion of the recall interview, \$20 debit cards mailed to each member completing a diary, and a \$20 debit card mailed to the respondent after completion of the records interview with an additional \$20 debit card mailed to respondents who used at least one record. Following completion of all sections of the interview, FRs administered a series of debriefing questions to the respondent within the CU who completed the records interview. FRs additionally answered FR debriefing questions at the end of each visit.

IV. Sample Description

a. Sample design

The POC was fielded out of four regional offices (ROs): Atlanta, Chicago, Denver, and New York during quarter 3 of 2015. The sample was drawn from a test sample based on the 2000 CE sample design, which was available at the time of planning. However, since CE production moved to the 2010 sample design in 2015 prior to the fielding of the POC test, only continuing counties within the test sample were included. This allowed field representatives (FRs) who were already trained on the CE Quarterly and Diary surveys to be used for the test.

In addition to using continuing counties from the 2000 sample design, counties were targeted for the sample based on their prevalence of English-speaking households and internet use. The decision to target English-speaking households was to limit the effects of a language barrier on the test. The decision to target areas with a high rate of home internet access was to sample in areas where more respondents would be eligible for the web diary – a feature of the new design that is intended to increase contemporaneous reporting, as well as allow for easier entry of expenditures. Although these types of areas were *targeted* for the sample, all CUs were eligible to participate in the POC. As a result of the targeted sample and use of only continuing counties, no rural counties were included in the POC sample.

The target number of completed interviews was between 520 and 565 – with a complete interview defined as a completed recall interview, at least one completed diary, and a completed records interview. In order to obtain these completed cases, an estimated starting sample size of 1,200 was drawn. The final number of complete POC test cases met the target with 520 completed cases.

Throughout the analysis, the POC data are compared against production data for interview and diary collected in the same time period as the POC test (2015 quarter 3) and in *only* the same counties sampled for the POC test. This ‘Restricted Production’ (RP) sample was composed of 1,483 interview cases (*across all waves*) and 645 production diaries (used for response rate comparisons) and 315 first week diaries (used for analysis in order to be more comparable to the one-week POC diary). For select analyses of the diary, the sample was further restricted to nonblank diaries, resulting in a sample size of 291 diaries. These instances are noted.

b. Sample characteristics

Looking at the demographics of the survey respondents, the POC sample mostly had a very similar distribution to those of the RP-Interview and RP-Diary samples (Table 4a). Owners (59 percent) were less likely to participate in the POC compared to RP-Interview (63 percent) and RP-Diary (63.2 percent). This could have been the result of incentives motivating a higher proportion of lower-income individuals who are more likely renters to participate in the survey. On the other hand, there was a very close distribution of education for the respondents compared to the RP sample, which is also typically correlated with income. Further analysis of income distributions will be available in the third report and we will attempt to uncover any potential bias resulting from the use of incentives.

Table 4a: Sample characteristics

	POC	RP - Interview	RP – Week 1 Diaries*
No. of Consumer Units	520	1483	315
Race of Respondent*			
White	82.5	79.2	73.0
Black	11.9	15.0	12.1
Other	5.4	5.8	5.7
Hispanic Origin of Respondent*	10.8	12.1	9.8
Gender of Respondent*			
Female	54.8	54.2	47.6
Male	45.0	45.9	43.2
Age of Respondent*			
Under 25 years	6.7	5.0	5.4
25-34 years	16.4	17.1	15.6
35-49 years	29.4	25.8	22.9
50-64 years	26.4	18.8	24.8
65 years and older	21.0	23.5	22.2
Education*			
Less than high school	7.5	8.3	7.3
High school grad	21.9	21.0	16.2
Some college	31.9	29.8	23.9
College grad	38.5	40.9	43.5
CU Size			
Single person	30.4	30.8	34.9
2-3	46.2	47.9	45.1
4+	23.5	21.3	20.0
Housing Tenure			
Renter	41.0	37.0	36.8
Owner	59.0	63.0	63.2

*Race, Hispanic Origin, Gender, Age, and Education are unknown for 29 production diary cases (9.2%) and 1 POC test case (0.2%).

V. Overall Response Rates

The final response rate for the POC test that BLS will use is based on the Consumer Unit's (CU's) completion of all three components of the POC: a complete recall interview, at least one expenditure in a diary (or indication of a week without expenditures)¹, and a complete records interview. This definition differs from the preliminary response rate reports from Census that reported on completeness of the recall and records interview, but not the diary. As a result the final response rates are lower than the Census reports. Details on the response rate calculations are presented in Table 5a. In the analysis plan for the POC, the team set guidelines as a starting point for the evaluation of success. The guidelines of the response rate was "higher than 60 percent *or* less than 10 percent below production". Historically, field tests have not performed as well as production, so this was taken into consideration when defining goals. The team had expected a higher response rate than typical field tests, likely due to the use of incentives. The final response rate did not meet the 60 percent goal (Table 5a.), but was close to the 10 percent threshold below production (13.1 lower than RP-interview and 8.8 percent lower than RP-diary).

Table 5a. Overall Response Rates Compared to Restricted Production (RP)

		POC	RP – Interview	RP – Diary
Starting Sample	(Type A + Type B/C + Complete ²)	1,229	2,805	1,314
Type B/C Cases		185	447	213
Type B/C Rate	$\left(\frac{\text{Type B/C}}{\text{Type A + Type B/C + Complete}}\right)$	15.1%	15.9%	16.2%
Eligible Cases	(Type A + Complete)	1,044	2,358	1,101
Type A Cases		524	875	456
Type A Rate	$\left(\frac{\text{Type A}}{\text{Type A+Complete}}\right)$	50.2%	37.1%	41.4%
Complete Cases/Diaries		520	1,483	645
Response Rate	$\left(\frac{\text{Complete}}{\text{Type A+Complete}}\right)$	49.8%	62.9%	58.6%

¹ In addition, the total number of expenditures reported via the diaries or diary recall needed to meet productions' processing thresholds for the minimal expenditure edit

² 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)

There were two types of reasons that a POC CU was classified as a “Type A” nonresponse: (1) either the CU refused or was unavailable to complete the interview from the beginning (Type A Refusal or Noncontact) or (2) the Interview was started, but not all of the components were completed (Type A Incomplete). For the POC, 36.4 percent of the eligible cases were refusals or noncontacts and 13.8 percent were incomplete (Table 5b). Looking more closely at the incomplete cases, the majority (53.5 percent) completed both the in person interviews, but did not complete a diary. Part of this can be explained by the placement date rules for the POC, modelled after the CED placement dates. The FRs were given 10 days to conduct the first visit and place a POC diary. Those placement dates were spread across the month to ensure that any cyclical diary expenditures were captured. If an FR missed the placement date, then we still requested that they attempt to complete the recall and records interviews with the goal of identifying what is feasible within the month. There were 59 cases that were incomplete due to the late placement of diary (41 percent of incomplete cases, Table 5b1). Of these 59 cases, 51 completed the records and recall interview, which suggests that they were otherwise cooperative respondents that likely would have completed the diary if it had been placed.

Beginning in 2017 production (and moving forward with the redesign), CE is moving away from rolling diary placement and instead allowing diary cases to be completed with similar protocols to the CEQ, with placements occurring at the FR’s discretion between the first and last the month. If restriction for earliest placement dates had been lifted in time for the redesign, the aforementioned 51 cases that had completed both the records and recall interview would likely have completed the diary as well and been considered complete – which effectively would have increased the response rate from 49.8 percent to 54.0 percent³.

Table 5b. Overall POC Dispositions

	N	Percent
Eligible Sample	1,044	100
Type A Refusal or Noncontact	380	36.4
<i>Refusal</i>	297	78.2
<i>Noncontact</i>	50	13.2
<i>Unknown</i>	33	8.0
Type A Incomplete (records, diary, or recall not complete)	144	13.8
<i>Completed recall and records, no diary</i>	77	53.5
<i>Completed recall and diary, no records</i>	2	1.4
<i>Completed recall, no diary no records</i>	65	45.1
Complete Cases	520	49.8

³ We estimate this rate by finding the percentage of households that completed at least one diary when placed with the household (86.3 percent) and applying that percent to the cases that were “placed too late”, but completed the recall and records interviews (51 cases).

Table 5b1. Incompletes Due to Late Diary Placement

	N	Percent
Incomplete	144	100
Diary placed too late	59	41.0
<i>Completed recall and records, no diary</i>	51	86.4
<i>Completed recall, no diary no records</i>	8	13.6

Another comparison to look at is the percent of CUs that refused the recall interview compared to the RP-Interview Wave 1 cases. This comparison helps to isolate the differences in the POC leading up to the contact with the CUs, such as the advance letter containing the prepaid \$2-bill incentive and the description of the interview process and incentive structure. For the RP-Interview, there was an advance letter that described the production interview, but no incentives. The analysis plan set a goal that the POC response rate for the recall interview should be no less than 10 percentage points lower than the RP-Interview wave 1 response rate. In this case, the POC recall interview response rate was nearly identical to the wave 1 response rate (63.6 percent compared to 64.4 percent, Table 5c). However, there were substantial differences in the types of nonresponse for each. In general, with incentives you would expect to see lower noncontact rates and lower refusal rates. However, in the POC, while the noncontact rates were 3.3 percentage points lower, the refusal rates were 5.3 percent higher. It's not possible to know definitively what caused this increase in the POC refusal rate; however, the description of the POC presented to the potential respondent at the time of the contact could have led to a refusal (a scenario that we hoped the incentives would prevent). There could have been other unknown differences between the production cases and the POC test that also led to a higher refusal rate for the POC recall. A potential explanation that has been proposed in the past is that FRs work harder to get production cases compared to test cases. Although, FR training emphasized the POC as a "trial run of the redesign" instead of a "test", it was still not production. There's also a question about the effect that incentives have on the FRs interaction with potential respondents – do FRs rely too much on the conditional incentive for persuading the CU to cooperate? If this is the case, we will learn more from the Incentives test that is being fielded in 2016.

Table 5c. Interview Dispositions, Recall Interview compared to Wave 1

	POC- Recall Interview	RP - Interview, Wave 1
N Eligible (Complete + Type A)	1,044	567
Noncontact rate (%)	4.8	8.1
Refusal rate (%)	28.4	23.1
Other Type A (%)	3.2	4.4
Response rate (%)	63.6	64.4

Similar to the attempt to compare the first contact with the CU, a research goal was to compare the second contact with the CU. For POC, this meant comparing all cases that had a complete recall interview at Visit 1 and looking at the refusal and noncontact rates for the subsequent records interview (Visit 2) against the wave 2 interviews in the RP-interview sample that completed a wave 1 interview (Table 5d). For the POC, we expected those that completed the recall interview and were handed an incentive after the interview would be more likely to cooperate in the records interview and this was the case, with only a 7.8 percent refusal rate for the records interview. A full 89.8 percent of CUs who completed a recall interview agreed to participate in the records interview. For the RP-Interview wave 2 cases, there was a higher refusal rate compared to the POC. This could either be due to the lack of incentive or because there were 3 months between the waves instead of 1 week. Despite the comparisons being similar, the consequence of the disposition at the second visit is very different for the POC (or redesigned CE survey) compared to production. For production, each wave is treated independently and if a respondent completes wave 1, but not wave 2, there is little effect on the data. For the structure of the redesign and the POC, if a respondent completes the recall interview and not the records interview (or the diary), then the full set of expenditures for that CU is not captured. So for POC, it's crucial to maintain cooperation through the records interview.

Table 5d. Interview Dispositions, Records Interview compared to Wave 2

	POC- records interview	RP - Interview, Wave 2, returning CUs
N Eligible (Complete Recall Interview)	655	362
Noncontact rate (%)	0.1	3.5
Refusal rate (%)	7.8	11.3
Other Type A (%)	1.7	2.2
Response rate (%)	89.8	82.9

VI. Respondent Cooperation and Incentives

a. Number/effectiveness of contact attempts

Making contact with the sample unit is the first step to gaining respondent cooperation, and the effort expended to make contact is a significant component of the survey budget. We used FRs' reported contact attempt information to proxy for the cost of effort.

Methodology. The POC design separated out the sections of the CEQ questionnaire into two mutually exclusive interviews with Recall sections and Records sections, respectively. Thus, the CEQ questionnaire content equivalent to POC's Visit 1 (Recall sections) and Visit 2 (Records sections) would be 1 wave of the CEQ, where all sections are asked within the same interview. The equivalent questionnaire content to POC's 1 week diary would be the CED's Week 1 diary. Thus for the purpose of examining contact attempts, the *comparable RP sample* to the POC was comprised of cases from CEQ Wave 1 plus CED Diary Week 1 (whose data collection period was similar to the POC). *Average counts per case* were defined as the following:

- *for the POC sample* = (total counts / number of POC cases)
- *for the RP sample* = (total counts for CEQ Wave 1 cases / number of CEQ Wave 1 cases) + (total counts for CED Week 1 cases / number of CED Week 1 cases)

The average total number of contact attempts for POC is *overall lower* compared to RP (6.1 contact attempts per case vs. 7.9 attempts, respectively); this holds when limiting the cases to completed interviews (5.3 vs 6.9). See Table 6a.

The FR's effort to attempt contact with the sample unit in person (in-person visit attempts) is more costly than a contact attempt by phone. The average total number of in-person visit attempts for POC was also *overall lower* than RP (3.8 contact attempts per case vs. 4.1 attempts, respectively). POC also required fewer contacts with sample unit members to complete interviews compared with RP (3.4 vs 4.7, respectively).

One indication of the effectiveness of the incentives in improving cooperation is whether there was a reduction in the effort needed to secure "Visit 1" with the CU. For this analysis, Visit 1 was defined as the following: for a POC case, it was the date of the completion of the CU roster; for a CEQ Wave 1 case, it was the date of *the first occurrence* of at least a partial or complete interview (CTTYPE =1,2); for a CED case, it was the first diary placement date (PLACDTE1). In other words, we measured how much effort an FR took to "get their foot in the door" and start the interview by looking at the first occurrence

of a started interview (regardless of the final outcome of that interview). The average total number of contact attempts up to and including Visit 1 was *lower* for POC compared to RP (2.1 contact attempts vs. 6.7, respectively). This could potentially be an effect of including a \$2 prepaid incentive and informing respondents of the POC incentives structure in the advance letters – which could also explain the lower noncontact rate seen in table 5c.

The average number of contact attempts to final case disposition after Visit 1 and midweek attempts (‘residual number contact attempts’) was *higher* for POC than for RP (3.9 attempts vs 1.3, respectively), but this is unsurprising since it included the contact attempts for Visit 2 – a feature unique to the POC design⁴.

Table 6a. The average counts per case from contact attempt history data

	Interview	Mean Type A	ALL
No. POC cases	520	520	1,040
No. of RP cases (CEQ Wave 1 + CED Week 1)	673	360	1,033
Total no. of contact attempts			
POC	5.3	7.0	6.1
RP	6.9	9.9	7.9
Total no. attempts to Visit 1			
POC	2.9	1.4	2.1
RP	5.2	9.3	6.7
Total no. midweek (pCHI) contact attempts			
POC	0.1	0.0	0.1
RP (not applicable)			
Residual no. contact attempts (after Visit 1 and midweek attempts) to final disposition			
POC	2.2	5.6	3.9
RP (=Total no. contact attempts - Visit 1 attempts)	1.7	0.6	1.3
Total no. of in-person attempts			
POC	3.8	5.4	4.6
RP	4.1	5.5	4.6
Total no. contacts made with sample unit member			
POC	3.4	2.2	2.8
RP	4.7	3.9	4.5

See Appendix A for note and assumptions behind Table 6a.

⁴ For the RP sample, this residual value represents any contact attempts FRs needed to make following the first interview contact (e.g., attempts to complete a partially-completed interview) up through the final disposition.

b. Individual diary placement and pickup process

The placement of the POC individual diaries was done at the end of Visit 1. The FRs were instructed to have the diarist gather any available CU members to participate in the placement process. Any respondent age 15 and older was eligible and asked to keep a diary.

Upon determination of eligibility, a CU member could either refuse or agree to participate in the diary task. A large share of eligible members agreed to have a diary placed (Table 6b).

Table 6b. Diary Placement Process

Total eligible members	1,008
Total diaries placed	991
Average # of CU members eligible for diary	1.9
Average # of CU members present at placement	1.5
% of multi-member CUs with all eligible present at placement (N=375)	56.6

Among CUs participating in diary keeping, there were an average of 1.5 members of the CU present when placement occurred. Furthermore, slightly more than half of complete CUs had all eligible members present at placement (56.6 percent).

Part of the placement process was to determine the diary mode for each eligible member. The placement questions were asked to (or about) each eligible member to determine whether they would be given the option to keep a paper or an online diary. Each eligible member was asked whether he or she had access to the internet at home either through a computer or mobile device and how often he or she accessed the internet. If the member wasn't present, then the respondent answered the related questions on that member's behalf. Despite concerns that the respondent wouldn't know the answers for other CU members, there were no cases where the respondent didn't know or refused to answer the internet ability and access questions on behalf of other non-present household members. There was only one household where the respondents (both eligible members present at placement) refused to answer the internet ability questions.

Table 6c. Diary Placement Process	Percent
Internet Access (asked of eligible members, N=1,008)	
Yes	80.0
No	19.9
Internet Ability (asked if Internet Access = yes, N=806)	
Daily	68.9
A few times per week	5.1
A few times per month (ineligible for online diary)	2.5
Less than a few times per month (ineligible for online diary)	3.4
Eligible for online diary (of all eligible members , N=1,008)	74.1

If home internet access was available and the member used the internet at least a few times per week, then the member was provided the option to keep either an online or a paper diary. Otherwise, they were only provided the paper diary to keep during the week. Of all eligible members, 74.1 percent were given the option of keeping the online diary (Table 6c). Of those that were given the option, only 46.2 percent chose to keep an online diary, with others choosing to keep a paper diary or refusing. Overall, there were only 345 members that kept an online diary, 34.2 percent of all eligible members (Table 6d). For those choosing an online diary, FRs were instructed to assist the diarists with logging in to their diary for the first time, to troubleshoot any problems diarists may have encountered with credentials or the website. Section VIII provides more information on the few online diarists who switched to using a paper diary.

Table 6d. Diary mode of eligible members (N=1,008)	Percent
Paper (ineligible for online)	25.2
Paper (eligible for online)	38.9
Online	34.2
Refused	1.6

c. Effectiveness of midweek contact attempts

After the initial interview and diary placement, FRs were instructed to follow up with respondents and diary keepers during the week (“midweek contacts”). At a minimum, the FR was instructed to contact the main respondent mid-week to encourage the diary keeping task and to remind him/her of the records collection task. For any member that provided an email address and was placed with a web diary, the FR was asked to send an email with links to the diary and other useful information. The FR was also encouraged to contact other CU diary keepers at their discretion to remind them to record their expenditures or to solve any technical issues arising from the online diary (as applicable). For each of

these contact attempts, the FR was asked to record the contact attempt in the contact history instrument, which was designed for the POC to capture contact history at the CU and person level (pCHI).

Analysis showed that online diarists who logged in earlier in the reporting period (on the 1st or 2nd day) had a higher number of entries than those who did not log in until later than the 2nd day (Table 6e).

Table 6e. Difference in number of online diary entries by period before first login (N=283)*

	N	Mean	Median	Min	Max
Within 1st 2 days	184	17.4	12	0	100
After 2nd day	99	11.9	8	0	66

*excludes diarists without a successful login and total recall diaries (includes eventual non-participating diarists)

There were a large proportion of online diarists (among those with a successful login) logging in shortly after the reporting period began (65 percent of total)⁵. These diarists were seen to have a significantly higher number of logins⁶ than those who did not log in until the 3rd day (averaging 17.4 and 11.9 respectively). These findings suggest the importance of either initial FR logins with the respondent or the making of midweek contacts to ensure that diarists were participating earlier in the week and that any difficulties they may have had with the diary keeping task were being addressed.

The number of midweek contact attempts for POC was computed at the CU-level, i.e. multiple contact attempts (pCHI) occurring on the same visit date between Visit 1 and Visit 2 for a case would be counted as 1 attempt. Data on midweek contact attempts indicated that there were 86 of the 520 CUs, or 16.5 percent, in which the FR made a midweek contact attempt⁷. This signifies a limited FR compliance with either the protocol of attempting midweek follow-up contacts that they were trained to make at a minimum of day 3 of the reporting period or of recording mid-week contacts into the pCHI. When occurring, we examined what role successful contact attempts had on diarist behavior. We examined this in two ways – how contacted diarists generally differed from non-contacted diarists in the number of overall entries they recorded (Table 6f), and how receipt of midweek contacts affected the login behavior of online diarists (Table 6f1).

⁵ Among online diarists with logins, the average period from diary placement to the first login was 3 days, with a median lag of only 1 day.

⁶ Wilcoxon-Mann-Whitney; H=-3.2, Pr>Z 0.0014

⁷ This was little different than the comparable 14 percent rate for the [Incentive Diaries Feasibility Test](#).

Table 6f. Effectiveness of midweek diary contacts on number of diarist entries* (N=891)

	N	Mean Entries	Median Entries
With Contact	88	23.0	19
No Contact	803	21.5	14

*made by diarists (so excludes recall/receipt entries)

The data in Table 6f reflects the low proportion of diarists receiving midweek contacts (or having a record of receiving them), with only 88 of the 891 diarists placed with diaries and having a CU member with entries receiving a successful midweek contact. Diarists receiving a midweek contact had slightly more entries on average than those who did not (23 and 21.5 respectively), a difference that was not statistically significant⁸. The second analysis focused only on online diarists, but more closely associated the midweek contact with diarist behavior. This was possible through paradata, which collected the timestamp of when a diarist logged in to the online diary. This allows distinguishing activity before and following the midweek contacts (whereas expenditure entries can be recorded up to several days after their actual occurrence). We focused on online diarists to determine: 1) if contacted diarists had an increase in logins subsequent to contact, and 2) if contacted diarists had more logins than non-contacted diarists (regardless of the time of contact) (Table 6f1).

Table 6f1. Effectiveness of mid-week contacts in obtaining additional logins among online diarists

	N (Diarists)	Mean Logins Prior	Mean Logins Subsequent*	Mean Total	Min Total	Max Total
Contact	25	2.2	2.2	4.4	0	16
No Contact	301	NA	NA	3.9	0	26

*when diarists received multiple midweek contacts, the earliest contact was used for this analysis

There was no evidence that midweek contacts affected the subsequent login behavior of contacted diarists, though the small number of diarists contacted limits the conclusions we can draw. Table 6f1 indicates contacted diarists had an average of 4.4 logins and those not contacted had an average of 3.9 logins⁹. It is not clear whether the late timing of some contacts (after day 3) meant diarists had already entered most of their expenses or whether diarists had any negative reactions to the type of contact (e.g., telephone, text or email¹⁰) that they received which limited their subsequent participation.

⁸ Wilcoxon-Mann-Whitney; $H=1.05$, $Pr>Z$ 0.29. Separate analysis that examined contact attempts found that a FR merely attempting contact (and possibly leaving a message) was not associated with a higher number of diarist entries; diarists only 'receiving' contact attempts actually had slightly fewer entries than those who did not.

⁹ Difference not statistically significant. Wilcoxon-Mann-Whitney; $H=0.66$, $Pr>Z$ 0.509

¹⁰ As noted in Table 6a, some midweek contacts also took the form of a personal visit.

d. Extent of diarist participation

Among the 520 CUs with an overall POC disposition of complete, compliance with recording entries in the diary was widespread. As seen in Table 6g, 87 percent of eligible diarists completed a diary – either by making entries or designating that they had no expenses to record¹¹.

Table 6g. Diary status of eligible diary recipients

	n	Percent
Overall	1,008	100%
Refused at placement	17	1.7%
Assigned diary	991	98.3%
Diary - not completed	88	8.7%
Diary - completed	903	89.6%
Completed - with total recall	28	2.8%
Completed - diary entries*	875	86.8%

* 'completion' includes legitimate blank diaries (i.e. diaries that had a “no expenses” checkbox checked)

Diarists who were assigned a diary but never made entries or participated in any form of recall (e.g., ‘total recall’¹²) made up the largest component of non-completion for the diary component of the POC. These 88 diarists were evenly comprised of those not present at the FR visit and those present. Part of the high completion rate can be attributed to the definition of POC complete cases, which required at least one diary in a CU to be completed. We expect many CU members participated in recording entries because they were informed they would be mailed a \$20 incentive for participating. When breaking out CUs by size, over 70 percent of the 343 multi-diary CUs involved all eligible diarists directly recording an expenditure or designating that they had no expenses to record (Table 6h).

Table 6h. Percent of CUs with 100% diary completion, by size*

	N	Percent
Single Diary	177	99%
Multi-Diary	343	71%
Total	520	80%

* 'completion' includes legitimate blank diaries, but excludes diaries with all entries from receipts/recall (the later were acceptable for a CU to be considered “complete” for response rate purposes).

¹¹ 18 legitimate blank diaries (i.e. diaries that had a “no expenses” checkbox checked) counted as completing.

¹² Total recall involves the respondent only providing expenditure entries through the recall and receipts process with the FR at the start of the second visit.

e. Effectiveness of incentives

As discussed in the overview, incentives were incorporated into the design of the POC as outlined in the redesign plan, including a prepaid \$2 bill incentive mailed with the advanced mailing, a \$20 debit card handed to the respondent upon completing the recall interview, \$20 debit cards mailed to each diarist completing a diary, and a \$20 debit card mailed to the respondent after completing the records interview with an additional \$20 debit card included if records were used.

In the respondent debriefings, CUs that indicated he/she received the advanced letter were asked whether or not they received the \$2 bill incentive. The biggest concerns with mailing a cash incentive are that it will get thrown away, or that someone other than the respondent will receive it. In the POC team's analysis plan, the team determined "More than 50% of CUs say yes" to be the evaluation of success, but a much larger percentage, 84.7 percent, of CUs that responded to this debriefing question responded "Yes" (Table 6i). This is a positive result; however, it's important to note that we don't know the outcome of the non-responders in the POC since only responders were asked the debriefing questions. The result does suggest that among participants completing the debriefing that the prepaid incentive was memorable.

Table 6i. Percent receiving \$2 incentive, CUs completing the debriefing (asked if respondent indicated he/she received the advanced letter) (N=530)

Question	Response options	% of CUs
In the first letter we sent you, explaining the study and asking for you to participate, we included a two dollar bill. Do you remember getting this money?	Yes	84.7%
	No	14.5%
	Don't Know	0.4%
	Refused	0.4%

All complete CUs, by definition, should have completed each component of the POC interview. As expected based on this definition, 100 percent of complete CUs received the \$20 recall incentive and at least the partial (\$20) records incentive (Table 6j). Of all complete CUs, 89.6 percent received the full records incentive (\$40), indicating that they not only completed the records interview sections, but also provided at least one record. Looking at Type A CUs, 27.5 percent received the recall incentive and 14.7 percent received the records incentive. These percentages demonstrate the amount of incentives that went to incomplete interviews.

Table 6j. Percent of CUs receiving conditional incentives

	Complete (N=520)	Type A (N=524)	All (N=1229)
Recall Incentive	100%	27.5%	54.2%
Records Incentive	100%	14.7%	48.6%
Partial (\$20)	10.4%	3.4%	5.9%
Full (\$40)	89.6%	11.3%	42.7%

For the diary incentive, each eligible CU member (at least 15 years old) received the incentive if he/she entered at least one expenditure or marked that no expenditures were made for the week. Of the 1,008 eligible members of complete CUs, 87.1 percent received the diary incentive (Table 6k). Breaking this number down further, 95.9 percent of main respondents and 77.8 percent of other CU members received the diary incentive.

Table 6k. Percent of CUs receiving diary incentive

	N	% Receiving Diary Incentive
Eligible Members of Complete CUs	1,008	87.1%
Primary Respondents	517	95.9%
Other CU Members	491	77.8%

While difficult to accurately measure the effect of incentives on CU participation, the effect was estimated by a question in the respondent debriefing that asked respondents if the money he/she received or expected to receive impacted his/her decision to complete the various survey components. Of the 591 CUs that responded to this debriefing question, 42.6 percent reported “Yes,” that the incentives did have an impact, while 56.7 percent reported “No” (Table 6l). Social desirability and other factors may have influenced some CUs’ responses to this debriefing question, but the responses reveal that for the majority of CUs, incentives did not impact their decision to respond to the survey or complete the diary. Despite this, a large percentage of CUs (42.6 percent) still indicated that the incentive influenced their decision to participate.

Table 6l. Effect of incentives on CU participation (N=591)

Question	Response options	% of CUs
Did the money you received or expect to receive impact your decision to respond to the survey or complete the diary?	Yes	42.6%
	No	56.7%
	Don’t Know	0.7%

VII. Records and Recall Interviews

a. Length of the interview elements

The redesign plan indicates that Visit 1 and Visit 2 are each intended to last 45 minutes, on average. In the analysis plan for the POC, the team provided a little more leeway to account for test conditions and set a guideline of “less than 60 minutes” for these visits. This 60 minute duration was also the threshold for the recall and record interviews, which are components of Visit 1 and Visit 2, but are not required to be collected in a fixed order or all in one visit. These guidelines were met for all of these POC survey components, and the longest component, with an average time of approximately 50 minutes (Table 7a), was Visit 1. Visit 1 included identifying CU members and characteristics, conducting the recall interview, placing diaries, instructing the main respondent on record collection, and some follow up items like scheduling the subsequent visit. The mean total time of both visits, including the debriefings, was 116.7 minutes, just under 2 hours. Visit 2 included the pickup of diaries and any associated recall (e.g., recording entries through “total recall”), conducting the records interview, and follow-up in the “back” section (which included a thank you screen as well as administrative type questions involving FR feedback). The mean duration was greater than the median duration for every survey component, indicating that the means could be affected by a small number of large values (e.g. the maximum total time to complete the survey is 281.2 minutes, or, 4.7 hours).

Table 7a. POC survey components: length in minutes (N=520)

	Mean	Min	Median	75 th Percentile	90 th Percentile	Max
Visit 1	50.1	5.6	44.3	64.1	84.5	169.3
Front/Control/Coverage	16.7	2.1	11.0	21.7	36.9	130
Recall Interview	15.8	0.9	12.8	20.8	31.7	70.1
Diary Placement	9.9	0.2	8.4	13.1	20.0	37.7
Records Instructions	1.9	0	0.5	2.6	5.7	18.5
Back (Minus Debriefing)	5.8	0.3	4.9	8.2	11.4	20.3
Visit 2	49.6	9.3	44.3	62.5	85.7	170.3
Diary Pickup	2.8	0	1.1	3.0	7.3	31.2
Add Diary Record	1.2	0	0	0	0.1	83.0
Records Interviews	39.9	6.9	35.1	52.7	72.3	121.5
Back (Minus Debriefing)	5.8	0.3	4.9	8.2	11.4	20.3
Respondent Debriefing	6.7	1.0	5.8	8.1	11.7	31.7
FR Debriefing	10.3	1.5	8.2	13.3	19.4	44.0
Total Time	116.7	23.1	109.4	146.6	188.4	281.2

When looking at the time to complete the POC interviews, it’s important to note that the POC deviated from the redesign plan by excluding some expenditure sections from collection. This was done with the expectation that when the redesign is fielded in production, the total number of questions will be reduced since less detail is required on expenditures and the instruments will be designed more efficiently. These two factors lead to the expectation that the total amount of time to complete the in-person portions of the

interview will be reduced. This should be kept in mind when comparing the times for each visit to the intended time of 45 minutes. Until the full set of redesign questions is complete and a new instrument programmed, we won't have a completely accurate picture of the length of the planned interviews.

Comparison of the POC recall interview sections with the same sections in RP shows a small difference between the two (Table 7b). While the mean time is greater in POC than in RP for every section except section 19, the maximum time is greater in RP than in POC for all sections. The resulting difference in the mean total time is 3 minutes, with a difference in the median total time of 2.5 minutes. While this difference between the POC and RP total recall interview times appears minor, the difference was found to be statistically significant.¹³ This result is surprising given that each recall interview section is in the same format as the current production CEQ section that it is being compared to. While we are concerned with time as it relates to respondent burden, we also acknowledge that longer interview times are associated with more reporting of expenditures. Since the administration of the POC recall sections was very comparable to RP, a higher amount of time may be the indication of additional expenditure reports, perhaps due to an increase in respondent motivation caused by incentives. The analysis of expenditure reports will be completed for Report 3 and will shed additional light on the topic.

Table 7b. POC and Production recall interview sections: length in minutes

Section	Section Title	POC (N=520)				RP (N=1,483)			
		Mean	Min	Median	Max	Mean	Min	Median	Max
6	Appliances and Household Equipment	3.5	0.3	2.6	23.5	2.8	0.1	2.0	25.2
8	Home Furnishings	3.0	0.2	2.4	19.6	1.9	0.1	1.3	29.8
17	Entertainment Expenses	3.0	0.1	2.4	14.2	2.4	0.1	1.6	27.8
18	Trips and Vacations	3.7	0.1	1.5	32.4	2.7	0.0	0.8	38.8
19	Miscellaneous Expenses	2.5	0.1	2.0	21.9	3.1	0.2	2.3	23.7
All Recall Sections		15.8	0.9	12.8	70.1	12.8	0.8	10.3	68.2

The records interview is a new interview format and a key component of the CE redesign. To gain more insight into the performance of the records interview, we compare the records section lengths to the same sections in RP as well as to production cases that have Records = 1, an indication that the CU “uses records always or almost always” (Table 7c)¹⁴.

¹³ Wilcoxon-Mann-Whitney; H=5.92, Pr>Z <0.0001

¹⁴ Note that due to a limited sample size of RP cases, this comparison uses all production CEQ cases with record use indicated.

Table 7c. POC and Production records interview sections: length in minutes

Sect.	Section Title	POC (N=520)				RP (N=1483)				Production, Records = 1 (uses records always or almost always) (N = 885)			
		Mean	Min	Median	Max	Mean	Min	Median	Max	Mean	Min	Median	Max
1	Housing Characteristics*	1.6	0.2	1.1	19.3	0.3	0	0	10.6	0.3	0	0	10.6
2	Rented Homes	1.7	0.2	1.2	18.6	1.0	0	0.5	19.5	0.7	0	0	6.8
3	Owned Homes	5.3	0.1	3.8	28.0	3.0	0	1.6	26.7	4.2	0.0	2.8	50.0
4	Utilities and Fuels	6.2	0.4	5.2	37.7	4.8	0.2	3.7	40.9	7.1	0.4	5.9	43.8
10	Rented and Leased Vehicles	0.9	0.1	0.5	19.6	0.6	0.1	0.4	16.0	0.6	0.1	0.4	5.4
11	Owned Vehicles	3.8	0.1	2.9	26.5	1.5	0.0	0.4	28.1	2.0	0.0	0.6	22.5
13	Insurance Other than Health	3.4	0.1	2.5	15.8	1.8	0.0	1.2	25.3	2.9	0.1	2.1	25.3
14	Health Insurance	2.9	0.2	2.1	24.9	1.7	0.0	1.1	20.3	2.3	0.1	1.6	21.0
15	Medical Expenses	2.6	0.1	1.7	24.4	2.0	0.0	1.3	18.9	3.4	0.1	2.4	26.8
21	Income*	7.7	0.6	6.2	30.4	3.7	0	2.3	36.2	4.6	0	3.1	39.4
22	Assets and Liabilities*	3.8	0.3	3.0	27.5	0.8	0	0	20.1	1.2	0	0	22.5
All Records Sections		39.9	6.9	35.1	121.5	21.0	2.6	17.6	109.3	29.2	2.6	24.8	115.5

*In RP and production samples, Section 1 is only asked in the first interview and for new CUs, Section 21 is asked in the first and fourth interview as well as for new CUs, and Section 22 is asked only in the fourth interview

As the team expected based on previous research indicating longer interview times when records are used, the total length of the records interview is longer than the total of the respective sections times in RP, with the mean time of every section being greater in POC than in RP. The mean total time of the records sections in POC is 18.9 minutes greater than that of RP, a statistically significant difference.¹⁵ The difference in the mean total time between POC and Records = 1 production cases, 10.7 minutes, is not as large but still statistically significant.¹⁶

¹⁵ Wilcoxon-Mann-Whitney; H=19.61, Pr>Z <0.0001

¹⁶ Wilcoxon-Mann-Whitney; H=9.86, Pr>Z <0.0001. For some sections (e.g., Utilities and Fuels) the RP with records group had a longer average duration.

An important difference between POC and production, however, is that for production, certain sections are only asked in the first or last interview. Housing characteristics (Section 1) is asked in the first interview and for new CUs, income (Section 21) is asked in the first and fourth interviews as well as for new CUs, and assets and liabilities (Section 22) is only asked in the fourth interview. For a given production quarter, interview number varies across CUs, so the lengths shown in Table 7c take into account all cases, even those that were not asked a particular section during that quarter. This explains the unusually low numbers, including zeroes, in some of the sections' minimum and even median lengths for RP and Records=1 Production. Despite this survey design difference, we compare POC to all RP and Records=1 Production cases instead of sub-setting by interview number because the ultimate goal is to compare the average POC interview to the average production interview. Through this comparison, it is evident that the POC records interview is in fact longer than the total of the respective sections in production, but still within the guideline set by the team of "60 minutes or less."

b. Use of records

In Visit 1 of the POC test's survey design, CUs were given instructions for collecting records in preparation for the records interview in Visit 2. To determine respondent cooperation, we first look at the FR debriefing question: "About how many records did the respondent have ready at the start of Visit 2?" According to the responses (excluding the 19 "other" responses) approximately 91 percent of CUs provided at least one record, a percentage far exceeding the guideline of 75 percent set by the POC team in the analysis plan (Table 7d). Additionally, the majority of the 19 "other" responses had indication of record use, mostly electronic record use, in the follow-up text question asking the FRs to specify.

Table 7d. Number of records provided by CUs (N=518*)

# of Records	Number of CUs	Percent of CUs
None	48	9.3%
One or Two	64	12.4%
Three to Five	94	18.2%
Five to Ten	106	20.5%
Ten to Fifteen	86	16.6%
More than Fifteen	101	19.5%
Other	19	3.7%

*These are complete cases that also had an answer in the FR debriefing for this question.

However, there are lower rates of records use in any one given section compared to record use at the level of the interview as a whole. Looking only at the CUs reporting expenditures for each section and the percent of them that provided at least one record, the percentages range from 39.2 percent for section 2 on

rented homes to 78.6 percent for section 4 on utilities and fuels (Table 7e). Text analysis of some of the respondent debriefing questions could give further insight into which records were difficult to collect and why.

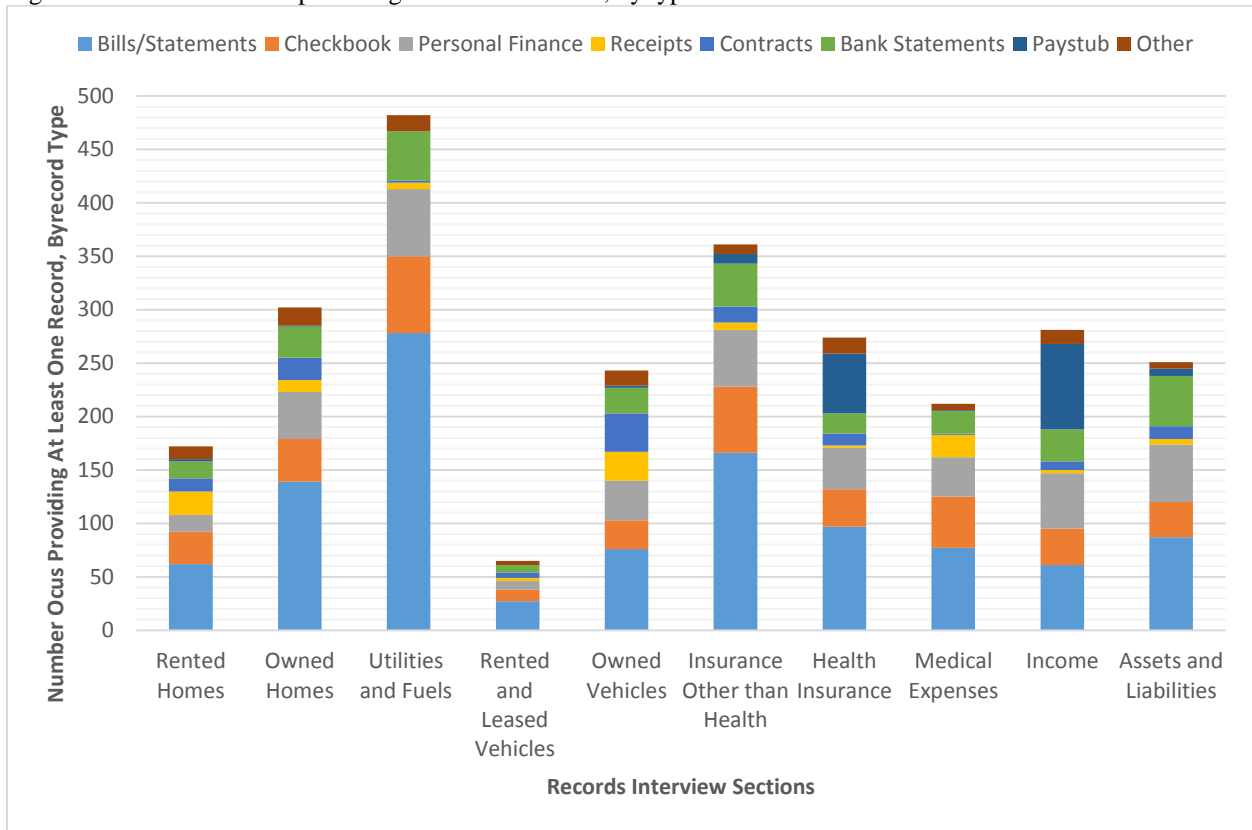
Table 7e. Percent of CUs reporting expenditures that provide at least one record, by interview section

Section	Section Title	N	% (of N) Providing at Least One Record
2	Rented Homes	222	39.2%
3	Owned Homes	318	64.5%
4	Utilities and Fuels	510	78.6%
10	Rented and Leased Vehicles	57	49.1%
11	Owned Vehicles	439	40.1%
13	Insurance Other than Health	449	66.8%
14	Health Insurance	294	66.0%
15	Medical Expenses	313	52.1%
21	Income	502	45.0%
22	Assets and Liabilities	471	41.2%

Figure 7a breaks record use down even further, looking at interview section as well as record type.

Certain types of records will provide more accurate data for the survey. For example, bills should provide the exact expense at the level of detail required in the CE. On the other hand, bank statements may only have the total amount for multiple items purchased at a store, where we would want each item reported separately. For this reason, use of records such as bills/statements, receipts, contracts, and paystubs are all more desirable. Despite this, quality can still be gained by use of other types of records. For example, the use of checkbook registers or credit card statements may cue the respondent to recall additional expenses that would otherwise have been forgotten.

Figure 7a. Number of CUs providing at least one record, by type of record and interview section



VIII. Individual Diaries

a. Diary completion

Diarists had the option of completing the diary via either online or paper mode, and some diarists opted to switch from the online mode to paper. Table 8a shows the final mode of completion for diarists assigned either an online or paper diary.

Table 8a. Diary type of completion by assigned diary mode

Individual diary assigned	Paper	Online	Mix Paper/Online	Total Recall	Total*
Paper diary (N=646)	578	NA	NA	8	586
Online diary (N=345)	14	257	8	20	299
Total	592	257	8	28	885

*total indicates number of diarists with at least 1 expenditure (so excludes 18 legitimate blank diaries)

Of note, a small proportion – 4 percent – of online diarists switched to completing a paper diary, and a similar number – 6 percent – ended up completing the diary only through the recall and receipts process ('Total Recall'). Fewer paper diarists – about 1 percent – provided total recall diaries. The limited extent of online diarists switching to paper may call into question the need to provide paper diaries as backups for online diarists. For most of the 14 diarists who switched to making all entries using a paper diary, there was no indication in debriefing responses as to what accounted for their switch (8 diarists/57 percent). For 4 diarists, they were unable to login either when they tried in the FR's presence or subsequently (29 percent). The remaining diarists either mentioned having a computer virus or simply neglected to make online entries. Among diarists making some entries using both modes, when indicated, debriefing notes suggested they were uncomfortable using online diaries, preferred paper, or lost their login credentials.

Also of interest is the frequency with which diarists resorted to telling FRs their expenses instead of recording them (Table 8b).

Table 8b. Extent of total recall individual diaries*

	N	Percent (of N)
POC	885	3.2%

*among diarists with any expenditures provided

Only 3.2 percent of diarists with some expenditure information provided diaries involving total recall instead of recording entries themselves. Their awareness of the \$20 incentive for participation in the diary-keeping task may have limited the number without entries in their diaries (though FRs were not instructed to communicate to diarists the specific mechanisms for qualifying for the diary incentive).

Analysis of total recall at the CU-level also reveals a very low prevalence of this process (Table 8c).

Table 8c. Prevalence of all total recall CUs

	N	Percent (of N)
POC	520	2.3%
RP-Diary	291	8.9%
IDFT	210	21.9%

Only 2.3 percent of POC CUs provided diaries involving total recall, compared to over 3 times that number (8.9 percent) for the comparable restricted production diary group. To establish equivalent comparison groups with the Individual Diaries Feasibility Test (IDFT), it is necessary to limit the POC CUs to the 148 CUs only assigned online diaries. Among this comparison group, only 4.7 percent of POC CUs provided diaries involving total recall compared to 21.9 percent of IDFT CUs.

Table 8d further examines diary completion. Among individuals assigned paper diaries, 89 percent completed the diary by making one or more entries. Very few who were assigned paper diaries did not participate in this mode (7 percent). Among those assigned online diaries, 81 percent completed the diary by making one or more entries, and 13 percent did not provide any diary entries (through recall or receipts) nor indicate to having no expenditures during the week.

Table 8d. Individual diary completion

Mode Assigned	Completed	Uncompleted		% Total Recall	Total
		% Non- Participation*	% Legitimate Blanks		
Paper (N=646)	89%	7%	2.6%	1%	100%
Online (N=345)	81%	13%	0.3%	6%	100%

*online diarists who switched to participating by paper diaries are not counted among '% Non-Participation'

Table 8e examines diary completion at a CU-level, among CUs with multiple diaries placed. The table indicates the percent of CUs with all of their placed diaries completed.

Table 8e. Percent of CUs with full participation, among multi-diary CUs*

	N	Percent (of N)
POC	324	77%
POC online-only	91	75%
IDFT	112	43%

*excluding total recall-CUs, non-placed diaries (and not counting diaries with only recall as participation)

When excluding CUs with total recall, 77 percent of the 324 multi-diary POC CUs had all diarists who were placed with a diary completing a diary. To enable comparisons with the IDFT, the rate is 75 percent for online-only multi-diary CUs in the POC. This compares favorably with the 43 percent of the online, multi-diary CUs with full completion in the IDFT.

In conclusion, the extent of participation within CUs was much broader for the POC than in prior tests. This may be attributable to the incentives offered for diary completion.

b. Respondent experience

i. Diary-level preference

Table 8f shows the distribution of responses to the two questions in the respondent debriefing related to the individual diary, a key component of the CE redesign. Based on the 374 CUs that responded to these debriefing questions, 68.2 percent of CUs actually prefer a single diary for the whole household, the way production CED is currently designed, over the new individual diary design. Further, 29.4 percent of diarists reported recording expenditures on behalf of other household members in the individual diary. However, this does not necessarily indicate that the respondent recorded expenses for other eligible diarists, as “other household members” could include ineligible CU members (children under 15). Furthermore, it is important to keep in mind that these respondents did not actually experience both survey designs, but only the individual diary design.

Table 8f. Respondent preference for CU vs. individual diary (asked if there is more than 1 CU member) (N=374)

Question	Response options	% of CUs
Would you prefer having a single diary for the whole household or having an individual diary for each member of the household?	Prefer single diary for the whole household	68.2%
	Prefer individual diaries	30.2%
	Don't know	1.1%
	Refused	0.5%
Did you record any expenditures on behalf of other household members into your diary (excluding shared household expenses)?	Yes	29.4%
	No	70.6%

ii. Online diaries

Changing Username/Password

The tables below include all online diaries that had a paradata record, but are not necessarily online diaries from CUs defined as “complete” for the POC. To assess the usability of the diaries that were fielded, we examined the ease of certain tasks – changing usernames or passwords, logging in, or being able to login after one or more initial failures – as well as debriefing responses from FRs.

Table 8g presents the frequency of username or password changes. According to the data very few (18) of the approximately 350 online diarists tried to change their username or password – 5 percent. Given the variety of characters and digits in the provided usernames and passwords, this was a surprising finding.

Table 8g. Percent of diarists who were able to change their username or password

	N	Percent (of N)
Username or Password Change	18	39%

In order to successfully change either of their credentials, a diarist had to provide an email address, and a security question with response options (Figure 8a). If an email (or security question/answer) was entered at the change username screen, the diarist was not required to reenter these on the change password screen.

Figure 8a. Screenshot of change password diary page

United States
Census
Bureau

CONSUMER EXPENDITURE
DIARY SURVEY

Need Help? Call the Help Desk toll-free at 877-744-1522 or send an email to NPC.CE.POC.Helpdesk@census.gov

Change Password

Passwords must contain all of the following:

- At least 8 characters
- At least 1 uppercase letter
- At least 1 lowercase letter
- At least 1 number
- At least 1 special character from the following: ! # \$ * & ? ~

Current Password:

New Password:

Confirm New Password:

Please provide a valid email address, select a security question, and provide an answer for your security question.
Field is required.

Email:

Confirm Email:

Security Question:

Security Answer:

[Cancel](#) [Change password](#)

Fewer than half of diarists who tried to change their credentials (as defined by navigating to the change username or change password screens) eventually did change them. The requirement to provide an email may have dissuaded some diarists from completing the process if they were not comfortable providing this identifying information. Thus the low percentage of credential changes may reflect an unwillingness on the part of diarist to provide all of the needed information, and not an inability to fulfill the requirements of changing their credentials. We additionally examined the duration of time diarists spent on the change password page (Table 8h).

Table 8h. Time diarists spent on change password page (in minutes: seconds) (N=17¹⁷)

	Mean	Median	Min	Max
Overall	0:57	0:39	0:16	3:38

¹⁷ One diarist who never entered a new password (nor confirmation entry of new password, nor answer to the security question) was excluded from analysis.

The durations in Table 8h were calculated from the first entry on the page to confirmation entry of the new password or their answer to the security question (whichever came later, as some respondents had entered the later as part of the change username page). The change process did not take long for most diarists – averaging just under 1 minute – ranging from 16 seconds to 3-and-a-half minutes. Not surprisingly, diarists took longer entering their (new or old) password, which had complex combinations, and entering a response for the security question than they took entering their email address.

From their debriefing responses, most diarists reported not having problems with using the online diary nor entering their expenditures (Table 8i). However, there were indications of difficulty with the login process, as two-thirds of those responding to the question indicated having trouble logging into the diary.

Online diary issues

Table 8i. Online diary issues identified by the debriefing (asked if an electronic diary was placed with at least one person in the CU)

Question	N	Response options	% of Respondents
Did you have problems using the online diary?	207	Yes	24.2%
		No	75.8%
Did you have problems logging into the diary?	50	Yes	66%
		No	34%
Did you have any problems entering your spending?	32	Yes	31.2%
		No	68.8%

Log-in Successes & Failures

We further looked at how frequently online diarists logged in to their diary instrument by type of device they used (Table 8j). It should be noted that FRs were instructed to assist the diarists logging in initially at the end of the visit 1 interview, however it is unclear to what extent this occurred¹⁸.

Table 8j. Online diarist login success frequencies, by type of device*

	N	Mean	Median	Min	Max
Mobile Only	81	3.4	2	0	26
Mix	57	6.1	5	1	26
Desktop Only	188	3.5	3	0	13
Overall	326	3.92	3	0	26

*for those without successful logins, type of device assigned according to device used for login attempts

¹⁸ Debriefing responses asked about respondent logins in the presence of an FR, but were only answered by 52 percent of online diarists – of these, 55 percent reported logging in the presence of an FR.

Diarists logging on only by mobile diaries had the fewest average and median number of logins (3.4 and 2 respectively), though these numbers were similar to those for diarists logging on using desktop diaries. Of interest, diarists who were able to access both modes were seen to log in much more frequently than those using a single mode with over 6 logins on average and a median of 5. This comports with diarists' feedback from the IDFT that they wanted the option to login using multiple devices and represents a promising feature of the fielded diaries. Overall, 92.6 percent of diarists with paradata records had at least one successful login¹⁹.

We additionally analyzed the extent to which diarists had login failures (Table 8k). Login failures are defined as users that successfully entered a username, but not a password. If the user could not successfully enter a password, then they were not counted in the failure count.

Table 8k. Online diarist login failure frequencies

	N	Mean	Median	Min	Max
Those w/1+ Failure	154	3.3	2.0	1	21
All Online Diarists	326	1.6	0.0	0	21

There were an average of 1.6 login failures for diarists having paradata records. For diarists having at least one login failure, the average number was 3.3. As a proportion of all diarists with paradata, a little under half had one or more login failures (47.2 percent). Combined with the finding that 92.6 percent of diarists had at least one successful login, this gives an indication that most diarists were able to overcome any initial difficulties with the login process. This is illustrated in Table 8k1.

Table 8k1. Prevalence of “no successful log-in” among online diarists with at least 1 log-in failure, by mode

Online diary type	N	Percent (of N)
Mobile Only	30	23%
Mix	38	0%
Desktop Only	86	20%
Overall	154	16%

This table shows that only 16 percent of diarists (from 20 CUs) were unable (or unwilling) to log in to make entries in their online diary following a failure to login to the diary. As with logins (Table 8i),

¹⁹ Two factors had the potential to inflate login counts: 1) the FR logging in with the diarist at the end of visit 1, and 2) diary sessions would close after a 15-minute period of inactivity (which could potentially require a diarist to login again).

diarists using a combination of modes appeared most proficient in overcoming any initial difficulties they had, with all of these diarists eventually logging in successfully.

Distribution of Logins

One the goals of introducing an online diary to the redesign plan was to reduce measurement error by providing a method for diarists to report expenditures as they occur and help reduce recall error. We could examine this aspect of diary-keeping by using the timestamps associated with diary logins in the paradata. As an indication of recording entries close to when they were incurred, we would expect to see logins at all times of the day. The results are shown in Table 8l below.

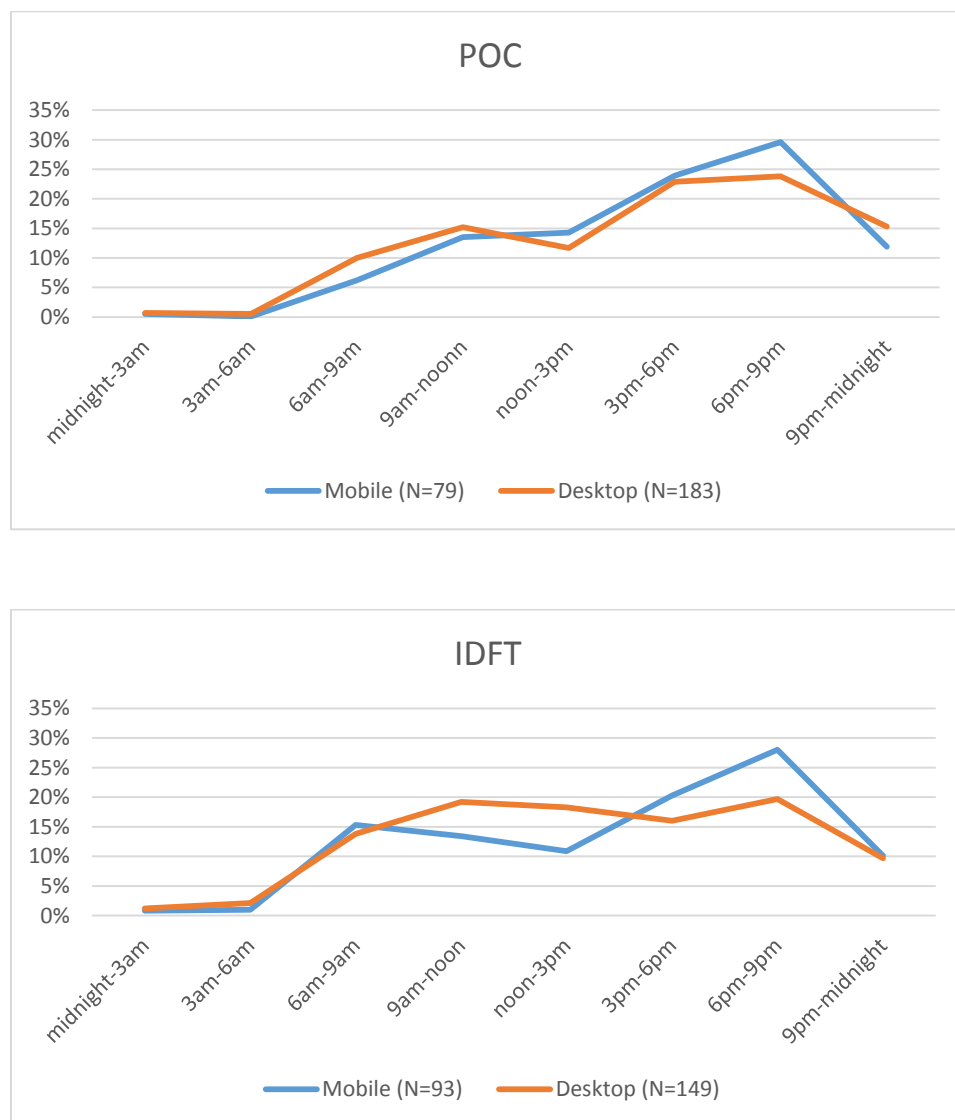
Table 8l. Extent of contemporaneous reporting: proportion of diarist's logins by segment of day

	POC*	IDFT
Midnight-3pm	38.1%	49.6%
3pm-midnight	61.9%	50.4%

* POC numbers exclude diarists whose logins spanned (mobile and desktop) modes & exclude recall entries.

Among POC online diarists with logins in either mobile or desktop mode (but not both), there were an average of 38 percent of logins occurring in the earlier portion of the day (from midnight to 3pm). This was a lower proportion than seen among IDFT online diarists, which had almost half of logins during this time period. The average proportion of CU logins at various time segments throughout the 24-hour period are displayed for those with desktop or mobile logins in Figure 8b.

Figure 8b. Proportion of diarist logins by time of day in the POC and IDFT tests



There was very little divergence in the pattern of logins between mobile and desktop diarists in the POC. We would have expected more mobile logins earlier in the day, as these diarists would be expected to have access to their diaries in those hours. Comparing POC and IDFT patterns, the overall patterns differed in magnitude but not trends, with the largest share of logins occurring between 6pm and 9pm. One possible reason for more diarist logins later in the day for POC would be that diarists waited to enter larger receipts until they had access to a desktop computer, whereas they gave these receipts to the FR for the IDFT. Whatever the case, there was no evidence to suggest this had a harmful impact on data quality.

IX. Individual Diaries Data Quality

This report examines the data quality resulting from individual diaries based on the number of entries provided, and the extent of item non-response for the various fields diarists were asked to fill in for their expenditures. At the expenditure category level, diarist entry frequencies for POC are presented in Table 9a, excluding diaries in which all entries were entered by the FR through total recall.

Table 9a. POC Diary entry frequencies, by category (N=858)*

	Mean	Median	Min	Max
Food (Home)	10.9	4	0	126
Food (Away)	4.2	3	0	59
Clothing	1.0	0	0	18
Other	6.6	5	0	59
Total	22.7	15	1	161

*Excludes total recall diaries

There were an average of 22.7 entries among the 858 diarists. Almost half of the entries were made for food for home consumption, with about a third made in the ‘other’ category.

Table 9b presents these findings at the CU-level for all complete CUs, regardless of total recall status, with week 1 expenditures from the RP sample appearing in the subsequent table.

Table 9b. POC CU entry frequencies, by category (N=520)*

	Mean	Median	Min	Max
Food (Home)	18.4	12	0	197
Food (Away)	7.0	5	0	67
Clothing	1.6	0	0	52
Other	11.2	8	0	118
Total	38.2	29	1	434

Comparable CED (Restricted Production) CU entry frequencies, by category (N=291)*

	Mean	Median	Min	Max
Food (Home)	16.3	12	0	113
Food (Away)	4.9	3	0	32
Clothing	1.4	0	0	31
Other	8.7	7	0	36
Overall	31.3	28	1	182

*Both tables exclude non-eligible (e.g., '217') cases and legit blank diaries for POC & RP; RP is only week 1 entries

There were an average of 38.2 entries among POC CUs, compared to 31.3 entries for the equivalent RP CUs. The median entry totals were more similar between POC and RP samples (29 and 28 entries, respectively), with the overall difference in total entries between samples being statistically significant²⁰. At the section-level, although there were small differences in the entries across the four categories compared to the RP sample, the mean and median number of entries were at least the same if not larger in the POC. In addition, the largest differences were in the “Food Away from Home” Category and “All Other Products, Services, and “Expenses” category, two categories that we suspect would be underreported in when relying on a proxy reporter as is the case in the RP sample. Relatedly, the greater number of entries for the ‘All Other Products, Services, and Expenses’ category in the POC compared to the RP sample suggests that POC diarists were accounting for common household goods in their entries; in IDFT there was a concern that these goods were being underreported.

Another data quality measure is the extent of item non-response among diarists. In the diary, it is not possible to know if a diarist failed to report an expenditure (as one would typically think of as item non-response); however, if an entry was made, we do know when a field related to that entry was left blank. This is how item non-response is evaluated for our purposes. Respondents needed to enter only one field for an item, if additional fields were not entered, this did not trigger an error message. These data on non-response are presented in Table 9c for all entries made by diarists having a POC outcome of complete.

Table 9c. Missing element rates by mode

	Online	Paper
No. of Entries	4,378	14,883
Item Cost*	0.0%	0.5%
Item Description*	0.6%	0.2%
Date	0.0%	NA

*Cost: excludes entries of '0' from tally of non-response

*Description: excludes records where respondents typed 'NOTHING' as the description

Comparable rates for IDFT

	Mobile	Desktop	Paper (IDFT RP)
No. of Entries	3,123	6,481	15,689
Item Cost	0.7%	2.8%	0.7%
Item Description	1.5%	2.5%	0.2%

[Recall item non-response examined separately for IDFT and date item non-response was not available]

²⁰ Wilcoxon-Mann-Whitney; H=-2.95, Pr>Z 0.003

There were very low rates of missing data in the POC diaries, regardless of whether they were online or paper diaries. The only major field exhibiting any item non-response for online diarists was for the description of the expenditure, where there was slightly more missing data than among paper diarists.

Compared to the IDFT (and its ‘paper’ RP comparison group), the POC had lower rates of missing data, although rates were low for all modes. We additionally examined the minor fields diarists would fill out depending on skip patterns determined by expenditure type; those item non-response rates are shown in Table 9c1.

Table 9c1. POC item non-response on follow-up questions

Topic	Online	Paper
MLS (N=3,565)		
Meal type	9.7%	9.2%
Vendor	NA	8.8%
Alcohol type (N=232)	0.0%	19.4%
Alcohol cost* (N=232)	0.0%	4.2%
FDB (N=9,234)		
Package type	5.7%	6.8%
CLO (N=815)		
Sex	3.7%	17.7%
Age	5.6%	15.3%

*Alcohol Cost: excludes entries of '0' from non-response

There were slightly higher rates of missing data for the type of meal among POC online diarists compared to those with paper diaries (9.7 and 9.2 percent respectively). For all other categories, online diarists left less missing data. Among the 232 diarists with alcohol expenditures, none in the online diary mode provided missing data for the type of alcohol purchased, compared to almost a fifth of paper diarists. Similarly, paper diarists were more likely to leave missing data for the characteristics of the individuals (male/female, under 2, 2-15, 16 and older) for whom clothing was purchased.

We additionally analyzed the proportion of receipts that were handed to the FR (instead of being recorded into the diary by the diarist), that contained 5 or more items (Table 9d).

Table 9d. Average percent of CU receipts with 5 or more items*

	N	Percent
Overall	145	46.9%

*from FR Debriefing: 29.8% of CUs had receipts for expenditures that they did not enter in the diary (145 CUs).

From the FR debriefing, almost half of the receipts that CUs gave to FRs contained 5 or more items. FRs reported that 29.8 percent of CUs had receipts (of any size) that they handed to the FRs. Of these CUs who provided receipts, the average CU provided 6.3 receipts (and gave an average of 3.1 receipts containing 5 or more items). Overall, although most diarists recorded entries themselves, when receipts were provided to FRs as part of the receipt and/or recall process, they tended to have at least 5 items listed.

In summary, POC diarists provided entries that matched or exceeded those made by their paper diary RP counterparts. Additionally, POC diarists completing online diaries had item nonresponse rates that tended to be lower than their POC paper counterparts. Given this, POC diarist performance indicates eligibility criteria to complete an online diary would not need to be tightened, and could potentially be expanded. The absence of diarists being randomly assigned to mode, however, means that differences in sample composition could have accounted for the performance differences observed between paper and online diarists.

X. Overall Respondent Experience

a. Analysis of respondent debriefing questions

Respondent perception of survey time and burden

While not the primary goal of the redesign, the reduction of respondent burden is a secondary objective. As such, respondents were asked about perceived burden in the POC debriefing. The debriefing questions were asked of a respondent within each CU at the end of the records interview, regardless of whether they were a complete CU for the POC. Debriefing respondents therefore may not have completed all components of the POC.

When asked in the respondent debriefing to estimate the amount of time spent on survey-related activities, respondents from complete CUs reported an average time of 35.6 minutes spent filling in the paper or online diary throughout the week. Looking at complete CUs that provided records, these respondents reported an average of 41.8 minutes spent collecting records (Table 10a). The median reported time for both of these survey activities was 30 minutes. The vast majority of respondents that answered these debriefing questions also reported that the amount of time they spent on these activities was reasonable or somewhat reasonable – 95.7 percent for filling in the diary and 97.8 percent for collecting records (Table 10a2).

Table 10a. POC respondents' estimation of time spent on survey-related activities (minutes)

Question	N	Mean	Min	Q1	Median	Q3	Max
How much time did you spend filling in your diary throughout the week?	505	35.6	1	15	30	45	360
How much time did you spend collecting records?	458*	41.8	1	10	30	60	480

*This is a subset of complete cases that responded to this debriefing question. CUs that responded 0 minutes, indicating that no records were collected, were excluded.

Table 10a1. POC respondents' assessment of burden from survey-related activities

Question	N	Response options	% of Respondents
Would you say that the amount of time you spent filling in the diary was	530	Reasonable	85.7%
		Somewhat Reasonable	10.0%
		Somewhat Unreasonable	2.1%
		Unreasonable	2.3%
Would you say the amount of time you spent collecting records was	520	Reasonable	86.5%
		Somewhat Reasonable	11.3%
		Somewhat Unreasonable	1.5%
		Unreasonable	0.6%

Although it is not possible to determine how much time respondents spent on paper diaries, paradata allows for estimation of time spent in the online diary. Looking at diarists placed with the online diary²¹, estimates from the paradata reveal the mean time spent in the online diary to be 18.23 minutes (Table 10b).

Table 10b. Time spent in online diary, device time (N=326)

	Time (Minutes)					
	Mean	Min	Q1	Median	Q3	Max
Overall	18.23	0	5.20	13.45	25.10	97.60

The paradata estimates were then directly compared to the self-reported estimates from the respondent debriefing. Table 10b1 shows the distributions of the self-reported time, the time estimated from paradata, and the paired difference between the two times for the 154 respondents that made entries using the online diary and also self-reported an estimate of time spent in the diary in the debriefing. The mean paired difference between the self-reported and paradata times is 11.8 minutes, with 35.3 minutes as the mean self-reported time and 23.5 minutes as the mean paradata time. While there were respondents that reported a time lower than the paradata-calculated time, the data reveals that for the majority of cases, respondents' perception of the time spent in the diary exceeded the actual time spent. This finding is not surprising, given that paradata can only give estimates based on the amount of time the respondent spent actually in the online diary and does not take into account things like respondents preparing/organizing receipts to enter information into the diary or simply even planning when to sit down during the day to log into the diary. The fact that the median difference between the self-reported and paradata time, 6.3 minutes, is much smaller than the mean difference shows that the mean is skewed by certain cases with a very large positive difference between the two times.

Table 10b1. Length of time spent on online diary in minutes: respondents' estimation vs. device time (N=154)

	Mean	Min	Q1	Median	Q3	Max
Self-reported	35.32	0	15	30	45	360
Paradata from device	23.52	0	10.6	18.4	30.4	97.6
Difference (self-reported–paradata)	11.80	-57.8	-7.7	6.3	19.1	324.2

²¹ This includes a small number of diarists who were placed with online diaries but later switched to using paper diaries. Those cases were excluded from table 10b1.

Records

Analysis on records use (Section 7b) found that a high percentage of respondents provided records. Nonetheless, the percentage of record use varied by interview section, and looking at the debriefing questions helps understand respondents' experiences with records collection. According to the respondent debriefing, 57.8 percent of respondents had electronic financial records such as online bank statements, utility or mortgage statements (Table 10c). Of these 341 respondents that reported having electronic financial records, though 21.1 percent did not access any of their electronic records before the interview.

Table 10c. Use of electronic records

Question	N	Response options	% of Respondents
Are any of your financial records electronic? Such as online bank statements, utility or mortgage statements?	590	Yes	57.8%
		No	41.9%
		Don't know	0.3%
(If answered 'Yes' to question above) Did you access any of your electronic records before our interview?	341	Yes, printed them out	33.4%
		Yes, available electronically	45.5%
		No	21.1%
		Don't know	0%

Furthermore, 16.7 percent of respondents answered that there were records that were difficult to collect, and 8.7 percent of respondents reported that there were records they found, or could have found, that were not used in the interview (Table 10d). Both of these numbers are positive, with not too large of a percentage of respondents reporting "Yes" for either of these questions. However, social desirability to answer "no" for these questions could have influenced responses. Of the 51 respondents that chose not to use certain records in the interview, though, 43.1 percent chose "confidentiality" as the reason.

Table 10d. Respondents' experience with collecting records

Question	N	Response options	% of Respondents
Were there any records that were difficult to collect?	588	Yes	16.7%
		No	83.0%
		Don't know	0.3%
Were there any records that you found, or could have found, that you chose not to use in the interview?	588	Yes	8.7%
		No	91.0%
		Don't know	0.3%
(If answered 'Yes' to question above) What was the reason that you did not use those records?	51	Confidentiality	43.1%
		Not able to download/print	17.6%
		Not able to find	17.6%
		Other	21.6%

Debit card

As discussed in the overview as well as in Section 6e, CUs were handed a \$20 debit card upon completion of the first interview. Then, in the respondent debriefing at the end of the second interview, respondents were asked if there were any problems using the debit card. The responses indicate that 43 percent of CUs did try using the debit card between the first and second interviews, regardless of whether there were issues using it (Table 10e). Since the period of time between the two interviews was the diary keeping week, the prevalence in debit card use could raise concerns about the debit card affecting diary expenditures.

Table 10e. Respondents' use of debit card (N=596)

Question	Response options	% of Respondents
For completing the first interview, you received a \$20 debit card. Did you have any problems using the debit card?	Yes	9.1%
	No	33.9%
	Haven't tried to use it	55.9%
	Don't Know	0.7%
	Refused	0.5%

Help desk

Only 33 respondents, 15.9 percent of those who were asked about help desk use, reported calling the help desk during the week (Table 10f). Most issues related to incorrectly transcribed or forgotten credentials. While only a small number of respondents called the help desk, more than half responded that they did not get the information/assistance that they needed. Based on the call log maintained by the help desk staff, this may have been because respondent calls regarded issues the help desk was not in a position to resolve (e.g., debit cards). Only in a few cases did it appear that respondents did not have their in-scope issue resolved – respondents entering incorrect start dates and not being able to change them, entering an email address incorrectly, or unclear URL error messages. Other unsatisfied requests may have involved FRs calling on behalf of respondents (which could have led to miscommunications), or respondents providing insufficient information regarding problems in voice messages or getting cut off when relaying issues.

Table 10f. Respondents' use and assessment of Help Desk (asked if an electronic diary was placed with at least 1 person in the CU)

Question	N	Response options	% of Respondents
Did you call the help desk during the week?	208	Yes	15.9%
		No	84.1%
(If answered 'Yes' to question above) Did the help desk give you the information or assistance you were looking for?	33	Yes	45.5%
		No	54.5%

Assessment of survey completion instructions

Diarists largely reported that the diary instructions were useful – 95.6 percent finding them ‘Very Useful’ or ‘Somewhat Useful’ (Table 10g). This did not vary significantly by the type of diary that was placed.

Table 10g. Respondents’ assessment of the usefulness of instructions.

Question	N	Response options	% of Respondents
I gave you instructions on how to use the diary, how useful were the instructions?	207	Very useful	76.8%
		Somewhat useful	18.8%
		Not very useful	2.4%
		Not at all useful	1.0%
		Don’t know	1.0%
(Asked if an online diary was placed with at least 1 person in CU) I also left a user guide for help with using the online diary. How useful were those pages?	206	Very useful	63.6%
		Somewhat useful	22.3%
		Not very useful	3.9%
		Not at all useful	8.3%
		Don’t know	1.9%
(Asked if a paper diary was placed with at least 1 person in CU) I provided instructions on how to fill out the paper diary, how useful were the instructions?	381	Very useful	81.4%
		Somewhat useful	15.2%
		Not very useful	0.8%
		Not at all useful	2.1%
		Don’t know	0.5%

Similarly, FRs felt that respondents understood the procedures for diary-keeping (Table 10h).

Table 10h. FRs’ assessment of respondents’ understanding of instructions

Question	N	Response options	% of FRs
How well did the respondent(s) seem to understand the diary process during pickup?	535	Very well	76.1%
		Pretty well	20%
		Not very well	3.0%
		Not at all	0.9%
How well did the respondent(s) seem to understand the diary instructions provided?	597	Very well	71.9%
		Pretty well	25.6%
		Not very well	2.0%
		Not at all	0.5%
How well did the respondent(s) seem to understand the records instructions provided?	664	Very well	70.2%
		Pretty well	27.1%
		Not very well	1.7%
		Not at all	1.1%

XI. Conclusion

The analysis for this report found that the various aspects of the POC – the interview visits, diary placement and pickup, and incentive delivery – were carried out more or less as intended and would be feasible to implement as part of a redesigned survey. Response rates were not as high as desired, but the provision of incentives may have led to improvements in CU participation once CUs were recruited – record use, within-CU diary participation, and diary entries were all high. Interview durations were significantly longer than RP-equivalent durations, but respondents did not report that the time spent on survey activities was unreasonable. Few diarists were placed with online diaries, and there continued to be technical issues with the login process for online diaries, yet overall, diaries were completed by the diarists themselves at high rates. The findings reported here note some minor issues meriting future examination, but, on the whole, the findings are promising as the CE survey moves forward with the redesign plan.

APPENDIX A: CHI

NOTES and ASSUMPTIONS behind Table 6a

The Restricted Production (RP) sample comprises restricted production CEQ Waves 1 & plus CED Week 1 cases that were eligible for survey participation during the POC data collection period.

- For both CEQ and CED, only cases with at least one contact attempt record (CTTYPE=1,2,3) were included for analysis (similar to POC)
- The number of RP cases for the denominator in the computation of the mean statistics for contact attempts: = No. of CEQ Wave 1 (565 cases) + No. of CED Week 1 (468 cases) = 1,033
- CED Week 1 cases:
 - Due to the incorrect allocation of contact attempts to Week 1 and Week 2 diaries in the current Post Phase 2 requirements, we used Unbox CHAI data for this analysis.
 - For consistency with the use of Post Phase 2 data for POC cases, the Post Phase 2 case final disposition, Outcome was used to classify the cases as “Interviews” or “Type As” (instead of the Unbox data final disposition PICK_UP1).

The aggregate counts used to compute the average count per case in Table 6A are in the subsequent Appendix A tables.

Assumptions applied to both POC and RP contact attempts

A1. Only cases with at least 1 contact attempt (cntctyp =1/2/3) are included for this analysis.

A2. A contact attempt for a case is identified as a unique visit date-time CHAI record.

A3. Contact attempts are counted as those attempts made with the CU. This is to avoid penalizing POC where attempts are also made at the member level.

- 1 CHAI record is created when pCHI is launched. The CHAI record has the same date time stamp as the pCHI record(s) generated.

Assumptions applied to POC member-level contact attempt records

A4. Counting member-level attempts in general

- These contact attempts are sourced from the pCHI records.
- Assumptions A2 & A3 imply that if multiple member-level contact attempts are reported on the same date-time stamp[(regardless if the attempt is to the same or multiple CU members), these pCHI attempts are counted as 1 contact attempt for that CU made at that specific visit date-time.

A5. Identifying Visit 1 and 2 dates

- Visit 1 date is defined as the date of the CU roster.
- Visit 2 date is defined as the earliest date of diary pickup date OR the completion of Section 1.

A6. A contact attempt from pCHI is counted as a *midweek attempt* if:

- It has a visit date time occurring between Visit 1 and Visit 2 dates, OR
- if there is no Visit 2 date, then it has a visit date time after Visit 1

Assumptions applied to Restricted Production contact attempt records

A7. CEQ Visit 1 date: this is the visit date when the 1st occurrence of CTTYPER=1,2 for the case occurs

- Count cases with Visit1 date *before* final visit date for the case, OR cases with Visit 1 date *equal* final visit date for the case

A8. CED cases:

- To determine contact attempt records for Week 1, we subset the Unbox CHAI data for records with contact attempt dates occurring prior and up to PICKDTE1 (the date when the final disposition for the Week 1 diary (PICK_UP1) is assigned).
- To determine contact attempt records for Visit 1, in-scope Unbox CHAI records were those with contact attempt dates occurring prior and up to PLCEDAT1 (the date when the final disposition for placement of Week 1 Diary (INSTAT1) is assigned).

Computation of means from aggregate counts for Table 6a

	Aggregate Counts			Interview w	Mean	
	Interview	Type A	ALL		Type A	ALL
No. POC cases	520	520	1,040	520	520	1,040
No. of RP cases	1,017	577	1,594	1,017	577	1,594
CEQ Wave 1	365	200	565			
CEQ Wave 2	344	217	561			
CED Week 1	308	160	468			
Total no. of contact attempts						
POC	2,736	3,624	6,360	5.3	7.0	6.1
RP	3,888	3,180	7,068	3.8	5.5	4.4
CEQ Wave 1	1,465	1,588	3,053			
CEQ Wave 2	1,548	1,277	2,825			
CED Week 1	875	315	1190			
Total no. of in-person attempts						
POC	1,998	2,808	4,806	3.8	5.4	4.6
RP	2,105	1,844	3,949	2.1	3.2	2.5
CEQ Wave 1	1,100	1,066	2,166			
CEQ Wave 2	678	752	1,430			
CED Week 1	327	26	353			
Total no. contacts made with sample unit member						
POC	1,743	1,139	2,882	3.4	2.2	2.8
RP	2,339	981	3,320	2.3	1.7	2.1
CEQ Wave 1	698	398	1,096			
CEQ Wave 2	772	270	1,042			
CED Week 1	869	313	1182			
Total no. attempts to Visit 1						
POC	1,504	712	2,216	2.9	1.4	2.1
RP	1,818	1,795	3,613	1.8	3.1	2.3
CEQ Wave 1	1,392	1,515	2,907			
CEQ Wave 2 not applicable						
CED Week 1	426	280	706			
Total no. midweek (pchai) contact attempts						
POC	56	1	57	0.1	0.0	0.1
RP (n/a)						
CEQ Wave 1						
CEQ Wave 2 n/a						
CED Week 1						
Residual no. contact attempts (after Visit 1 and midweek attempts) to final disposition						
POC	1,129	2,907	4,036	2.2	5.6	3.9

Computation of means from aggregate counts for Table 6a

	Aggregate Counts			Interview w	Mean Type A	ALL
	Interview	Type A	ALL			
RP (=Total no. contact attempts - Visit 1 attempts)	2,070	3,421	6,762	2.0	5.9	4.2
CEQ Wave 1 (after Visit 1)	73	2,109	3,453			
CEQ Wave 2	1,548	1,277	2,825			
CED Week 1 (after Visit 1)	449	35	484			

Table 6a2: POC contact attempts: aggregate counts

	ITYPE_BLS		All
	BLS Good	BLS TypeA	
N (no. unique CTRLNOs)	520	520	1040
Total no. contact attempts	2,736	3,624	6,360
Up to and including Visit1	1,504	712	2,216
Midweek pchai attempts	103	5	108
After Visit 1 & mid-week attempts	1,129	2,907	4,036
Total no. contacts with sample unit member	1,743	1,139	2,882
Up to and including Visit1	826	300	1,126
Midweek pchai attempts	56	1	57
After Visit 1 and mid-week attempts (if any)	861	838	1,699
Total no. attempts by visit	1,998	2,808	4,806
Up to and including Visit1	1,295	579	1,874
Midweek pchai attempts	14	1	15
After Visit 1 & mid-week attempts	689	2,228	2,917

Table 6a3: RP CEQ contact attempts: aggregate counts

	interview	Type A	All
Wave 1			
No. unique NEWIDs	365	200	565
Total no. contact attempts	1,465	1,588	3,053
Total no. of contacts with sample unit member	698	398	1,096
Total no. of contact attempts by visit	1,100	1,066	2,166
Wave 1 total no. attempts through Visit 1	1,392	1,515	2,907
Cases with Visit 1 dates=final attempt date (44 unique NEWIDs)	1,300	1,460	2,760
Cases with Visit 1 dates before final attempt date (521 unique NEWIDs)	92	55	147
Wave 1 no. in-person visit attempts through Visit 1	1,079	1,036	2,115
Cases with Visit 1 dates=final attempt date	1,015	1,002	2,017
Cases with Visit 1 dates before final attempt date	64	34	98
Wave 1 No. contacts with unit member through Visit 1	652	380	1,032
Cases with Visit 1 dates=final attempt date	610	352	962
Cases with Visit 1 dates before final attempt date	42	28	70
Wave 2			
No. unique NEWIDs	344	217	561
Total no. contact attempts	1,548	1,277	2,825
Total no. of contacts with sample unit member	772	270	1,042
Total no. of contact attempts by visit	678	752	1,430
Total sample size (wave 1 & 2)	709	417	1,126

Table 6a4 RP CED Week 1 contact attempts: aggregate counts

	Interview (outcome=201)	Type A	ALL
No. CED Cases (with at least 1 record with CTTYPER=1,2,3)	308	160	468
Total no. contact attempts to Week 1 final disposition (attempt date <=PICKDTE1)	875	315	1190
Total no. contacts with sample unit member	869	313	1182
Total no. in-person attempts	327	26	353
Total no. contact attempts up to Visit 1 (attempt date <=PLCEDTE1)	426	280	706
Total no. of attempts to Week1 diary placement (INSTAT=201)	426	41	467