



# COVID-19 CE DIARY FINAL REPORT: Mode Impacts in Late 2020

Brett McBride & Nikki L. Graf  
Melissa Edwards & Meaghan Smith, Consultants  
*Consumer Expenditure Surveys Program Report Series*

## I. Introduction

Following the outbreak of the COVID-19 pandemic in the U.S. in early 2020, the Census Bureau adopted a number of measures to revise data collection for surveys, including procedures for the Consumer Expenditure (CE) Diary Survey. Under normal diary data collection operations, respondents give diaries with two weeks of recorded expenditures to the interviewer. Prompted by pandemic measures, starting around March 19, 2020, the Census Bureau changed data collection procedures. Census suspended personal visits and instructed interviewers to call to recruit respondents and verbally collect expenditures in two one-week diaries with the aid of receipts and a diary form sent by mail or email. In a second period, starting in June 2020, interviewers provided URLs to English-speaking respondents who usually access the internet multiple times a week to enable them to enter expenditures in an online diary. For those without internet or who preferred not to use an online diary, interviewers used phone transcription to record expenditures. In a third period, starting in July 2020, the period of time covered in this report, interviewers continued to either offer the online instrument for diary completion or collect the expenditures using the phone transcription procedures. Additionally, from July through the end of the year, Census relaxed COVID-19 regulations that suspended personal visits. In areas where health guidelines permitted it and households were receptive, interviewers resumed the placement and collection of paper diaries, using the traditional procedures that were in place prior to March 19, 2020. This report documents how these revised data collection procedures implemented in response to the pandemic affected spending behavior, diary reporting, and data quality in the second half of 2020. An earlier report examined research into these areas for data from the first half of 2020.<sup>1</sup>

## II. Data and Methods

This report covers CE diary data from the July through December 2020 collection months. Data used primarily consisted of post (processed) edited data. This report, examining the resumption of selective in-person visits in “Period 3,” follows the earlier periods (i.e., Baseline, Period 1, and Period 2) covered in the first report.

---

<sup>1</sup> [“COVID-19 CE Diary Preliminary Report: Response and Data Quality Impacts.”](#) McBride, B., and N. Graf. Consumer Expenditure Surveys Program Report Series. June 2021.

- Period 3 – July 1-December 31, 2020. Referral of respondents to online diaries with an alternative of phone transcription. Gradual resumption of in-person visits (e.g., for diary placement and pickup), but generally maximum telephone posture.

For most analyses, diary weeks were subset to completed interviews (INTRVIEW = 1), included in the fully processed Edit and Estimation Systems (EES) data.<sup>2</sup> The counts of completed Consumer Units (CUs) and diaries are displayed below (Table 1).<sup>3 4</sup>

Table 1. Sample of complete EES CUs and diaries by week

	CUs	Diaries with 1+ entries
Both weeks not blank	2,499	4,998
Blank 1st diary week	58	58
Blank 2nd diary week	149	149
Total	2,706	5,205

In total, there were 5,205 complete diaries that contained at least one entry across the two-week reporting period. As the mode by which potential respondents are recruited to participate and complete entries affects the participating sample composition and reporting behavior, we focused on the various modes used during this part of the pandemic.<sup>5</sup> Furthermore, since it is at the diary pickup stage where the reminding to report forgotten expenditures and collection of receipts, two important survey components, occur, we generally characterized diary mode using the mode at point of collection. Diary mode is defined by the three classifications below:

- 1) Traditional – The traditional mode involved in-person diary pickups (presumably with recall of forgotten entries and receipt collection), which in July 2020 selectively resumed in areas where health guidelines permitted it.<sup>6</sup> It then became a diary collection option nationwide in September 2020, before being restricted in specific geographies in December 2020.
- 2) Phone transcription – This involved interviewer transcription of respondent-reported entries by phone (regardless of whether a paper diary was initially placed or downloaded by the respondent); this was the mode of pickup for a plurality of diaries as the Census Bureau was in a “maximum telephone posture” for most of the second half of 2020.

<sup>2</sup> EES encompasses the diary editing involving imputation, allocation, and the assignment of completion status, among other processes.

<sup>3</sup> A CU need not provide diary entries in either week to be classified as a complete in EES data (e.g., if the CU was away during the reporting period, or legitimately had no expenditures). 35 complete CUs with no entries in either week were not included in these counts.

<sup>4</sup> A Consumer Unit consists of families, single persons living alone or sharing a household with others but who are financially independent, or two or more unrelated persons living together who pool their income to make joint expenditure decisions. Complete diaries do not include some diaries that were filled out but had implausibly low or no reported expenditure amounts.

<sup>5</sup> See: Jason M. Ward, Kathryn Anne Edwards, “CPS Nonresponse During the COVID-19 Pandemic: Explanations, Extent, and Effects,” *Labour Economics*, Volume 72, 2021, 102060, ISSN 0927-5371, <https://doi.org/10.1016/j.labeco.2021.102060>.

<sup>6</sup> The Census Bureau regional offices permitted in-person diary pickups in the South and West (excluding Washington state).

3) Online – Unlike the other groups, this referred more to how the respondent recorded entries (i.e., in the online instrument) regardless of the type of pickup interview. A CU had to have made at least one entry in the online diary to be categorized into this group.

At the diary level, the share of diaries collected in each mode is shown below (Table 2).

Table 2. Mode breakout of diary completion

	N (diaries)	Percent
Traditional	1,649	31.7%
Phone transcription	2,212	42.5%
Online	1,344	25.8%
Total	5,205	100%

The table below shows how the mode of diary completion for CUs evolved in late 2020 (Table 3).

Table 3. Mode distribution by month

	N (CUs)	Traditional (N=853)	Phone Trans. (N=1,153)	Online (N=700)
July	357	8.1%	61.6%	30.3%
August	412	16.5%	50.0%	33.5%
September	524	37.0%	40.8%	22.1%
October	503	42.2%	33.6%	24.3%
November	463	43.8%	33.1%	23.1%
December	447	32.9%	42.7%	24.4%
Total	2,706	31.5%	42.6%	25.9%

As Census lifted restrictions on using personal visits for data collection, interviewers transitioned away from the phone transcription mode (61.6 percent of CUs in July down to 42.7 percent in December) and back toward using the traditional in-person mode (8.1 percent in July, 32.9 percent in December). During this time period, the monthly share of CUs using online diaries varied between roughly 22 and 34 percent. Among CUs placed with online diaries, 86 percent went on to provide at least one entry in this mode (10 percent switched to complete the diary using phone transcription, and 4 percent switched to complete by the traditional mode).

The assessments of the effects of revised data collection procedures related to COVID-19 on the Consumer Expenditure Diary Survey are based on the following analyses and methods:<sup>7</sup>

1) Completion Rates. Using Initial Edit System (IES) data, which included non-complete cases, we compared completion rates for diaries placed as online diaries to those placed as paper/non-online diaries during this period.<sup>8</sup> In addition, we examined how often respondents selected the “no expenses” checkbox in online diaries.

<sup>7</sup> See appendix for detailed methods information about regression model-building decisions.

<sup>8</sup> IES reflects the specific diary entries indicated by the interviewer or respondent prior to any allocation or imputation.

2) Demographic Analysis. We examined how the demographic composition of the responding samples varied across modes. We analyzed demographics at the CU level, including CU size, pre-tax CU income, housing tenure, family composition, the region and type of area where the CU resides (urban/rural), and the following characteristics of the respondent: age, gender, race, Hispanic origin, and educational attainment. The demographic analysis used post-processed data, including income imputations.

3) Recalled Expenses and Receipt Usage. For complete diaries, we examined rates of recall by mode. For diaries with at least some recall entries, we identified the share providing receipts. Limited receipt provision, especially for online diaries that were not picked up in person, could indicate problems with interviewer receipt-collection protocols (e.g., transcribing receipts and recalled expenses into a separate paper diary form), compared with other modes.

4) Non-itemization and Allocation Rates. We analyzed respondent lack of adherence to instructions about diary reporting – namely, not itemizing expenses but instead using general descriptors – across modes, making use of information about allocated expenditures. This was analyzed specifically for food expenditures, where the presence of at least one entry containing general descriptors was identified using item code 100090.<sup>9</sup> Additionally, we examined how often diaries contained any entries requiring allocation (due to non-specificity) by mode.

5) Drop-Off in Weekly Reporting Behavior. We analyzed respondent reporting behavior over the two-week period by examining the drop-off in expenditure entries using minimally processed (IES) data.<sup>10</sup> We measured this by calculating the proportion of two-week diary expenditure entries reported in week one (i.e., the number of expenditure entries in week one over the number of expenditure entries reported in both weeks) for each CU that had at least one entry over the two-week period. This measure was compared across collection modes.

6) Expenditure Counts: By Section and By Mode. We used minimally processed (IES) data to examine the number of entries by diary week, broken out by section and by mode, and subset to diary weeks containing one or more entries.<sup>11</sup> For section-level analysis, using zero counts for sections with no expenditures, we calculated the share of diaries containing no entries in each section. We carried out ordinary least squares (OLS) regressions of the number of diary entry counts on mode and control variables to identify the impact of diary mode (in-person, phone transcription, and online) on reporting. For control variables, we incorporated processed data including imputed income as well as interview and respondent characteristics.

7) Item Nonresponse. We examined fields requiring a response, identifying those where the respondent did not know the answer or refused to provide one. We also analyzed the rate of item nonresponse in the cost and item description fields (excluding the Meals Away section) by mode.<sup>12</sup>

---

<sup>9</sup> This code is assigned when the diary contains a vague food entry description, such as “food,” “groceries,” or “junk food.”

<sup>10</sup> Minimally processed data (or IES data) are used in many parts of the report to avoid counting expenditures that were created during the allocation process.

<sup>11</sup> The Consumer Expenditure Diary consists of four sections: “Food and Drinks for Home Consumption” (or “Food for Home”), “Meals, Snacks, and Drinks Away from Home” (or “Meals Away”) which includes delivered or take-out food, “Clothing, Shoes, Jewelry, and Accessories” (or “Clothing”), and “All Other Products, Services, and Expenses” (or “Other”).

<sup>12</sup> In the Meals Away section, respondents do not enter a description beyond providing the name of the vendor.

8) Expenditure Levels. We examined the total reported expenditure amounts at the diary level, using EES data, by mode. We further investigated mode differences by comparing expenditure amounts by diary section. To control for demographic differences between CUs across modes, we estimated OLS regressions. Regression models also incorporated interview characteristics such as indicators for allocation and interviewer-prompted recall.

### III. Findings

#### 1. Completion Rates

To examine differences by mode, we compared completion rates for diaries placed as online diaries to those placed as paper/non-online diaries during the period from July through December of 2020. Information on changes in overall nonresponse and nonresponse category (e.g., noncontact, refusal) is available in the Annual CE Data Quality Profile.<sup>13</sup>

In contrast to other analyses presented in this report, the data used for this analysis included non-complete cases and focused on the mode of placement, rather than the mode at point of collection. Therefore, the online group here includes all diaries initially placed as online diaries, regardless of whether the respondent actually recorded their entries in the online instrument or finished in a different mode.

Diaries placed as online diaries had significantly lower completion rates than those placed as non-online diaries: 66.9 percent of diaries placed as online diaries resulted in completed interviews, compared with 81.5 percent of diaries placed as non-online diaries.<sup>14</sup>

In addition, we expanded the comparison to examine diaries that had some data entered, even though their interview outcome was subsequently reclassified as nonresponse during the editing process because the information provided in the diaries was viewed as incomplete. When these diaries with some data entered were combined with completed interviews, diaries placed as online diaries remained far less likely than those placed as non-online diaries to have any data entered (67.4 percent vs. 82.7 percent).<sup>15</sup>

In online diaries, respondents could indicate whether they had no expenditures in one or both weeks by marking a checkbox, which appeared on the last day of the diary week for diaries that had no expenditures up to that point, if the respondent had already logged in to their online diary at least once. To explore how often respondents used this checkbox, we examined unprocessed data. During the time period analyzed in this report (July through December 2020), only two online diaries (in two separate CUs) had the “no expenses” checkbox marked. One of these two diaries had expenditures subsequently entered and was included in our analysis as a complete diary, while the other was indeed a blank diary week. Additional analysis using login date paradata would be needed to determine the share of respondents who actually viewed the “no expenses” checkbox and had the option of selecting them, since a respondent’s login would need to occur on or after the last day of the diary week for them to potentially see the checkbox.

---

<sup>13</sup> [The Annual CE Data Quality Profile - 2020](#)

<sup>14</sup>  $Z = -13.66, p < 0.0001.$

<sup>15</sup>  $Z = -14.46, p < 0.0001.$

## 2. Demographic Analysis

In this section, we examined how the demographic composition of the responding samples varied across modes. Specifically, we compared CUs who were eligible for and agreed to use the online mode and made at least one entry in the online diary to CUs with traditional and phone collection modes (Table 4). These latter groups included those who were ineligible for the online diary or otherwise declined or were unable to use the online diary. Therefore, the compositional differences across modes reflect the eligibility requirements to some extent, including having adequate internet access and appropriate technology, which can vary by socioeconomic factors and age.

Respondents in CUs completing online diaries were far more likely to be college graduates than those in CUs with traditional and phone collection modes (61.6 percent vs. 37.9 percent and 38.2 percent, respectively).<sup>16</sup> Respondents in the online diary group were also generally younger; 51.9 percent were under age 50, compared with 38.5 percent in the traditional mode and 36.3 percent in the phone transcription mode.<sup>17</sup> The phone transcription group, on the other hand, had the highest share of respondents ages 65 and older (36.7 percent vs. 31.3 percent in the traditional mode and 21.9 percent in the online mode).<sup>18</sup> The shares of female respondents were similar across the traditional and phone transcription groups (55.1 percent and 56.1 percent, respectively), though the phone transcription group had a significantly higher share of female respondents than the online diary group (51.3 percent).<sup>19</sup>

Racial and ethnic composition also varied across modes. Online diary CUs were less likely to have a respondent who reported being of Hispanic origin than those with traditional and phone collection modes (7.3 percent vs. 14.3 percent and 14.1 percent, respectively).<sup>20</sup> It should be noted that non-English-speaking Spanish speakers were not given the online diary option, which factors into these differences. In addition, the traditional and phone transcription groups had higher shares of respondents identifying as Black or African American (11.7 percent and 11.2 percent, respectively) than the online diary group (6.0 percent), while the online group had higher shares of respondents identifying as other races or multiple races (12.4 percent vs. 7.9 percent in the traditional mode and 9.8 percent in the phone transcription mode).<sup>21</sup> <sup>22</sup> Shares of respondents identifying as White were not significantly different across modes.

Income levels differed substantially between online and other modes, with online diary CUs reporting considerably higher pre-tax incomes on average than those in the traditional and phone collection modes. Online diary CUs had a mean pre-tax income of \$114,136, whereas the corresponding means for CUs in the traditional and phone transcription modes were \$77,905 and \$77,477, respectively.<sup>23</sup> Similar income gaps were evident when comparing median pre-tax income (\$90,800 for CUs in the online mode vs. \$54,695 for those in the traditional mode and \$53,164 in the phone transcription mode). Likewise,

---

<sup>16</sup> Test for the equality of proportions of respondents with a bachelor's degree or higher – Online vs. Traditional:  $Z = 9.30$ ,  $p < 0.0001$ ; Online vs. Phone transcription:  $Z = 9.75$ ,  $p < 0.0001$ .

<sup>17</sup> Online vs. Traditional:  $Z = 5.18$ ,  $p < 0.0001$ ; Online vs. Phone transcription:  $Z = 6.48$ ,  $p < 0.0001$ ; Traditional vs. Phone transcription:  $Z = 1.01$ ,  $p = 0.31$ .

<sup>18</sup> Phone transcription vs. Traditional:  $Z = 2.51$ ,  $p = 0.01$ ; Phone transcription vs. Online:  $Z = 6.69$ ,  $p < 0.0001$ ; Traditional vs. Online:  $Z = 4.17$ ,  $p < 0.0001$ .

<sup>19</sup> Traditional vs. Phone transcription:  $Z = 0.45$ ,  $p = 0.65$ ; Phone transcription vs. Online:  $Z = 2.02$ ,  $p = 0.04$ ; Traditional vs. Online:  $Z = 1.50$ ,  $p = 0.13$ .

<sup>20</sup> Online vs. Traditional:  $Z = -4.37$ ,  $p < 0.0001$ ; Online vs. Phone transcription:  $Z = -4.47$ ,  $p < 0.0001$ .

<sup>21</sup> Test for the equality of proportions of respondents identifying as Black or African American – Online vs. Traditional:  $Z = -3.89$ ,  $p < 0.0001$ ; Online vs. Phone transcription:  $Z = -3.74$ ,  $p = 0.0002$ .

<sup>22</sup> Test for the equality of proportions of respondents identifying as other races or multiple races – Online vs. Traditional:  $Z = 3.00$ ,  $p = 0.003$ ; Online vs. Phone transcription:  $Z = 1.77$ ,  $p = 0.08$ ; Traditional vs. Phone transcription:  $Z = -1.51$ ,  $p = 0.13$ .

<sup>23</sup> Comparisons of pre-tax CU income by mode used the mean of imputation iterations.

looking at shares with pre-tax incomes of \$100,000 or more, 44.7 percent of online diary CUs had incomes at this level, compared with 28.6 percent in the traditional mode and 25.4 percent in the phone transcription mode.<sup>24</sup> In addition to having higher incomes, online diary CUs were also more likely to be homeowners than those in the traditional and phone collection modes (73.6 percent vs. 68.7 percent and 65.0 percent, respectively).<sup>25</sup>

Although CU sizes were generally similar across CUs completing online diaries and those in the traditional mode, both tended to be larger than CUs in the phone transcription mode. For example, 21.7 percent of CUs completing online diaries and 20.6 percent of those in the traditional mode had four or more persons, compared with 15.8 percent of CUs in the phone transcription mode.<sup>26</sup>

Regional differences across modes were also apparent. For example, the traditional mode had a higher share of CUs in the South (38.1 percent) than did the phone transcription mode (29.1 percent) or online mode (22.6 percent).<sup>27</sup>

---

<sup>24</sup> Online vs. Traditional:  $Z = 6.59$ ,  $p < 0.0001$ ; Online vs. Phone transcription:  $Z = 8.59$ ,  $p < 0.0001$ ; Traditional vs. Phone transcription:  $Z = 1.60$ ,  $p = 0.11$ .

<sup>25</sup> Online vs. Traditional:  $Z = 2.10$ ,  $p = 0.04$ ; Online vs. Phone transcription:  $Z = 3.82$ ,  $p = 0.0001$ ; Traditional vs. Phone transcription:  $Z = 1.71$ ,  $p = 0.09$ .

<sup>26</sup> Online vs. Traditional:  $Z = 0.52$ ,  $p = 0.60$ ; Online vs. Phone transcription:  $Z = 3.22$ ,  $p = 0.001$ ; Traditional vs. Phone transcription:  $Z = 2.80$ ,  $p = 0.005$ .

<sup>27</sup> Traditional vs. Phone transcription:  $Z = 4.22$ ,  $p < 0.0001$ ; Traditional vs. Online:  $Z = 6.58$ ,  $p < 0.0001$ ; Phone transcription vs. Online:  $Z = 3.10$ ,  $p = 0.002$ .

Table 4. Sample CU characteristics by mode (excludes CUs with no expenditure entries)

	Traditional (n=853)	Phone Trans. (n=1,153)	Online (n=700)
Race of Respondent	%	%	%
White	80.4	79.0	81.6
Black or African American	11.7	11.2	6.0
Other (incl. Asian, multi, other)	7.9	9.8	12.4
Hispanic Origin of Respondent			
Hispanic or Latino	14.3	14.1	7.3
Not Hispanic or Latino	85.7	85.9	92.7
Gender of Respondent			
Female	55.1	56.1	51.3
Male	44.9	43.9	48.7
Age of Respondent			
Under 25	3.6	3.6	3.6
25-34	12.0	11.5	19.0
35-49	22.9	21.2	29.3
50-64	30.2	27.0	26.3
65 and older	31.3	36.7	21.9
Education of Respondent			
Less than high school	10.2	8.9	1.6
High school graduate	23.9	22.5	9.9
Some college or associate's deg.	28.0	30.3	27.0
Bachelor's degree or higher	37.9	38.2	61.6
CU Size			
One person	29.1	35.6	26.4
2 persons	34.9	35.0	37.4
3 persons	15.4	13.6	14.4
4+ persons	20.6	15.8	21.7
CU Composition			
One person	29.1	35.6	26.4
Married couple only	25.0	24.8	28.9
Married couple with children only	21.8	17.8	26.3
One parent with children	4.0	4.1	3.3
Other composition	20.2	17.8	15.1
CU Income Before Tax			
<\$30,000	25.4	26.9	13.0
\$30,000-49,999	21.2	20.6	12.6
\$50,000-99,999	24.7	27.1	29.7
\$100,000+	28.6	25.4	44.7
Mean Income Before Tax	\$77,905	\$77,477	\$114,136
Median Income Before Tax	\$54,695	\$53,164	\$90,800



Table 4 (cont.).

	Traditional (n=853)	Phone Trans. (n=1,153)	Online (n=700)
<b>Housing Tenure</b>			
Non-Homeowner	31.3	35.0	26.4
Owner	68.7	65.0	73.6
<b>Area Type</b>			
Urban	92.4	95.9	93.6
Rural	7.6	4.1	6.4
<b>Region</b>			
Northeast	12.2	15.9	20.3
Midwest	23.2	19.7	23.7
South	38.1	29.1	22.6
West	26.5	35.3	33.4

### 3. Recalled Expenses and Receipt Usage

The number of diary entries is composed of those expenditures respondents reported throughout the week as well as forgotten expenditures that interviewers are instructed to collect (or prompt respondents to enter) through “recall” at the end of the reporting period. In the traditional mode, recalled expenditures are entered into the respondent’s diary at pickup, but for CUs where collection had to occur over the phone, they were collected by the interviewer as part of the phone transcription process. In the case of CUs placed with online diaries, interviewers were instructed to enter recalled expenses in a separate paper diary, especially if the CU decided to discontinue online diary entry. Regardless of mode, interviewers were asked to record whether CUs: 1) provided any entries in the diary, and 2) had any entries collected by recall. In this analysis, a diary with no entries except for those collected by recall was defined as a “total recall” diary, and one that had initial and recalled entries was a “partial recall” diary.

We examined the shares of diaries with total and partial recall by mode, among diaries that contained at least one entry (Table 5). Diaries completed by phone had the highest rates of any recall (36.8 percent), followed by diaries completed in the traditional mode (23.2 percent).<sup>28 29</sup> Only 5.3 percent of online diaries involved any recall.<sup>30</sup> Interviewers may have classified the standard expenses recorded via phone transcription as recalled expenses, which could contribute to the higher rates observed for that mode. Similar patterns emerged by mode for total recall: 16.4 percent of diaries in the phone transcription mode, 10.9 percent of diaries in the traditional mode, and just 1.3 percent of online diaries involved total recall.<sup>31</sup>

Table 5. Percent of diaries with total and partial recall by mode (among diaries with at least one entry)

	N (diaries)	Any Recall	Total Recall	Partial Recall
Traditional	1,649	23.2%	10.9%	12.3%
Phone Transcription	2,212	36.8%	16.4%	20.4%
Online	1,344	5.3%	1.3%	4.0%
Total	5,205	24.3%	10.7%	13.6%

For diaries with at least some recall entries, we identified the share providing receipts, since limited receipt provision, especially for online diaries that were not picked up in person, could indicate problems with interviewer receipt-collection protocols, compared with other modes. However, results showed that these shares were largely similar across modes (Table 6).

<sup>28</sup> Phone transcription vs. Traditional:  $Z = 9.06$ ,  $p < 0.0001$ .

<sup>29</sup> This was noted to be a higher rate of any recall for diaries with phone transcription than was observed when that mode was first introduced from March to May in 2020 (30.6 percent) and in June (25.8 percent); no traditional diaries were placed in those months.

<sup>30</sup> The comparable rate for the initial online diaries in June was 10.3 percent.

<sup>31</sup> Phone transcription vs. Traditional:  $Z = 4.82$ ,  $p < 0.0001$ ; Phone transcription vs. Online:  $Z = 14.14$ ,  $p < 0.0001$ ; Traditional vs. Online:  $Z = 10.59$ ,  $p < 0.0001$ .

Table 6. Percent of recall diaries with receipts by mode (among diaries with at least one entry)

	N (diaries)	With Receipts
Traditional	382	79.6%
Phone Transcription	814	81.8%
Online	71	81.7%
Total	1,267	81.1%

#### 4. Non-itemization and Allocation Rates

A potential indicator of poorer data quality involves diary entries that contain general descriptors. In these situations, the BLS needs a greater level of specificity than the CU provided in order to characterize the CU’s expenditures with sufficient accuracy. One way to identify these situations is through identifying vague descriptions of food purchases, such as “food” or “groceries,” that indicate the CU did not itemize their food purchases item-by-item, as instructed. Rates of any Food for Home section non-itemization, excluding diaries with no entries in that section, are shown by mode in Table 7.

Table 7. Non-itemization rates in the Food for Home section by mode (among diaries with at least one entry in the Food for Home section)

	N (diaries)	Percent
Traditional	1,443	15.4%
Phone transcription	1,768	24.0%
Online	1,221	46.4%
Total	4,432	27.4%

Table 7 indicates that only 27.4 percent of diaries had an entry in the Food for Home section that was not itemized. In other words, most diaries did not have any of the general food descriptions that would require those entries being allocated to more specific food items. This low rate of diaries with non-itemized descriptions did vary significantly by mode, however.<sup>32</sup> Diaries that were completed online were more likely to contain a non-itemized entry, and this was the least prevalent among diaries completed in the traditional mode. Looking at non-itemization in a different way, there were 19.9 percent of diaries completed online that contained nothing but non-itemized entries in that section (contrasted with a little under six percent for the traditional mode). These findings suggest that interviewers may need to provide a greater emphasis on the itemization instructions when providing online diaries in order to maximize the accuracy of food data without resorting to allocation edits.

Allocation is an indicator that the BLS needed to enhance respondent-provided data to make it fit for use.<sup>33</sup> Between food non-itemization (about a quarter of diaries) and entry allocation (over three-quarters of diaries) there was an overlap, but many diaries with allocation did not have any food non-itemization (55.5 percent). We examined the percent of diaries containing any allocation by mode (Table 8).

<sup>32</sup> ANOVA; Significant difference  $F=182.41$ , 2 df,  $Pr>F < 0.0001$ . Student-Newman-Keuls (SNK) test; online is significantly higher than phone transcription which is higher than the traditional mode ( $p<0.05$ ).

<sup>33</sup> Specifically, allocation makes amounts not reported according to the CE item structure fit into the CE item structure (e.g., “bread” being allocated to a specific bread type).

Table 8. Diaries with any allocation by mode (among diaries with at least one entry)

	N (diaries)	Percent
Traditional	1,649	79.4%
Phone transcription	2,212	72.6%
Online	1,344	88.1%
Total	5,205	78.7%

The findings indicate that allocation is a fairly common edit needed for diaries but that it occurs somewhat more frequently when diaries are filled out online.<sup>34</sup> We examine allocation further in regressions predicting weekly diary counts and reported expenditure amounts.

### 5. Drop-Off in Weekly Reporting Behavior

Respondents may experience fatigue from the diary reporting process and report fewer expenditure entries over time. This “drop-off” was measured for each collection mode by calculating the proportion of two-week diary expenditure entries reported in week one (i.e., the number of expenditure entries in week one divided by the number of expenditure entries reported in both weeks) for each CU that had at least one entry over the two-week period. Table 9 shows both the average entries per week, and the computed average within-CU drop-off.

Table 9. Weekly entries and within-CU drop-off in weekly reporting (among CUs with at least one entry) by collection mode

	Traditional (n=853)	Phone Trans. (n=1,153)	Online (n=700)
Week 1 mean entries	30.9	21.2	23.8
Week 2 mean entries	28.9	18.3	21.1
Mean entries (total)	59.8	39.5	44.9
Mean (CU-level) drop-off	52.7%	53.9%	56.1%

Since proportions greater than 50 percent reflect a drop-off in entries across the 2 weeks, Table 9 suggests the potential of a reduction in data quality regardless of collection mode. The differences between collection modes were significant with drop-off higher for the online mode.<sup>35</sup> However, the observed drop-off in the counts of entries in the second half of 2020 was in line with levels measured earlier in the year.<sup>36</sup> Calculations in Table 9 include CUs not providing any entries in one of the two diary weeks. Accordingly, the findings in Table 9 can be attributed to a greater proportion of CUs with online completion not having entries in the second week (of online CUs missing entries in one week, 93 percent missed them in the second week), compared to lower corresponding rates for CUs with traditional and phone collection modes (65 percent and 64 percent, respectively).<sup>37</sup>

<sup>34</sup> ANOVA; Significant difference  $F=61.97$ , 2 df,  $Pr>F < 0.0001$ . Student-Newman-Keuls (SNK) test; online is significantly higher than traditional, which is higher than the phone transcription mode ( $p<0.05$ ).

<sup>35</sup> ANOVA; Significant difference  $F=4.79$ , 2 df,  $Pr>F = 0.0084$ . Student-Newman-Keuls (SNK) test; online is significantly higher than other modes ( $p<0.05$ ).

<sup>36</sup> Drop-off was around 53 percent in the first half of 2020 when calculated to include CUs not providing entries in one of the two diary weeks.

<sup>37</sup> After accounting for these differences (representing a small share of all online diaries), there is no longer a significant difference in drop-off across collection modes (ANOVA;  $F=1.10$ , 2 df,  $Pr>F = 0.3321$ ).

## 6. Expenditure Counts: By Section and By Mode

For the expenditure count analysis we examined counts at the diary level using minimally processed data. One indicator for potential differences in data quality would be lower expenditure reports for certain modes. We examined the counts of entries by mode for diaries containing at least one entry (Table 10).

Table 10. Count of weekly Diary entries by mode

	N (diaries)	Mean	Median	Min	Max
Traditional	1,649	30.9	26	1	173
Phone Transcription	2,212	20.6	14	1	147
Online	1,344	23.4	16	1	232
Total	5,205	24.6	17	1	232

The median number of diarist entries ranged from 14 (phone transcription mode) to 26 (traditional mode); overall, diarists provided a median of 17 entries. These differences across modes were statistically significant.<sup>38</sup> Examining diary section-level counts across modes provides a sense of what may account for these differences (Table 11).

Table 11. Count of weekly Diary section entries by mode (including sections with no entries)

	N (diaries)	Food for Home		Meals Away		Clothing		Other	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
Traditional	1,649	19.1	15	2.6	1	1.0	0	8.2	6
Phone Trans.	2,212	12.1	6	1.7	1	0.6	0	6.2	4
Online	1,344	12.6	5	2.7	2	0.9	0	7.3	5
Total	5,205	14.4	8	2.3	1	0.8	0	7.1	5

We can see from Table 11 that the preponderance of the differences between modes was due to differences in expenditure reports in the Food for Home section. These significant differences within the Food for Home section point to evidence of underreporting in diaries completed using non-traditional modes.<sup>39</sup> This could also be a result of the finding in Table 7 that non-itemization rates in that section were higher in the non-traditional modes. CUs completed in those modes would have reported a larger number of food items had they been willing (or able) to itemize as frequently as did CUs in the traditional mode. Table 12 shows the percent of diaries containing no entries in a section by mode.

<sup>38</sup> ANOVA; Significant difference  $F=99.68$ , 2 df,  $Pr>F < 0.0001$ . Student-Newman-Keuls (SNK) test; traditional is significantly higher than online which is higher than phone transcription mode ( $p<0.05$ ).

<sup>39</sup> ANOVA; Significant difference  $F=94.73$ , 2 df,  $Pr>F < 0.0001$ . Student-Newman-Keuls (SNK) test; traditional is significantly higher than online and phone transcription modes ( $p<0.05$ ).

Table 12. Percent of weekly diary sections containing no entries by mode (among diaries with at least one entry)

	N (diaries)	Food for Home Percent	Meals Away Percent	Clothing Percent	Other Percent
Traditional	1,649	12.5%	37.9%	68.7%	7.9%
Phone Trans.	2,212	20.1%	46.6%	78.1%	7.0%
Online	1,344	9.2%	27.8%	65.0%	8.9%
Total	5,205	14.9%	37.9%	71.7%	7.9%

Sections with no entries, or “blank sections” occurred for only 8 percent of diaries in the Other section. In contrast, most diaries – 72 percent – had a blank section for Clothing. Notable variation in blank sections occurred for Meals Away, ranging from 28 percent for online to roughly 47 percent for phone diaries.

To determine the association of mode with the entries recorded, accounting for the impact of differences in the demographics of the responding sample, we conducted a regression analysis. The regression (including non-significant variables) involved mode, demographics, interview characteristics, and several control variables. The regression excluded diaries for which there were no entries (for which mode of completion was not expected to be a factor). It was based on minimally processed (e.g., pre-allocation/imputation) count data, representing 2,706 CUs (providing 5,205 weekly diaries).<sup>40</sup> The regression (also showing a reduced model that excluded non-significant variables) is shown below (Table 13).

---

<sup>40</sup> See appendix for more details about the model methods. Data on CU characteristics were fully processed data and included imputed income values.

Table 13. Expenditure counts regression model findings (n=5,205)

	Full Model		Reduced Model	
	Coefficient	SE	Coefficient	SE
Intercept	3.42***	(0.075)	3.46***	(0.062)
Phone (v. Traditional)	-0.43***	(0.054)	-0.42***	(0.053)
Online (v. Traditional)	-0.54***	(0.054)	-0.54***	(0.053)
Recall	-0.22***	(0.033)	-0.22***	(0.033)
No allocation	-1.21***	(0.059)	-1.22***	(0.059)
Race: Black	-0.13**	(0.048)	-0.16***	(0.047)
Race: Other	-0.21***	(0.046)	-0.23***	(0.046)
Ethnicity: Hispanic	0.07	(0.044)	-	-
Male respondent	-0.00	(0.027)	-	-
Age: Under 35	-0.18***	(0.046)	-0.18***	(0.046)
Age: 35-49	-0.06	(0.044)	-0.06	(0.043)
Age: 50-64	0.00	(0.037)	0.00	(0.037)
Education: HS or less	-0.20***	(0.059)	-0.19***	(0.058)
Education: Some college	-0.16**	(0.060)	-0.18**	(0.060)
CU Size: 2 members	0.31***	(0.039)	0.28***	(0.036)
CU Size: 3 members	0.28***	(0.077)	0.32***	(0.047)
CU Size: 4+ members	0.45***	(0.076)	0.49***	(0.047)
CU married couple with children	0.13	(0.072)	-	-
CU one parent with children	-0.16	(0.089)	-	-
CU other composition	-0.09	(0.053)	-	-
CU pre-tax income: <\$30K	-0.13**	(0.048)	-0.16***	(0.047)
CU pre-tax income: \$30K-<\$50K	-0.11*	(0.045)	-0.13**	(0.044)
CU pre-tax income: \$50K-<\$100K	-0.07	(0.038)	-0.09*	(0.037)
Housing tenure: Not homeowner	-0.08*	(0.034)	-0.09**	(0.033)
Rural	-0.03	(0.059)	-	-
Region: NE	-0.11*	(0.042)	-0.11**	(0.042)
Region: MW	-0.10*	(0.038)	-0.10*	(0.038)
Region: SO	-0.22***	(0.036)	-0.21***	(0.036)
Month: August	-0.08	(0.050)	-	-
Month: September	-0.04	(0.049)	-	-
Month: October	0.03	(0.050)	-	-
Month: November	-0.00	(0.050)	-	-
Month: December	0.07	(0.050)	-	-
Inter: Phone-No allocation	0.31***	(0.075)	0.31***	(0.075)
Inter: Online-No allocation	0.70***	(0.101)	0.70***	(0.101)
Inter: Phone-<HS education	-0.08	(0.075)	-0.09	(0.075)
Inter: Phone-Some college	-0.00	(0.078)	-0.01	(0.078)
Inter: Online-<HS education	-0.24*	(0.102)	-0.27**	(0.102)
Inter: Online-Some college	-0.02	(0.085)	-0.03	(0.085)
Adjusted Model R <sup>2</sup>		0.264		0.259

\*p<.05; \*\*p<.01; \*\*\*p<0.001

Based on the regression, it is evident that non-traditional data collection modes were associated with lower weekly entry counts – diaries completed by phone having roughly 35 percent fewer entries and those completed online having roughly 42 percent fewer – relative to traditional in-person diary collection, holding all other factors constant. Most of the CU characteristics traditionally associated with diary entries were found to have expected associations in our model, with CU size, respondent educational attainment, and income, all positively associated with entries (after accounting for differing reference groups used for different characteristics). Respondents who did not own their homes were found to have 7 percent fewer entries than those who did. White respondents and those older than age 64 were also found to have higher entry counts (but for age, this was higher only relative to those in the under 35 age category). Neither sex, Hispanic ethnicity, nor urban/rural status was significant. Although variables representing CU composition were included in the full model, these involved a complicated interaction with the CU size variables such that single CUs and married-only CUs served as reference groups. These variables were of only borderline significance, despite suggesting those married with children tended to report more relative to single-member CUs. It was interesting to note geographic variation, with CUs in the west providing significantly more entries relative to those in other regions.

Controlling for these demographic characteristics, a CU in which no allocation occurred was strongly associated with fewer entries recorded. This is not surprising if we consider the frequency of entries needing allocation (see Table 8), such that diaries having more entries would have a higher likelihood of having one of those entries allocated. Thus, allocation is not necessarily an indicator of poorer data quality in the diary. Another element of the diary-keeping process is the collection of entries not recorded during the diary-keeping period through interviewer-prompted recall. As described earlier, this was more prevalent for diaries completed by phone or by the traditional mode (see Table 5). The model indicated that fewer entries were recorded, controlling for other factors, in diaries with recall. Unlike allocation, here it is likely that the need for recall is characteristic of CUs with weaker adherence with the diary keeping task. Despite including expenditures that the respondent was prompted to remember, these diaries were associated with approximately 20 percent fewer entries in total. Based on findings of a shift from phone to in-person diary completion over time (see Table 3) it was of interest whether entries varied by month, controlling for mode. Here, although it might be expected December entries would be higher with the presence of holiday shopping, this was not found to be significant, nor were any of the other months relative to July. For these models, allocation was examined in place of itemization for food for home expenses. This was due largely to the improved model fit when including the (non-) allocation variable. Diaries with allocation despite full itemization were more likely to be diaries with traditional pickup, which suggested that interviewers (in this mode) were able to prompt successful itemization of grocery expenses, while allocation still remained necessary in other areas of the diary. Unsurprisingly, models (not shown) including food non-itemization in place of allocation, found non-itemization to be associated with significantly fewer entries, even controlling for mode.

A number of interactions were examined. Of these, the interaction between mode and allocation stood out for its impact on entries. Diaries completed online without allocation had significantly more entries versus traditional diaries without allocation. It is unclear why this may be, although it could be an indicator of conscientious reporters in this group. The other interaction between mode and education level sheds light into how the mode used affected reporting. Less-educated respondents completing online diaries provided significantly fewer entries than traditional diaries with educated respondents. This was unique to the online interaction as the coefficient for less-educated phone respondents did not obtain significance. If we assume education levels to be strongly associated with technology use, this finding is intuitive.



## 7. Item Nonresponse

Aside from analyzing the number of entries respondents provided, item nonresponse gives an indication of how comprehensively a diary was completed. For some fields in the online diary an entry was required, but due to differences in what is collected by mode, we focused on missing rates for cost and item description. Item nonresponse rates for cost are shown by mode (Table 14).<sup>41</sup>

Table 14. Item nonresponse in the cost field by mode (excluding blank diary weeks)

	N (entries)	Item Nonresponse
Traditional	51,011	1.23%
Phone Transcription	45,560	2.32%
Online	31,444	0.24%
Total	128,015	1.37%

We can see that cost amounts were missing at various rates by mode.<sup>42</sup> Of interest, there were higher rates of missing cost data in diaries involving interviewer transcription than in paper diaries respondents filled out themselves. Diaries completed online had the lowest rates of missing cost data. Turning to item description, the Meals Away section does not involve respondents entering a description beyond providing the name of the vendor, and so this section was excluded from calculations (along with blank diary weeks) (Table 15).

Table 15. Item nonresponse in the item description field by mode (excluding blank diary weeks)

	N (entries)	Item Nonresponse
Traditional	46,736	0.22%
Phone transcription	41,694	0.46%
Online	27,872	0.66%
Total	116,302	0.41%

For item description there were also significant differences by mode.<sup>43</sup> Here, however, online diaries had the highest levels of missing data and traditional diaries the lowest. Since phone transcription began in the early period of the pandemic (from March 19th onward) we can note that missing data rates continued to climb among diaries completed in this mode.<sup>44</sup>

---

<sup>41</sup> In online diaries, respondents must fill out the business name, expenditure category and date fields, as well as one of the other fields. As a result, cost and description do not both need to be provided in order for the entry to be saved.

<sup>42</sup> ANOVA; Significant difference  $F=309.99$ , 2 df,  $Pr>F < 0.0001$ . Student-Newman-Keuls (SNK) test; phone transcription is significantly higher than traditional which is higher than the online mode ( $p<0.05$ ).

<sup>43</sup> ANOVA; Significant difference  $F=42.55$ , 2 df,  $Pr>F < 0.0001$ . Student-Newman-Keuls (SNK) test; online is significantly higher than phone transcription which is higher than the traditional mode ( $p<0.05$ ).

<sup>44</sup> Earlier in 2020, there were 1.6 percent of entries missing cost in the phone transcription mode compared to 2.3 percent in the second half of 2020. There were 0.27 percent of phone-transcribed entries missing item description in the earlier period compared to 0.46 percent in the period covered by this report.

## 8. Expenditure Levels

We examined the total reported expenditure amounts at the diary level, using processed (EES) data, by mode. Table 16 shows that CUs completing online diaries recorded a much higher average total amount of expenditures each week (\$1,059.01) than did those in the traditional and phone transcription modes (\$791.32 and \$737.14, respectively).<sup>45</sup> The difference in the mean amounts between the traditional and phone transcription modes was not statistically significant.<sup>46</sup> Similar patterns were evident for median amounts: \$540.81 for online diaries, compared with \$377.98 for those in the traditional mode and \$340.00 for the phone transcription mode. Therefore, despite online diaries having lower entry counts than those in the traditional mode, they had far higher expenditure totals recorded. And although phone transcription diaries also had lower entry counts than those in the traditional mode, the total expenditure amounts recorded in those modes were similar.

Table 16. Expenditure totals by mode

	N (diaries)	Mean	Median
Traditional	1,648	\$791.32	\$377.98
Phone Transcription	2,211	\$737.14	\$340.00
Online	1,343	\$1,059.01	\$540.81
Total	5,202	\$837.40	\$394.95

Examining expenditure amounts by diary section shows that this pattern of online diary CUs reporting higher amounts than traditional and phone transcription CUs holds across all sections, except for the Clothing section (Table 17). In the Clothing section, diaries completed online and in the traditional mode had similar mean expenditure amounts, and both were higher than the phone transcription mode.<sup>47</sup> In addition, in the Food for Home and Meals Away sections, diaries completed in the traditional mode had higher mean expenditure amounts than the phone transcription mode.<sup>48</sup>

Table 17. Expenditure totals by section and mode

	N (diaries)	Food for Home		Meals Away		Clothing		Other	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
Traditional	1,648	\$114.39	\$83.79	\$48.83	\$20.00	\$36.55	\$0.00	\$591.54	\$182.74
Phone Trans.	2,211	\$94.71	\$61.60	\$35.83	\$7.40	\$25.47	\$0.00	\$581.13	\$183.50
Online	1,343	\$154.20	\$115.29	\$59.65	\$36.27	\$38.48	\$0.00	\$806.68	\$272.87
Total	5,202	\$116.30	\$82.19	\$46.10	\$18.98	\$32.34	\$0.00	\$642.66	\$206.47

Given that the demographic composition of responding CUs varied substantially across modes, and that many of these demographic characteristics are strongly associated with expenditure levels, it is essential to account for these differences when examining the relationship between mode and expenditure amounts. Therefore, we estimated regression models predicting the log of total expenditure amounts at the diary

<sup>45</sup> Online vs. Traditional:  $t = -4.26$ ,  $p < 0.0001$ ; Online vs. Phone transcription:  $t = -5.42$ ,  $p < 0.0001$ .

<sup>46</sup> Traditional vs. Phone transcription:  $t = 1.26$ ,  $p = 0.21$ .

<sup>47</sup> Online vs. Traditional:  $t = -0.47$ ,  $p = 0.64$ ; Online vs. Phone transcription:  $t = -3.78$ ,  $p = 0.0002$ ; Traditional vs. Phone transcription:  $t = 3.03$ ,  $p = 0.0025$ .

<sup>48</sup> In Food for Home:  $t = 5.06$ ,  $p < 0.0001$ ; Meals Away:  $t = 5.76$ ,  $p < 0.0001$ .

level. Independent variables included mode, demographics, interview characteristics, and control variables, as examined in the entry counts regression models. The regressions were based on the 5,202 weekly diaries that included expenditure amount information (Table 18).<sup>49</sup>

---

<sup>49</sup> The sample size for the regression predicting expenditure totals (n=5,202) differed from that for the regression predicting entry counts (n=5,205) because there were three diaries that had expenditure entries but provided no information on the dollar amounts associated with those expenditures.

Table 18. Regression models predicting log of total expenditure amounts (n=5,202)

	<b>Model w/ Mode and Demographics</b>		<b>Full Model (incl. Allocation and Recall)</b>		<b>Reduced Model</b>	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
Intercept	6.279***	(0.088)	6.446***	(0.085)	6.455***	(0.057)
Phone (v. Traditional)	-0.096*	(0.041)	-0.028	(0.040)	-0.021	(0.038)
Online (v. Traditional)	0.126**	(0.047)	0.062	(0.046)	0.056	(0.044)
Recall	-	-	-0.068	(0.040)	-0.068	(0.039)
No allocation	-	-	-0.887***	(0.040)	-0.898***	(0.040)
Race: Black	-0.231***	(0.060)	-0.170**	(0.057)	-0.183**	(0.055)
Race: Other	-0.119*	(0.058)	-0.105	(0.056)	-0.120*	(0.054)
Ethnicity: Hispanic	0.071	(0.055)	0.058	(0.053)	-	-
Male respondent	0.007	(0.034)	0.028	(0.033)	-	-
Age: Under 35	-0.176**	(0.058)	-0.122*	(0.055)	-	-
Age: 35-49	-0.090	(0.056)	-0.051	(0.053)	-	-
Age: 50-64	-0.058	(0.047)	-0.038	(0.045)	-	-
Education: HS or less	-0.363***	(0.046)	-0.301***	(0.044)	-0.297***	(0.043)
Education: Some college	-0.293***	(0.042)	-0.264***	(0.040)	-0.268***	(0.040)
CU Size: 2 members	0.343***	(0.049)	0.271***	(0.047)	0.274***	(0.047)
CU Size: 3 members	0.352***	(0.096)	0.250**	(0.092)	0.243**	(0.091)
CU Size: 4+ members	0.700***	(0.096)	0.560***	(0.092)	0.552***	(0.090)
CU married couple with children only	-0.024	(0.090)	0.012	(0.086)	0.001	(0.085)
CU one parent with children	-0.047	(0.112)	0.047	(0.107)	0.023	(0.104)
CU other composition	-0.226***	(0.066)	-0.183**	(0.063)	-0.191**	(0.063)
CU pre-tax income: <\$30K	-0.608***	(0.060)	-0.566***	(0.058)	-0.558***	(0.055)
CU pre-tax income: \$30K-<\$50K	-0.466***	(0.056)	-0.433***	(0.054)	-0.430***	(0.052)
CU pre-tax income: \$50K-<\$100K	-0.251***	(0.047)	-0.236***	(0.045)	-0.238***	(0.044)
Housing tenure: Not homeowner	-0.212***	(0.042)	-0.210***	(0.040)	-0.232***	(0.037)
Rural	-0.159*	(0.074)	-0.099	(0.071)	-	-
Region: NE	0.128*	(0.052)	0.086	(0.050)	-	-
Region: MW	0.047	(0.048)	0.030	(0.046)	-	-
Region: SO	-0.025	(0.044)	-0.006	(0.043)	-	-
Month: August	-0.004	(0.063)	-0.016	(0.060)	-	-
Month: September	-0.016	(0.061)	-0.006	(0.058)	-	-
Month: October	0.003	(0.062)	-0.016	(0.059)	-	-
Month: November	0.052	(0.063)	0.030	(0.060)	-	-
Month: December	0.057	(0.063)	0.040	(0.060)	-	-
Adjusted R <sup>2</sup>	0.161		0.233		0.233	

\*p<.05; \*\*p<.01; \*\*\*p<0.001

First, we estimated an initial model with predictor variables that included mode, demographics, and month. In this model, the online mode was associated with higher expenditure totals and the phone transcription mode with lower expenditures totals, both relative to the traditional mode. Specifically,

diaries completed online had 13 percent higher expenditure totals recorded than diaries in the traditional mode, whereas diaries completed by phone reported 9 percent lower expenditure totals than those in the traditional mode, after controlling for demographic characteristics of the respondent and CU.

Next, we added indicators for allocation and interviewer-prompted recall to the model. Having no allocation was associated with significantly lower expenditure totals.<sup>50</sup> Recall, however, was not statistically significant. Once a combination of these factors (particularly, allocation) and demographics were accounted for, the associations between mode and expenditure totals were no longer statistically significant.

In alternative models (not shown) that included food non-itemization in place of allocation, we found that non-itemization was also highly significant (associated with higher expenditure totals). After controlling for non-itemization, the association between online mode and expenditure totals was reduced to non-significance in these models as well. However, the association between phone transcription and expenditure totals remained significant in these alternative models.

In the same manner as the prior regression models predicting entry counts, we estimated a reduced model that removed sets of demographic and control variables that did not reach statistical significance in the full model: Hispanic ethnicity, gender, age, type of area (urban/rural), region, and survey month.<sup>51</sup> While mode was not statistically significant in the full model, it is still included in the reduced model presented here because it is the central variable of interest.

As expected, income, education, housing tenure, and CU size were strongly associated with reported expenditure totals. The results also showed that CUs with respondents identifying as Black or other races reported lower expenditure totals than those with respondents identifying as White, net of other factors. In addition, multi-person CUs with other types of composition – which includes unmarried partners, married couples with additional people in their CU other than their own children, and all other compositions besides married couple only, married with children only, and single parents – had lower expenditure totals than single-person CUs and those consisting only of married couples (these latter two categories were both omitted from the model and together served as the reference group for the family composition variables).

Unlike the model predicting entry counts, neither the interactions between mode and allocation nor between mode and education were significant in the full model.<sup>52</sup>

#### **IV. Conclusion**

In this report we followed up on research into how the COVID-19 pandemic affected the Consumer Expenditure Diary Survey and focused on the second half of 2020. We examined whether the prior report

---

<sup>50</sup> In bivariate analysis (not shown), this association was also evident, with diaries with no allocation having lower average expenditure totals than those with allocation.

<sup>51</sup> Although the coefficient for the age under 35 group is statistically significant in the full model (where the reference category is 65 years and older), the set of age group variables were not jointly statistically significant in the full model ( $F(3, 5169) = 1.67, p = 0.17$ ) and therefore were not included in the reduced model.

<sup>52</sup> F-test for the interaction between mode and allocation (i.e., model without interaction terms vs. model with interaction terms):  $F(2, 5167) = 1.87, p = 0.15$ . F-test for the interaction between mode and education:  $F(4, 5165) = 1.52, p = 0.19$ .

findings for the first half of 2020 – of large reductions in diary entries and increased non-itemization but more stable patterns for other data quality measures (e.g., missing data) – persisted into the second half of 2020, and we expanded the analysis to consider additional measures. In this period, we found that the growing use of online diaries, in addition to phone transcription and traditional in-person diary collection, had varying implications for diary reporting and data quality in the second half of 2020.

A distinctive feature of the survey in the early COVID-19 time period was its use of phone transcription to collect expenses, with that mode making up a plurality of CU completions. This was less a choice than a contingency, necessitated by health guidelines discouraging in-person interactions. Demographically, these CUs' diaries came from those not traditionally associated with high levels of spending (e.g., CUs with respondents ages 65 and older and having few CU members). Yet even controlling for these and interview characteristics (e.g., the presence of recall or allocation), they tended to have lower diary entry rates and reported amounts no higher than diaries collected in other modes. Despite these CUs' diaries being most likely to have recalled expenses – at roughly 37 percent with 16 percent having all of their expenses recalled – and the recall process often being accompanied by receipts, overall, diaries with phone transcription had relatively high rates of non-itemized entries in the Food for Home section and elevated missing cost information. These findings suggest that the use of phone transcription in this period, with its potential absence of visual diary cues, likely led to underreporting of expenditures.

The second most frequently completed mode for diary was in-person, the traditional pre-COVID-19 method of diary collection, which occurred at an increasing rate during the second half of 2020. The CU characteristics of those completing diaries in person were similar to those completing diaries by phone transcription as far as their education and income levels, as well as race and ethnicity. The higher rate of weekly diary entries in this group may have been aligned with their lower rates of non-itemization of food expenses. This suggests that interviewer presence is a beneficial feature in identifying and rectifying non-itemized entries during diary pickup.

Diaries placed online were found to be completed at lower rates than diaries not placed online. It is unclear why diaries placed online tended to be completed at lower rates. Additionally, CUs with online diaries were more likely to not provide entries in the second week than in the first week. This first-week only pattern was higher than in CUs not using online diaries and may be a sign of technical difficulty prompting them to discontinue participating with online diaries. Online diaries used in this period were more likely to feature general food descriptions – with a fifth of them only containing such entries in that section. Online diaries were found to have significantly fewer entries after controlling for their demographic characteristics (e.g., younger, with higher incomes) and interview characteristics, yet did not have lower expenditure totals. These findings suggest these CUs may need more training on the need to itemize expenditures, assuming they are willing to do so. Given the findings of the entries model of fewer online entries for respondents with a high school education or less, it might be beneficial to target more technologically savvy CUs for these diaries. In this time period, there were some CUs placed with online diaries – 14 percent – that switched to report using other modes.

Now that the online diary has been incorporated into regular survey data collection for the CE Diary Survey, it is worth further examining whether the patterns in the second half of 2020 (e.g., higher rates of non-compliance with instructions to itemize food expenditures, lower entry counts, and low rates of recall relative to other modes) will persist. Similarly, now that the CE Survey Program regards phone transcription as a last resort and not part of a “maximum telephone posture,” it would be beneficial to investigate current data quality indicators – including item nonresponse, a drop-off in entries, use of recall and receipts – and how they associate with diary mode.

## Methods Appendix

We used unweighted data for this report, as in most analyses they intended to refer only to the universe of complete diaries (or in the analysis of completion rates, to placed diaries), and not to the larger U.S. population.

To carry out the analysis, we used SAS and Stata software in addition to common computer and calculator software. We employed a mix of tests for significance depending on the analysis. For completion rates, demographic analysis, and recall analysis, we used a test of the equality of proportions for most comparisons. For mean comparisons for income and expenditure amounts, we used t tests. Reported p-values refer to two-tailed tests. For other analysis, we used ANOVA tests to determine if there were differences between mode samples, and then often a Student-Newman-Keuls test. The Student-Newman-Keuls test is a stepwise multiple comparisons procedure used to identify specific sample means that are significantly different from each other.

Regression modeling revealed the dependent variables (counts, expenditure totals) led to violation of OLS model assumptions of normally distributed residuals with constant variance. To deal with this, natural log transformations of the dependent variable were carried out (this was further identified as the best transformation using Box-Cox methods). Although the counts regression model still had heteroscedasticity following the transformation it did help improve the distribution of model residuals. The caveat is that the standard errors of estimates may be underestimated due to residuals not being independent, yet this should not affect the significance of the mode findings. The models did not use weights as the findings were seeking to identify the impact of revised data collection procedures related to COVID-19 on respondents' diary-keeping practices and not trying to make inferences to the larger U.S. population.

Regression independent variables were determined in order to accurately control for demographic characteristics that, along with mode of diary completion, could impact the dependent variables. We based many of these on the respondent's characteristics (and not those of the reference person). Although we do not capture which CU member(s) fills out the diary (the respondent is the person who participates in the pickup interview), for most cases this is thought to be the same person; the person participating in the pickup interview was expected to be the relevant person for studying the impact of diary completion. Income variables were based on CE publication table categorizations, and specific ranges (which could be more comparable across models than quartiles) were applied, using the mean of imputation iterations, to the regression variables. This also avoided the problem of carrying out log transformations of zero or negative values. The non-itemization variable was taken from item code 100090 to identify the presence of one or more grocery-type responses in the Food for Home category. We used the variable 'DEG\_URBN' to define urbanicity. Allocation data was coded based on the presence of a post-processing allocation indicator in any of the diary sections. Tenure involved use of the variable 'TENURECU' (and not 'CUTENURE') to define homeownership. We did not use a three-category specification for tenure because for the model predicting log of total expenditures, owners without mortgages were not significantly different from owners with mortgages (reference group). For the model predicting log of count of entries, the coefficient indicated owners without mortgages had more entries than those with mortgages, but it was of borderline significance. Renters were significantly different from owners with mortgages only in their expenditure totals and not their count of diary entries. We included CU composition variables in conjunction with the CU size variables based on the assumption that spending decisions might be different if there were multiple adults/people with spending-decision authority. Such

situations could be associated with misreporting if proxy reporting does not fully capture the expenditures made by a CU.