Testing the Feasibility of Integrating Outlets into the CE Diary

Erica Yu and Brandon Kopp

CE Methods Symposium
July 14, 2015

All views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the U.S. Bureau of Labor Statistics
Outline

- Background: CE, TPOPS, and CPI
- Research questions
- Study methods
- Findings: CE data quality and burden
- Findings: Outlet data quality
- Summary
Background: CE, TPOPS, and CPI

CE
Consumer Expenditure Survey

CPI
Consumer Price Index

TPOPS
Telephone Point-of-Purchase Survey

Other Sources
Background: Telephone Point-of-Purchase Survey

- Primary source for outlet frame for CPI
- Interview is organized around expenditure categories
  - Expenditures collected for all in household
  - Categories are similar to those used in CE
  - Reference periods can be different to those used in CE; based on purchase frequency
Research Questions

If outlet questions are integrated into the CE Diary...

- What is the impact on data quality?
- What is the impact on respondent burden?

Answer: “It depends.”
Two Ways of Integrating Outlets in the CE Diary

<table>
<thead>
<tr>
<th>Item-Based</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Shirt</td>
</tr>
<tr>
<td>a. July 13th</td>
</tr>
<tr>
<td>b. $49.99</td>
</tr>
<tr>
<td>c. Macy's</td>
</tr>
<tr>
<td>d. Age 16+</td>
</tr>
<tr>
<td>e. Female</td>
</tr>
<tr>
<td>2. Blender</td>
</tr>
<tr>
<td>a. July 13th</td>
</tr>
<tr>
<td>b. $59.99</td>
</tr>
<tr>
<td>c. Macy's</td>
</tr>
</tbody>
</table>

- Embed the outlet field within the other follow-up questions
- Collect outlet for every item
Two Ways of Integrating Outlets in the CE Diary

Item-Based

1. Shirt
   a. July 13th
   b. $49.99
   c. Macy’s
   d. Age 16+
   e. Female
2. Blender
   a. July 13th
   b. $59.99
   c. Macy’s

Advantages
• Minimal disruption to original design

Disadvantages
• Additional burden for every item
Two Ways of Integrating Outlets in the CE Diary

- Organize the diary around transactions
- Collect outlet once per transaction regardless of the number of items

<table>
<thead>
<tr>
<th>Transaction-Based</th>
<th>1. Macy’s on July 13th</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a. Shirt</td>
</tr>
<tr>
<td></td>
<td>i. $49.99</td>
</tr>
<tr>
<td></td>
<td>ii. Age 16+</td>
</tr>
<tr>
<td></td>
<td>iii. Female</td>
</tr>
<tr>
<td></td>
<td>b. Blender</td>
</tr>
<tr>
<td></td>
<td>i. $59.99</td>
</tr>
<tr>
<td></td>
<td>c. Sports watch</td>
</tr>
<tr>
<td></td>
<td>i. $32.99</td>
</tr>
<tr>
<td></td>
<td>ii. Age 16+</td>
</tr>
</tbody>
</table>
Two Ways of Integrating Outlets in the CE Diary

Advantages
• Potentially better match with the way people recall expenses

Disadvantages
• Major disruption to the original Diary design

Transaction-Based
1. Macy’s on July 13th
   a. Shirt
      i. $49.99
      ii. Age 16+
      iii. Female
   b. Blender
      i. $59.99
   c. Sports watch
      i. $32.99
      ii. Age 16+
Experiment Design

- **“Control” diary**
  Current CE Diary, item-based, without outlets

- **Control “Plus” diary**
  Current CE Diary, item-based, plus an outlet field embedded in the other follow-up fields

- **“Transactions” diary**
  Redesigned CE Diary organized around transactions, with outlet as an initial field
Participant Task

1. Introduction to CE
2. Told about a fictional expense
   - Item description, price, date, outlet name and location
   - Paper receipt to refer to at any time
3. Record expense into diary
4. Repeat steps 2-3

SUREWAY

SUREWAY STORE 4035-43
184 125TH ST NE WASHINGTON, DC 19328

GROCERY
ARNLD SLC WHT BREAD 1.23 F
REG PRICE 1.55

***TAX 0.00 BAL 1.23
CASH 2.00
CHANGE 0.77

NUMBER OF ITEMS=1

HAVE A GOOD DAY!
12/14/14 8:01AM
Screenshot: “Control” Diary

Date: 12/2/2014
Description: Milk
Item Cost: $3.99

- Food and Drink Away From Home
- Food and Drink for Home Consumption
- Clothes, Shoes, Jewelry, and Accessories
- All Other Products/Services

Packaging:
- Fresh
- Frozen
- Bottled/Canned
- Other

Purchased for someone outside the household
Screenshot: “Control” Diary
Screenshot: “Plus” Diary

Date: 12/2/2014
Description: Milk
Item Cost: $3.99

Packaging: Fresh

Merchant Name: Safeway
City, State: Washington

Paid Online?: No
Screenshot: “Plus” Diary

Add an Expense

Summary

<table>
<thead>
<tr>
<th>Description</th>
<th>Item Cost</th>
<th>Delete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk</td>
<td>$3.99</td>
<td>X</td>
</tr>
<tr>
<td>Bread</td>
<td>$2.50</td>
<td>X</td>
</tr>
</tbody>
</table>
Screenshot: “Transactions” Diary

- Expense Date: 12/2/2014
- Merchant Name: Safeway
- Paid Online: No
- City, State: Washington, DC
- Total Cost: 6.49

Add Item(s)
Screenshot: “Transactions” Diary
Screenshot: “Transactions” Diary
Methods

- Cognitive lab study at the BLS

- $n = 60$ participants; 20 per diary
  - Convenience sample
  - Qualitative test of feasibility
  - Recruited from our database of participants, local community, contractors working in the building

- Task duration about 60 minutes

- Compensation of $40
RESULTS:
REPORTING EXPENDITURES
Data Quality: Item-level reporting

- Bundle “household goods”?
- Make one entry for total expense rather than separate entries for items?
Data Quality: Item-level reporting

- No differences in understanding across diary conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>n Total</th>
<th>n Adequate</th>
<th>n Inadequate</th>
<th>Proportion Adequate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>20</td>
<td>17</td>
<td>3</td>
<td>85%</td>
</tr>
<tr>
<td>Plus</td>
<td>20</td>
<td>15</td>
<td>5</td>
<td>75%</td>
</tr>
<tr>
<td>Transactions</td>
<td>20</td>
<td>19</td>
<td>1</td>
<td>95%</td>
</tr>
</tbody>
</table>

Fisher’s exact test, ns
# Data Quality: Meal-level reporting

**Harvest Table**  
465 14th St, SR  
Washington, DC 20058

<table>
<thead>
<tr>
<th>Item</th>
<th>Quantity</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Sodas</td>
<td>2</td>
<td>6.00</td>
</tr>
<tr>
<td>Artichoke dip</td>
<td></td>
<td>7.00</td>
</tr>
<tr>
<td>Spinach ricotta ravioli</td>
<td></td>
<td>12.00</td>
</tr>
<tr>
<td>Fish special</td>
<td></td>
<td>14.00</td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtotal</td>
<td>39.00</td>
<td></td>
</tr>
<tr>
<td>Tax</td>
<td></td>
<td>3.90</td>
</tr>
<tr>
<td>Total</td>
<td>42.90</td>
<td></td>
</tr>
</tbody>
</table>

**Balance Due**  
42.90

**Gratuity**  
8.40

**Total**  
51.30

- Itemize dishes?
- Omit tax or gratuity?
Data Quality: Meal-level reporting

- No differences in understanding across diary conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>n</th>
<th>Adequate</th>
<th>Inadequate</th>
<th>Proportion Adequate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>20</td>
<td>11</td>
<td>9</td>
<td>55%</td>
</tr>
<tr>
<td>Plus</td>
<td>18</td>
<td>11</td>
<td>7</td>
<td>61%</td>
</tr>
<tr>
<td>Transactions</td>
<td>18</td>
<td>13</td>
<td>5</td>
<td>72%</td>
</tr>
</tbody>
</table>

\[ \chi^2(2, \, N = 56) = 1.22, \, \text{ns} \]
Data Quality: Follow-up questions

- Pay more attention to entering outlets?
- Skip over CE follow-up questions?
Data Quality: Follow-up questions

- No differences in reporting across diary conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>n Total</th>
<th>n Adequate</th>
<th>n Inadequate</th>
<th>Proportion Adequate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>20</td>
<td>18</td>
<td>2</td>
<td>90%</td>
</tr>
<tr>
<td>Plus</td>
<td>19</td>
<td>17</td>
<td>2</td>
<td>89%</td>
</tr>
<tr>
<td>Transactions</td>
<td>20</td>
<td>20</td>
<td>0</td>
<td>100%</td>
</tr>
</tbody>
</table>

Fisher’s exact test, ns
Respondent Burden:
Time taken for single-item transaction

- Task: report a single item purchased from a previously reported outlet
- How much time is added by collecting outlets?
Respondent Burden: Time taken for single-item transaction

- Significantly more time in Transactions condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>n</th>
<th>Total</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>20</td>
<td>53.2</td>
<td>54.3</td>
<td>15.2</td>
<td></td>
</tr>
<tr>
<td>Plus</td>
<td>17</td>
<td>62.4</td>
<td>59.7</td>
<td>25.6</td>
<td></td>
</tr>
<tr>
<td>Transactions</td>
<td>20</td>
<td>81.1</td>
<td>70.7</td>
<td>38.2</td>
<td></td>
</tr>
</tbody>
</table>

\( F(2, 54) = 5.13, \ p < 0.009 \); Transactions > Control, Bonferroni \( p < 0.008 \).
Respondent Burden:
Time taken for multi-item transaction

** Val U Mart **
STORE 326
295 UTANA WAY
ARLINGTON, VA 43503

WELCOME!
LOW PRICES EVERYDAY!

985458 MILK 1GAL 2PCT  4.29
023134 GRNY SMTH APPLES 1.40 lb @ 2.29/lb  3.21
854873 NVL ORANGES 2.34 lb @ 1.59/lb  3.72
875498 CLEMENTINES 5LB  6.99
544893 BAG SPINACH 9OZ  2.99
475454 NBSC CRACKERS  3.49
436754 VALU COOKIES  1.09

**TAX 0.00**
TOTAL  25.71

CASH  30.00
CHANGE DUE  4.22

THANK U
FOR SHOPPING AT VAL U MART

12/13/14  22:21
TRNS 8548-6797

• Task: report multiple items purchased from a new outlet
• How much time is added by collecting outlets?
Respondent Burden: Time taken for multi-item transaction

- No difference in time taken across diary conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>n</th>
<th>Total</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>15</td>
<td></td>
<td>250.5</td>
<td>213.0</td>
<td>85.9</td>
</tr>
<tr>
<td>Plus</td>
<td>16</td>
<td></td>
<td>296.5</td>
<td>263.0</td>
<td>117.9</td>
</tr>
<tr>
<td>Transactions</td>
<td>19</td>
<td></td>
<td>252.8</td>
<td>223.0</td>
<td>113.3</td>
</tr>
</tbody>
</table>

\((F(2, 47) = 0.94, \text{ ns})\).
Respondent Burden: Participant ratings overall

- No difference in burden ratings across diary conditions

- Control (n=19)
  - Very easy
  - Somewhat easy
  - Neither easy nor difficult
  - Somewhat difficult
  - Very difficult

- Plus (n=19)
  - Very easy
  - Somewhat easy
  - Neither easy nor difficult
  - Somewhat difficult
  - Very difficult

- Transactions (n=18)
  - Very easy
  - Somewhat easy
  - Neither easy nor difficult
  - Somewhat difficult
  - Very difficult
Preference for Diary Design

- Out of 20 participants asked for their preference
  - 16 participants preferred Transactions diary
  - 4 participants preferred Plus diary
RESULTS:
REPORTING OUTLETS
Outlet Data Quality: Unusual outlet name

** Val U Mart **
STORE 326
295 UTANA WAY
ARLINGTON, VA 43983

WELCOME!
LOW PRICES EVERYDAY!

985458 MILK 1GAL 2PCT  4.29
023134 GRNY SMTH APPLES
   1.40 lb @ 2.29/1b  3.21
854873 NVL ORANGES
   2.34 lb @ 1.59/1b  3.72
875498 CLEMENTINES 5LB  6.99
541893 BAG SPINACH 9OZ  2.99
475454 NBSC CRACKERS  3.49
436754 VALU COOKIES  1.09

**TAX 0.00 **
TOTAL  25.78

CASH  30.00

CHANGE DUE  4.22

THANK U
FOR SHOPPING AT VAL U MART

12/13/14 22:21  TRNS 5548-6797

• Spell the outlet name accurately?
Outlet Data Quality: Unusual outlet name

<table>
<thead>
<tr>
<th>Reported Outlet Name</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val U Mart</td>
<td>34</td>
</tr>
<tr>
<td>grocery store</td>
<td>1</td>
</tr>
<tr>
<td>VAL U MART STORE 326</td>
<td>1</td>
</tr>
<tr>
<td>VAL-U-MART</td>
<td>1</td>
</tr>
<tr>
<td>ValUMart</td>
<td>1</td>
</tr>
<tr>
<td>Val U Maart</td>
<td>1</td>
</tr>
<tr>
<td>Vall U Mart</td>
<td>1</td>
</tr>
</tbody>
</table>
While you’re out that night [in Washington, DC], you had to pay for parking. You don’t get a receipt so you enter the information as soon as you get home – the meter cost $2.25
## Outlet Data Quality: Unknown outlet name

<table>
<thead>
<tr>
<th>Inadequate Outlet Names</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parking</td>
<td>3</td>
</tr>
<tr>
<td>Park Now</td>
<td>1</td>
</tr>
<tr>
<td>Parking fee</td>
<td>1</td>
</tr>
<tr>
<td>J Street Cinema Parking</td>
<td>1</td>
</tr>
<tr>
<td>HARVEST TABLE</td>
<td>1</td>
</tr>
<tr>
<td>BTC PARKING</td>
<td>1</td>
</tr>
<tr>
<td>Jay's Parking Garage</td>
<td>1</td>
</tr>
<tr>
<td>j street cinema</td>
<td>1</td>
</tr>
<tr>
<td>don't know</td>
<td>1</td>
</tr>
<tr>
<td>PARKING</td>
<td>1</td>
</tr>
<tr>
<td>J Street Cinema</td>
<td>1</td>
</tr>
<tr>
<td>other</td>
<td>1</td>
</tr>
</tbody>
</table>
Outlet Data Quality: Unknown outlet name

- No differences in reporting across diary conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>n Total</th>
<th>n Adequate</th>
<th>n Inadequate</th>
<th>Proportion Adequate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Plus</td>
<td>18</td>
<td>9</td>
<td>9</td>
<td>50%</td>
</tr>
<tr>
<td>Transactions</td>
<td>16</td>
<td>7</td>
<td>9</td>
<td>44%</td>
</tr>
</tbody>
</table>

\[ \chi^2(1, N = 34) = .13, \text{ ns} \]
Outlet data quality collected here is similar to data quality from TPOPS.

Outlet data quality and error types may change due to change in administration mode.
Summary

- The collection of outlet information did not substantively increase or decrease CE data quality
- Participant ratings of burden show no large effects due to outlets
- Possible increase in time taken to enter items in the diary
Next Steps

- Would these results differ for the mobile diary?
- Can usability of the diary designs be improved?
- What is the impact of outlet collection on the CE recall interview?
Contact Information

Erica Yu
Research Psychologist
Office of Survey Methods Research
www.bls.gov/osmr
202-691-7924
yu.erica@bls.gov