

# Results from Testing a Web Mode for the Consumer Expenditure Diary Survey

---

Ian Elkin, Laura Erhard, Brett McBride, and Dawn Nelson

Consumer Expenditure Survey Division/U.S.  
Census Bureau

AAPOR 69<sup>th</sup> Annual Conference

May 15, 2014



# Presentation Outline

---

- I. Web Diary Overview
- II. Results – Sample Performance
- III. Results – Expenditure Comparisons
- IV. Findings and Recommendations

# I. WEB DIARY OVERVIEW

# Consumer Expenditure Diary Overview

---

- The Consumer Expenditure Diary methodology uses PAPI – one week's worth of expenditure entries (x2)
- Census Field Representatives “place” the diary in-person and have up to two more visits
- The CED gathers more detail than the Consumer Expenditure Quarterly Interview
- The Diary is divided into four sections:
  - ▶ Food Away From Home
  - ▶ Food for Home Consumption
  - ▶ Clothing, Shoes, Jewelry and Accessories
  - ▶ All Other Products, Services and Expenses

# Web Diary Overview

---

## Motivation

- ▶ To encourage contemporaneous reporting, improve accuracy through access across locations and household members, provide a platform for individual diaries, allow for a multiportal approach, and to improve response rates by offering alternative response modes to the traditional paper diary.

## Objective

- ▶ To understand the operational issues regarding implementing a web mode for CE and to understand the data quality and response rates associated with web collection.

# Web Diary Logistics

---

- Data Collection Period – Jan, Feb, Mar 2013
- Separate online WD application
- All research cases are “double placed”
- FRs provide username and password to Respondent via CAPI instrument
- New outcome codes for Web Diary (final disposition)
- New materials
  - ▶ FR Talking Points (FR)
  - ▶ User Guide (FR)
- Test Limitations
  - ▶ English only
  - ▶ No Multi-CU HHs or Replacement HHs
  - ▶ No Reinterview

# Web Diary Eligible Cases & Production Sample

---

- Eligible Web Diary Cases
  - ▶ English only
  - ▶ No Multi-CU HHs or Replacement HHs
  - ▶ Home internet access via PC or tablet
- Full Production Sample
  - ▶ All CE Diary cases fielded between January and March 2013
- Restricted Production Sample
  - ▶ Excluded non-English language interviews
  - ▶ Excluded Spawned HHs and Replacement HHs
  - ▶ Excluded HHs reporting no internet access or internet access only through a mobile phone

## **II. RESULTS – SAMPLE PERFORMANCE & DEMOGRAPHICS**



# Sample Demographics & Characteristics

- Pre-tax Income Difference

- ▶ Web Diary Sample–Restricted Production = \$5,000

	Web Diary Sample (N = 456)	Restricted Production (N = 2,261)	Full Production (N = 3,251)
<i>Average Age</i>	50.1	49.6	50.1
<i>Average CU Size</i>	2.6	2.5	2.4
<b>Race (percent)</b>			
<i>White &amp; Other Race</i>	87.1	86.5	85.1
<i>Black</i>	6.8	9.2	10.9
<i>Asian</i>	6.1	4.2	4.0
<b>Ethnicity</b>			
<i>Hispanic</i>	12.7%	9.8%	14.6%

# Sample Demographics & Characteristics

	Web Diary Sample (N = 456)	Restricted Production (N = 2,261)	Full Production (N = 3,251)
<b>Gender</b>			
<i>Male</i>	48.5%	51.0%	49.6%
<i>Female</i>	51.5%	49.0%	50.4%
<b>Education</b>			
<i>Elementary</i>	0.9%	1.4%	3.6%
<i>High school</i>	21.1%	26.7%	33.2%
<i>College</i>	78.1%	71.8%	62.8%
<i>Never attended</i>	-	0.1%	0.4%
<b>Tenure</b>			
<i>Homeowner</i>	73.0%	72.3%	64.4%
<i>Renter</i>	27.0%	27.7%	35.6%

# Overall Response Rates

- Response Rates
  - ▶ Web Diary Sample – 32.0%
  - ▶ Restricted Production – 60.8%
  - ▶ Full Production – 64.5%

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

# **III. RESULTS – EXPENDITURE COMPARISONS**

# Diary Completion

- Analysis excluded non-respondents and removed diaries that FRs indicated had no entries upon pick-up
- Partial Recall – Diaries in which respondent provided entries are augmented by entries collected through the receipt/recall process

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

# Number of Diary Entries

- Total Entries among CUs entering 2 weeks of data without total recall

Total Entries				
	Web Diary Sample (n=155)	Restricted Production (n=801)	Difference (Test-RP)	Change
<b>Mean entries</b>	66.9	75.6	-8.7	-11.5%
<b>Median entries</b>	61.0	69.0	-8.0	-11.6%

- Total Entries among CUs entering 2 weeks of data without total recall, compared against double placed restricted production CUs

Total Entries compared against double-placed RP CUs				
	Web Diary Sample (n=155)	Double-Placed Restricted Production (n=281)	Difference (Test- RP)	Change
<b>Mean entries</b>	66.9	79.4	-12.5	-15.7%
<b>Median entries</b>	61.0	74.0	-13.0	-17.6%

# Amount of Expenditures Reported

- Total Expenditures among CUs entering 2 weeks of data without total recall

Total Expenditure Amounts		
	Difference (Test-RP)	Change
<b>Mean difference</b>	-\$301.82	-15.1%
<b>Median difference</b>	-\$228.16	-17.0%

- Total Expenditures among CUs entering 2 weeks of data without total recall, compared against double placed restricted production CUs

Expenditure Amounts compared against double-placed RP CUs		
	Difference (Test-RP)	Change
<b>Mean difference</b>	-\$138.26	-7.6%
<b>Median difference</b>	-\$196.17	-15.0%

# Number of Entries per Week

- Cases were subset to include CUs not having any entries from partial recall

Entries by CUs completing both weeks (no recall)				
	Web Diary Sample (n=126)	Restricted Production (n=244)	Difference (Test- RP)	Change
<b>Week 1 mean entries</b>	37.8	41.7	-3.9	-9.4%
<b>Week 1 median entries</b>	35.0	37.0	-2.0	-5.4%
<b>Week 2 mean entries</b>	31.7	38.3	-6.6	-17.2%
<b>Week 2 median entries</b>	29.0	32.5	-3.5	-10.8%
<b>Overall mean entries</b>	69.5	80.0	-10.5	-13.1%
<b>Overall median entries</b>	62.0	75.5	-13.5	-17.9%

Within-CU proportion of entries entered in first week by CUs completing both weeks (no recall)		
	Web Diary Sample (n=126)	Restricted Production (n=244)
<b>Proportion (Week 1)</b>	54.8%	53.1%



# Amount of Expenditures per Week

- Cases were subset to include CUs not having any entries from partial recall

Expenditure Totals for CUs completing both weeks (no recall)		
	Difference (Test-RP)	Change
Week 1 mean difference	-3.73	-0.4%
Week 1 median difference	-74.85	-12.8%
Week 2 mean difference	-244.32	-25.1%
Week 2 median difference	-93.84	-15.9%
Overall mean difference	-248.05	-13.2%
Overall median difference	-210.31	-15.5%

Within-CU proportion of expenditure amounts entered in first week of CUs completing both weeks (no recall)		
	Web Diary Sample (n=126)	Restricted Production (n=244)
Proportion (Week 1)	52.4%	52.1%

# Association of Respondent Log-in Patterns

## ■ Expenditures

- CUs that logged-in within the first three days and completed both diary weeks had larger median expenditure amounts (\$146 higher) than those that did not.
- CUs that logged-in within the first three days and completed both diary weeks had comparable median expenditure totals to those in the RP subsample.

## ■ Entries

- CUs completing two weeks of data entry and logging-in early in the reporting period had a greater average number of entries (77) compared to those that did not (65).
- CUs logging-in within the first three days of the reporting period had similar average number of entries (77) as those in the paper diary RP sample (76).

# Association of FR Follow-up Contacts

## ■ Expenditures

- CUs receiving a FR reminder contact had a slightly higher median weekly amount of reported expenditures (\$50 higher).

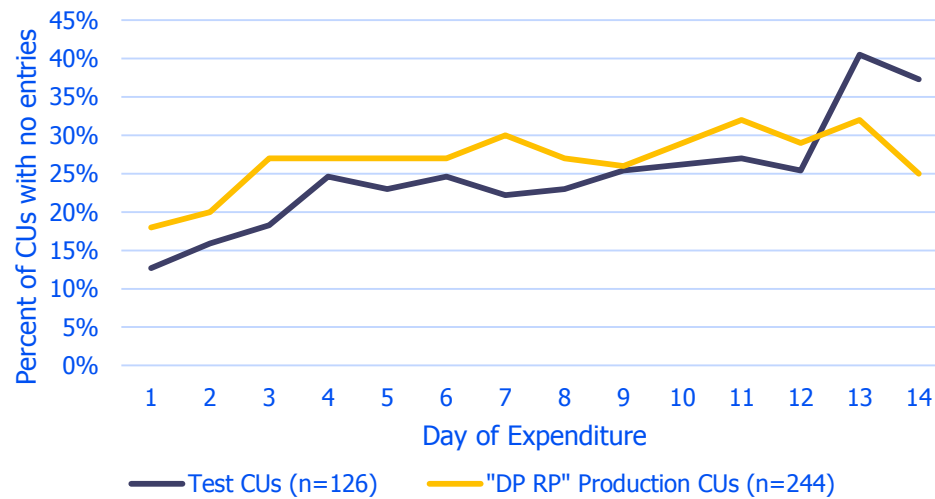
## ■ Entries

- CUs receiving a FR reminder contact had a slightly greater median number of entries than those that did not (33 and 29 respectively).

# Daily Expenditures between Test and RP Diaries

- Although, RP CUs have a higher number of entries on average than test CUs, a greater percentage of RP CUs had days with no expenditures

Change in Percent of CUs with No Daily Expenditures (among 2-week diary completers)



# **IV. FINDINGS & RECOMMENDATIONS**

# Findings

---

- Higher median expenditures for the following sections:
  - ▶ Clothing, Shoes, Jewelry, and Accessories
  - ▶ Food and Drinks for Home Consumption
- Lower item nonresponse, as defined for a diary instrument
- Week-to-week drop-off validates shift to one week collection period
- Key analysis – Expenditure reporting by log-in frequency

# Recommendations

---

## ■ Low Response Rates

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

# Recommendations

---

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



# Recommendations

---

## ■ Low Placement Rates

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

## ■ Higher Ineligible Rates

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

# Recommendations

---

- Other

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

# Acronym Definitions

---

- WD – Web Diary
- FR – Field Representative
- R – Respondent
- CU – Consumer Unit
- HH - Household
- CAPI – Computer-assisted Personal Interviewing
- PAPI – Paper-and-Pencil Interviewing

# Contact Information

---

**Ian J. Elkin**

**[elkin.ian@bls.gov](mailto:elkin.ian@bls.gov)**

**Full Report:**

**[http://www.bls.gov/cex/ce\\_statpa\\_pers.htm](http://www.bls.gov/cex/ce_statpa_pers.htm)**