Respondent Record Use in the US Consumer Expenditure Survey

Jennifer Edgar
Bureau of Labor Statistics

AAPOR 2010
Consumer Expenditure Survey

- National federal survey, average 65 minutes
- Panel survey, 5 quarterly waves
- Designed to be personal visit (~30% phone)
- Respondents report for the household
- Expenditure questions ranging from large, regular items (mortgage) to small, infrequent (clothing) items
  - Detailed follow-up questions (description, month, cost, sales tax)
- General philosophy: “more is better”
“Most questions are about expenses your household had or bills you’ve received. You will find it helpful to have your checkbook register, credit card statements and other records available as you answer the questions.”

“Please refer to any billing statements or other records you have when answering these questions.”
Respondent Record Use in the CEQ

Research on CEQ data shows a relationship between record usage and:
- Reporting levels (Safir & Goldenberg, 2008)
- Underreporting (Tucker, Meekins & Biemer, 2008)
- Data quality (Gonzalez & Edgar, 2009)
Data

April 2006 through March 2008
Waves 2 through 5

Interviewer question after interview asking how often respondent used records and what type of records were used

44,300 completed personal visit interviews
21,011 unique households
Method

- Compared record use by
  - Respondent characteristics
  - Household characteristics
  - Interviewer characteristics

- All statistical tests (Chi Square and ANOVA) were statistically significant and not reported
  - Likely due to sample size
Reported Record Usage

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always</td>
<td>15.2%</td>
</tr>
<tr>
<td>Almost always</td>
<td>16.2%</td>
</tr>
<tr>
<td>Mostly</td>
<td>10.8%</td>
</tr>
<tr>
<td>Occasionally</td>
<td>18.9%</td>
</tr>
<tr>
<td>Almost never</td>
<td>7.5%</td>
</tr>
<tr>
<td>Never</td>
<td>31.1%</td>
</tr>
</tbody>
</table>
Respondent Characteristics

- Included only unique households
- Several interesting trends in who does and does not use records
- Males were slightly less likely to use records than females
By Education

By Education

Less than High School
HS Graduate
Some college, no degree or AA degree
Bachelor's degree
Graduate degree
Household Characteristics: Family Size and Type

- As family size increased, the likelihood that a respondent will use records decreases
  - Exception: two-person households were the most likely to use records
- Husband and wife only households were most likely to use records
  - Single consumers and ‘other’ household types were least likely
Household Characteristics: Tenure

- Homeowners without a mortgage were most likely to use records.
- Households residing in student housing were least likely to use records.
Household Characteristics: Income Reporting

![Bar chart showing income reporting frequency]

- Complete Income Reporter
- Incomplete Income Reporter

- Always: 10%
- Almost always: 20%
- Mostly: 25%
- Occasionally: 35%
- Almost never: 50%
- Never: 50%
Interview Characteristics

- Interviews where:
  - The advance letter was reported to have been received were more likely to use records
  - The advance letter was not received were more likely to never use records
  - The Information Booklet was used were more likely to use records
  - The Information Booklet was not used were more likely to never use records
Interview Characteristics: By Wave

Always
Almost always
Mostly
Occasionally
Almost never
Never

0%
5%
10%
15%
20%
25%
30%
35%

BLS 15
Interview Characteristics: Converted Refusals

- Always: Yes 10%, No 40%
- Almost always: Yes 15%, No 35%
- Mostly: Yes 20%, No 20%
- Occasionally: Yes 25%, No 10%
- Almost never: Yes 30%, No 5%
- Never: Yes 45%, No 0%
Interview Characteristics: Number of Contacts

- Always: 17%
- Almost always: 20%
- Mostly: 15%
- Occasionally: 25%
- Almost never: 10%
- Never: 35%

Legend:
- 1
- 2
- 3 to 5
- 6 or more
Interview Characteristics: Interview Length

- Less than 41 minutes
- 41 to 59 minutes
- 60 to 81 minutes
- More than 81 minutes
Interview Characteristics: Section Length

- Some specific CEQ sections probably benefit most from the use of records, (utilities, health-insurance and income)
- Respondents who used records took longer to complete these sections than respondents who did not
- This trend found in sections not thought to benefit from records (entertainment) too
- Record usage relates to longer interviewing time, regardless of topic
Interview Characteristics: Reporting Rates

- Number of expenditures ranged from 0 to 179 items
  - Mean = 29.4
  - Standard deviation = 17.5
Reporting Rates

- First quartile
- Second quartile
- Third quartile
- Fourth quartile

- Always
- Almost always
- Mostly
- Occasionally
- Almost never
- Never
Expenditure Amounts

Total expenditure amount reported by each respondent ranged from $0 to $424,981

Mean = $5,180
Standard deviation = $8,566
Expenditure Amounts

- First quartile
- Second quartile
- Third quartile
- Fourth quartile

Bars represent different frequency levels:
- Always
- Almost always
- Mostly
- Occasionally
- Almost never
- Never

 (%)
Data Quality

- Any type of editing required (imputation or allocation) was identified per expenditure report and summed across an interview.
- Proportion of editing calculated.
  - Ranged from 0.0 (no editing required) to 1.0 (all reports required editing).
- Forty-two percent of interviews required no editing.
- The mean proportion was 0.12, with a standard deviation of 0.16.
Data Quality

[Bar chart showing frequency of data editing]

- Always
- Almost always
- Mostly
- Occasionally
- Almost never
- Never

- No Editing
- Second quartile
- Third quartile
- Fourth quartile
Certain types of respondents (females, older, highly educated) and households (two person, homeowners) were more likely to use records. Respondents who reported receiving the advance letter, and used the Information Booklet were more likely to use records. Respondents who had to be convinced to participate were less likely to use records. Using records is related to longer interviews, more reports and higher reports… but…
Conclusions

...record usage was not consistently related to higher quality data quality

So, what should we conclude?

- Record use likely increases respondent burden: is it worth it?
- Should we be collecting records rather than respondent answers?
- Is more really better?
Next Steps: Additional Analysis

- Existing data
  - Multivariate analysis: mitigating variables between data quantity and quality?

- New data
  - Section level record usage
  - Information about how respondents use records
  - What “always” vs. “almost always” means to interviewers
Jennifer Edgar
Edgar.Jennifer@Blsl.gov