

Data Quality in the Online Panel Self-Administered Diary Test of the Consumer Expenditure Survey

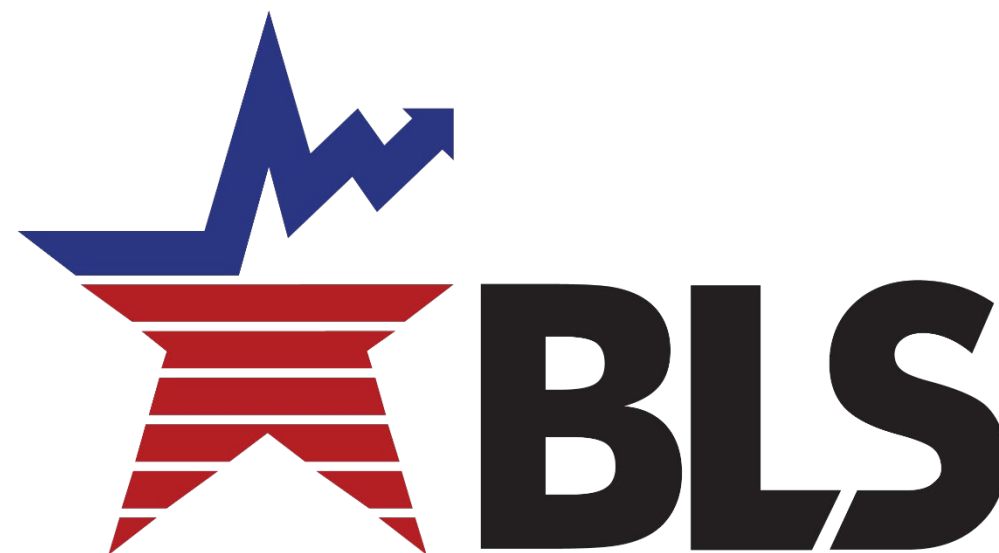
Gray Jones, Nikki L. Graf, Tucker Miller, and Wendy Carlton
U.S. Bureau of Labor Statistics

2023 FCSM Research & Policy Conference
October 25, 2023



Presentation Outline

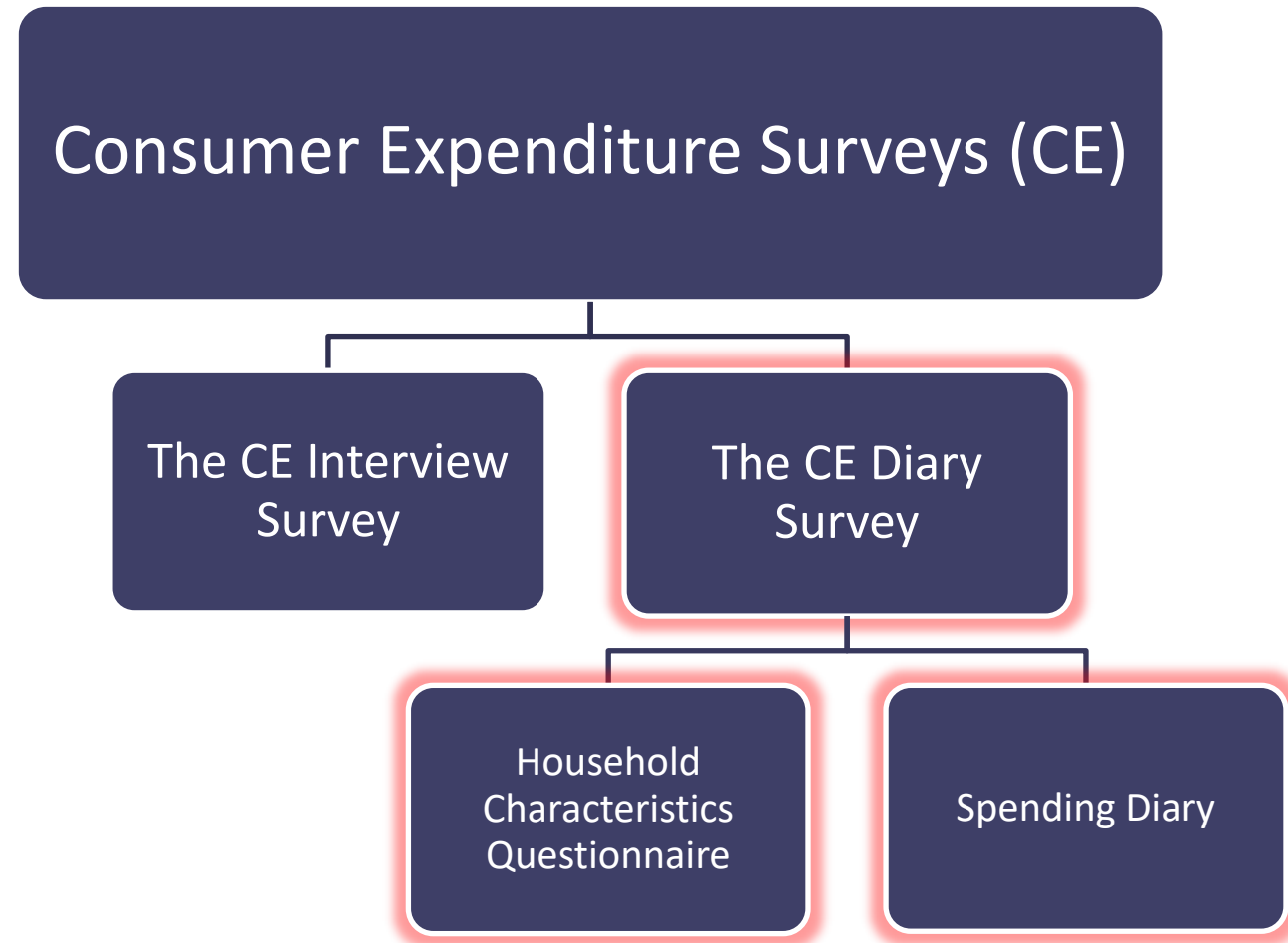
- Consumer Expenditure Surveys (CE) background
- Online Panel Self-Administered Diary (OPSAD) background
- Comparison OPSAD and CE
- Analysis of Data Quality
- Summary & Next steps



Consumer Expenditure Surveys (CE) Background



Consumer Expenditure Survey (CE) Background



CE Spending Diary Interviewer-Aided Protocol

Diary placement



Train respondents to use the diary

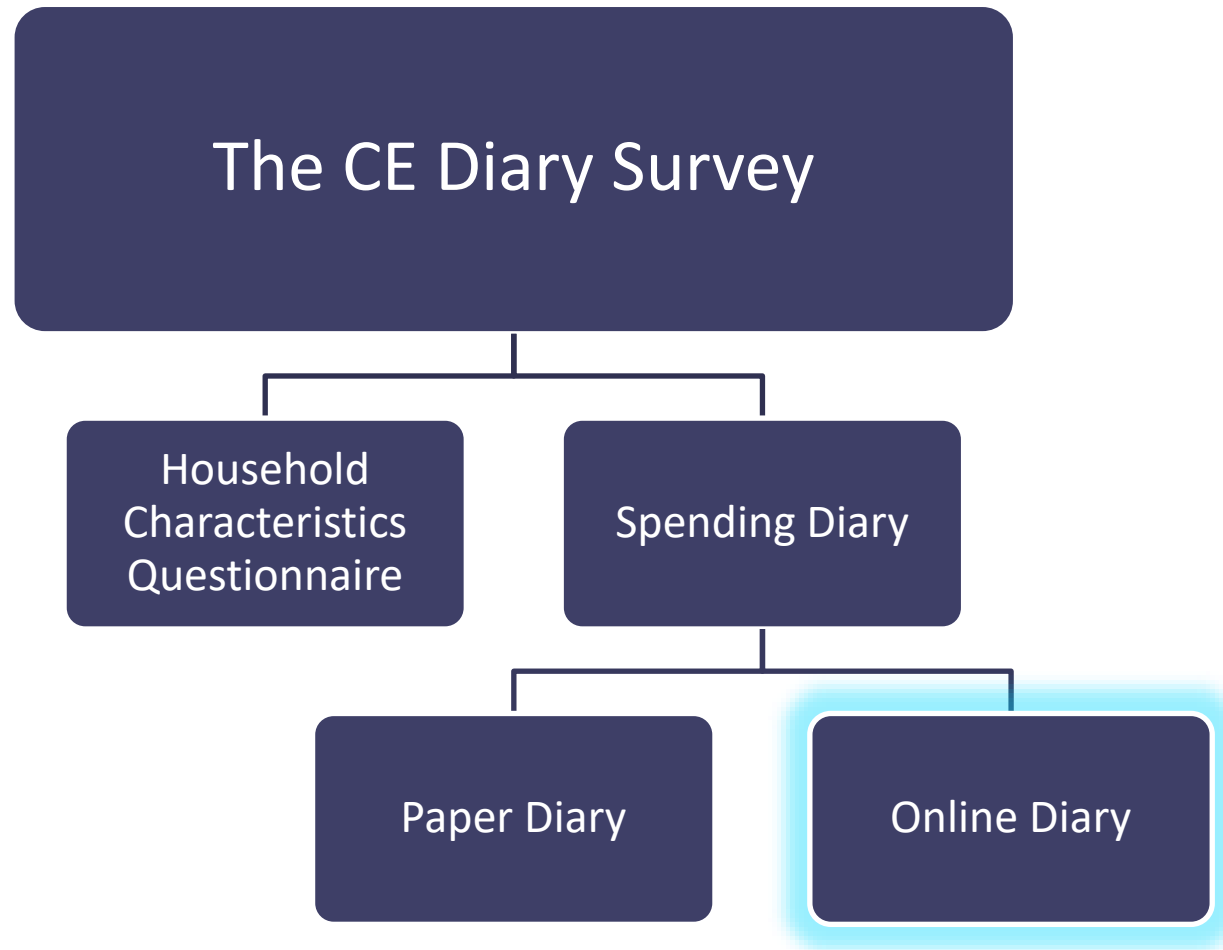


Follow-up to encourage diary keeping



Collect any forgotten expenditures at the end of the diary period

CE Spending Diary Online Mode



Online Panel Self-Administered Diary (OPSAD) Test Background



OPSAD Test Background

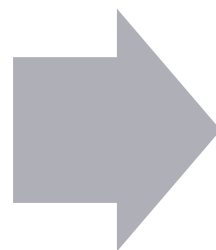
OPSAD Test Design

- Test conducted on the Ipsos KnowledgePanel®
- Fielded between November 2021 – January 2022
- Total starting sample of 12,700 adults representing their household
- Two stage survey design (based on the CE Diary Survey)
- Invitation and reminder emails sent for each stage
- Conditional incentives used at each stage

OPPSAD Test Background: OPPSAD Survey Components

Household
Characteristics
Survey (HCS)

- \$2 Incentive



Two-week
Consumer Spending
Diary

- \$50 Incentive

CE Diary Survey and OPSAD Comparison



CE Diary and OPSAD Comparison: Structure

CE Diary Survey

- Components of the survey design
 - Household Characteristics Questionnaire
 - Spending Diary
- Data fielded between November 2021 and January 2022
- **Interviewer-Aided**
- No incentives
- Regular CE Sample

OPSAD Test

- Components of the survey design
 - Household Characteristics Survey
 - Spending Diary
- Data fielded between November 2021 and January 2022
- **Fully Self-Administered**
- Conditional Incentives
- Sample from Ipsos KnowledgePanel®

CE Diary and OPSAD Comparison: Completion by Diary Week

Households with at least one entry in...	CE Diary Survey	OPSAD
Both diary weeks	1,586	2,702
Only week 1	15	424
Only week 2	21	28
Total with entries in either or both weeks	1,622	3,154



Analysis of OPSAD Data Quality & Comparison to the CE Diary Survey Data Quality



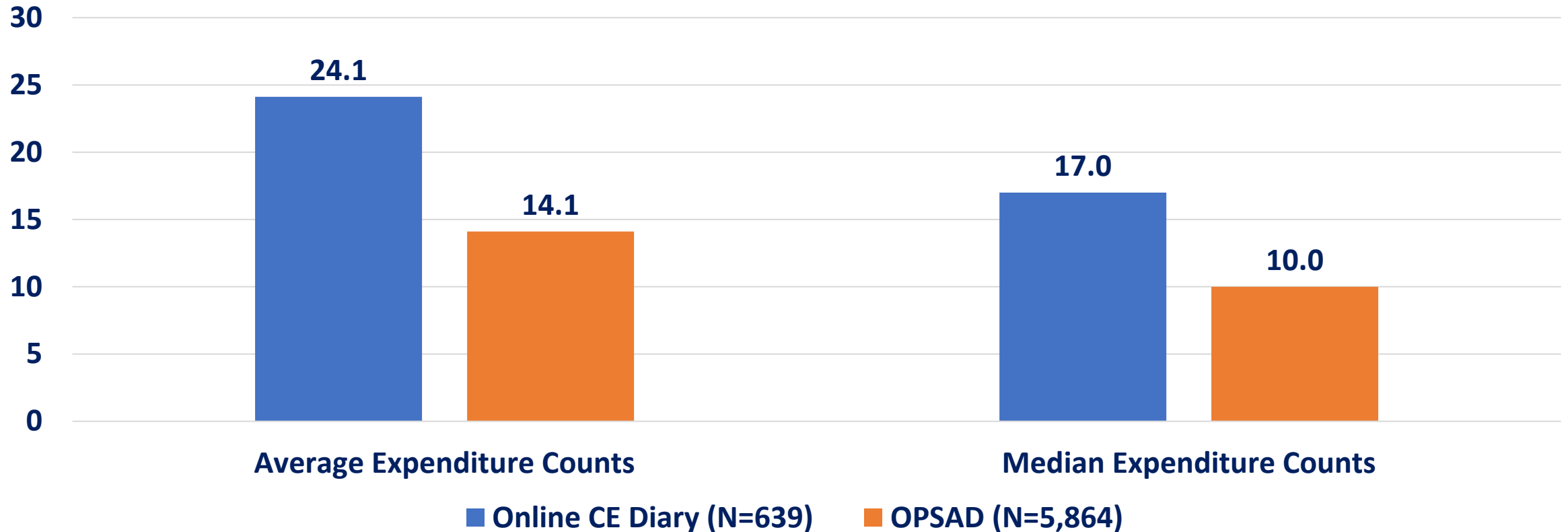
CE Diary and OPSAD Comparison: Sample Composition

- Comparison to focus on differences between CE “Online” Diary users and OPSAD users to control for mode effects.
- Past CE research has found meaningful differences between “Paper” and “Online” Diary users (McBride & Graf 2022).
- Preliminary results shown here are weighted.

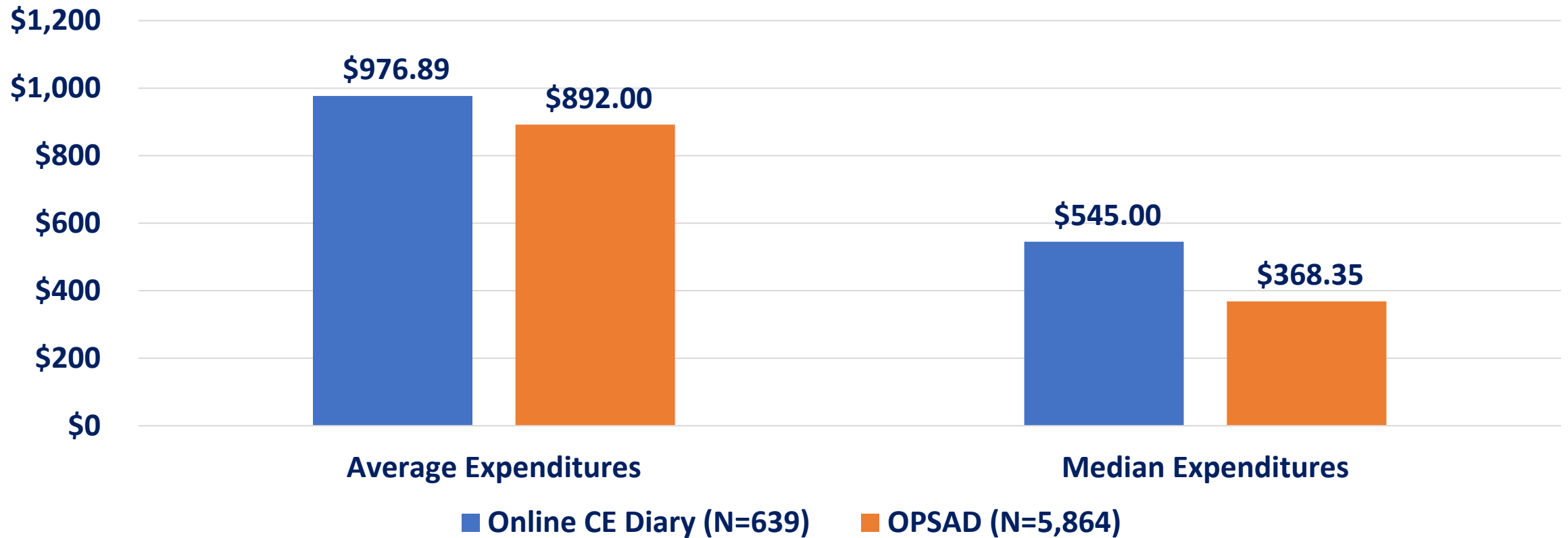
OPSAD Data Quality: Item Nonresponse

% of Diary Entries with...	Online CE Diary	OPSAD
Blank Cost Values	0.3%	2.07%
Blank Item Descriptions (Excluding Meals Away from Home)	0.0%	1.05%
Invalid Item Descriptions (Excluding Meals Away from Home)	0.1%	1.58%

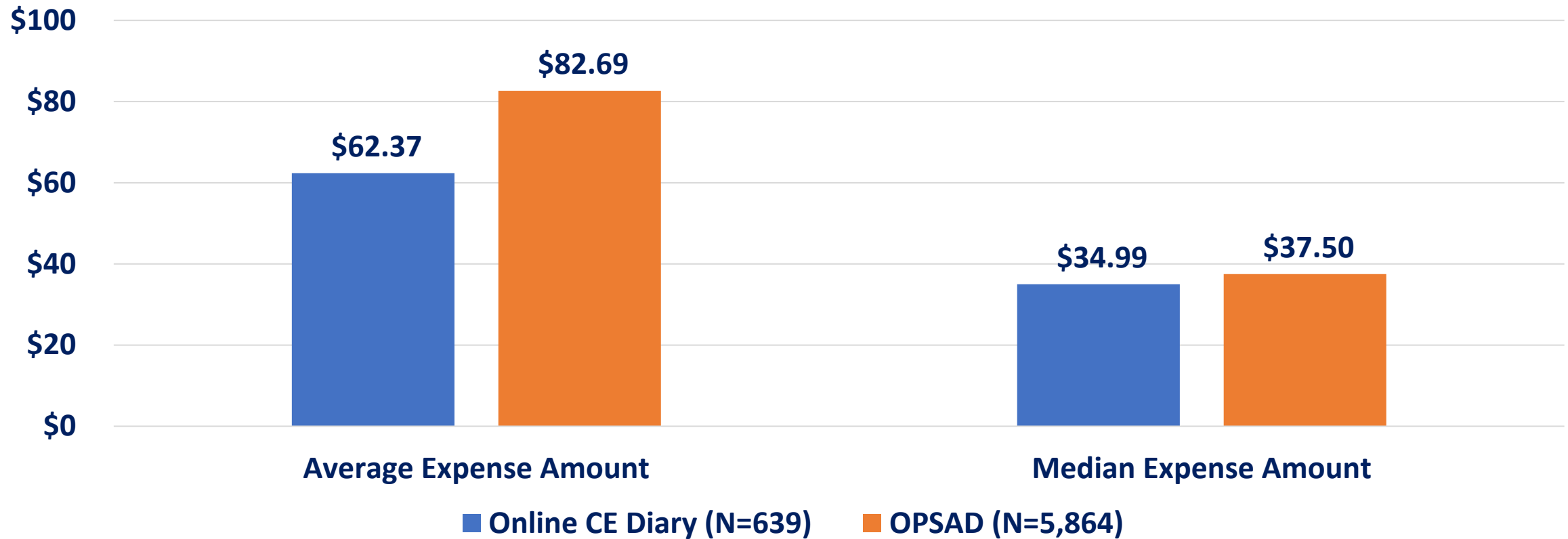
OP SAD Data Quality: Average Expenditure Counts



OPSAD Data Quality: Average Total Expenditures



OPSAD Data Quality: Average Expenditure Amounts



OPSAD Data Quality: Bundled Expenses

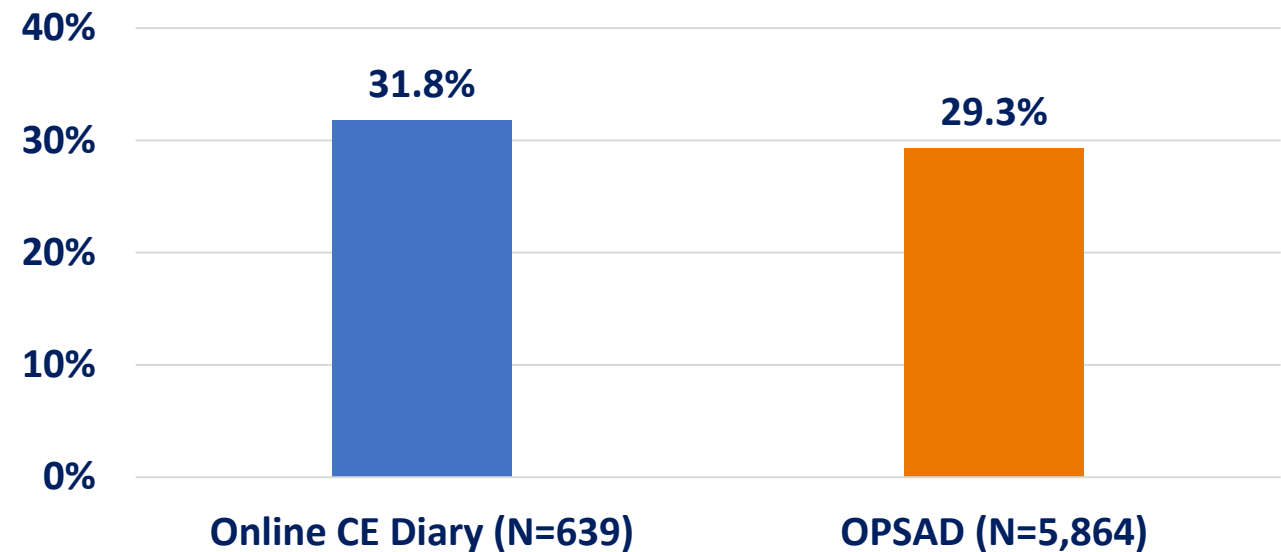
- Item descriptions containing the following strings were flagged as “bundled” entries:
 - ▶ GROCERIES
 - ▶ FOOD
 - ▶ CLOTHES
 - ▶ CLOTHING
 - ▶ HOME IMPROVEMENT
 - ▶ HOUSEHOLD
 - ▶ MISCELLANEOUS
 - ▶ MISC

OPSAD Data Quality: Expenditure Entry Bundling

- The rate of diaries with any bundled expenses

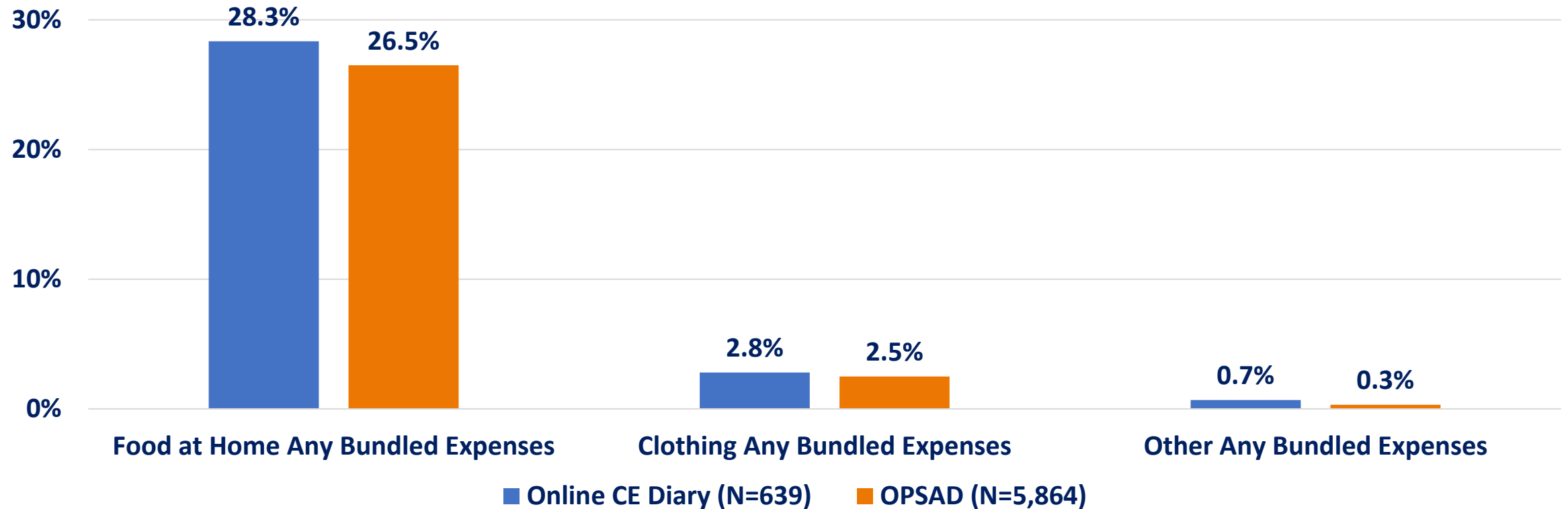
▶
$$\frac{\# \text{ of Diary Cases with Bundled Expenses}}{\text{Total \# of Diary Cases}}$$

Percentage of Diary Cases with Any Bundled Expenses



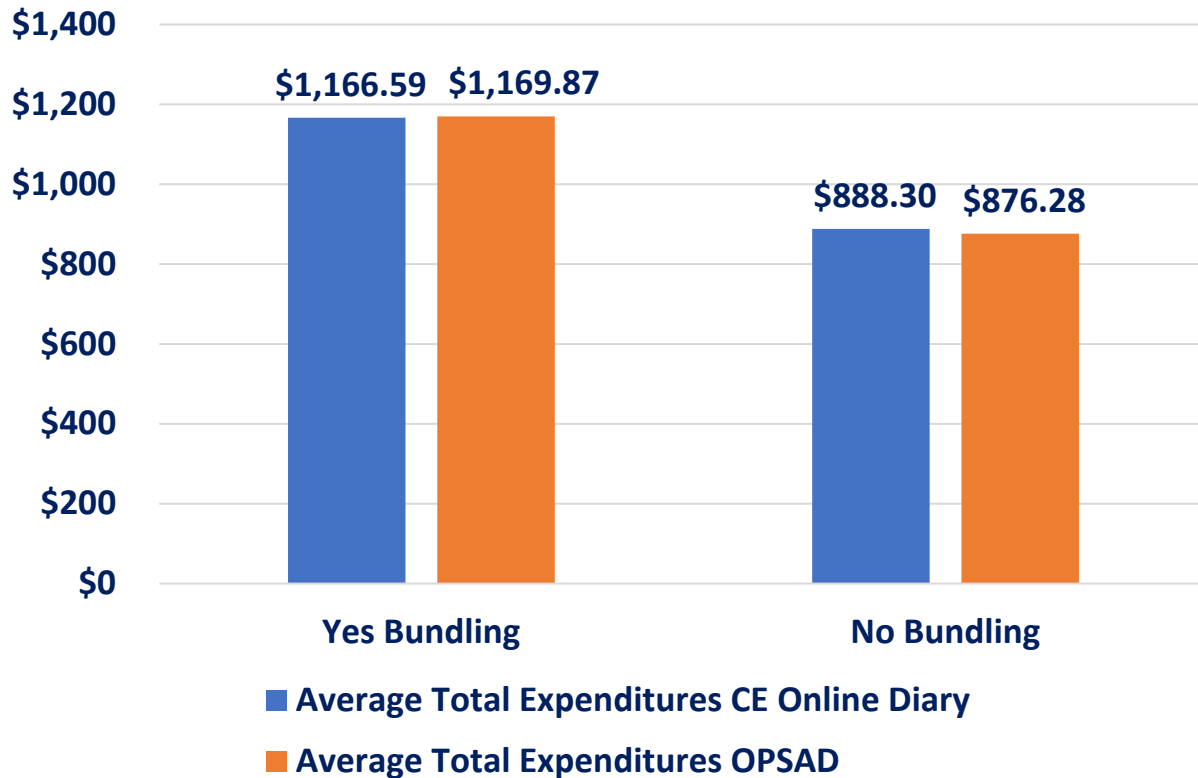
OP SAD Data Quality: Expenditure Entry Bundling

Percentage of Diaries with Bundled Expenses by Expenditure Category

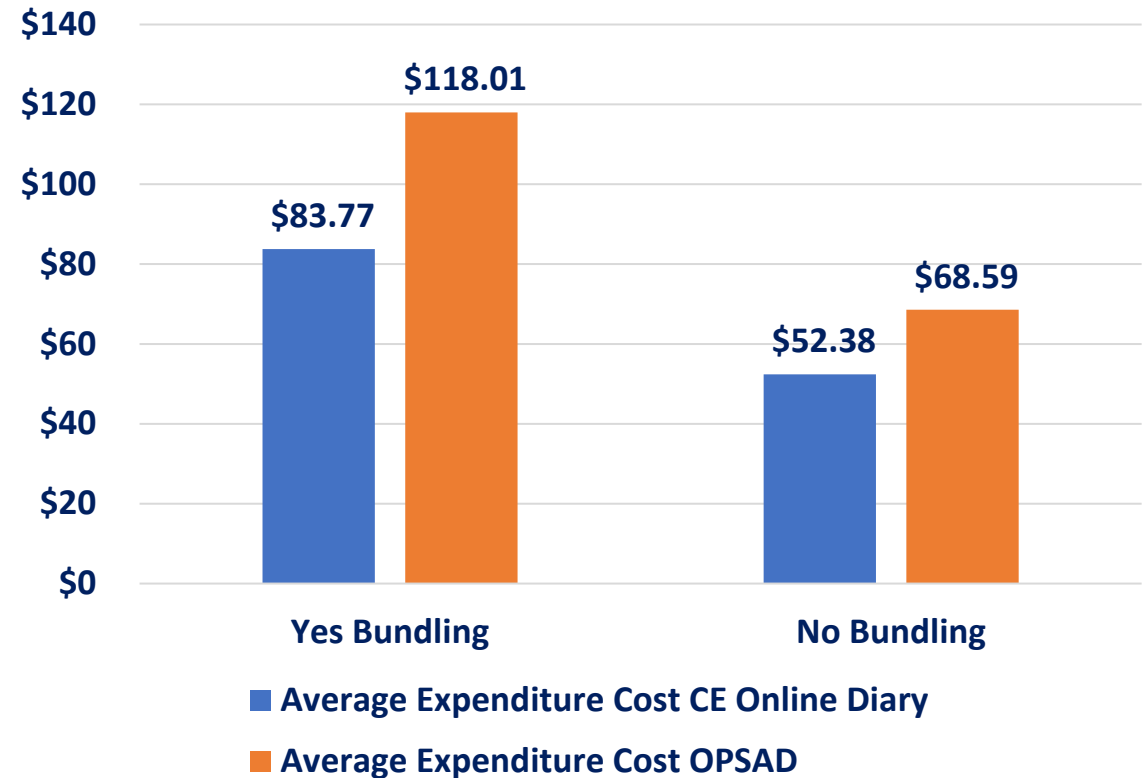


OPSAD Data Quality: Cost and Bundling

Comparison of Average Total Expenditures



Comparison of Average Expenditure Cost



Summary

- Measures of data quality were generally more favorable within the CE Production Diary data.
 - ▶ Lower Rates of Item Nonresponse
 - ▶ Higher Average Expenditure Totals
 - ▶ Higher Average Entry Counts
 - ▶ Similar Rates of Bundling Expenses
- These differences may be attributable to the lack of interviewer guidance in the OPSAD test.

Next Steps

- Further investigate and account for the differences in the underlying samples.
 - ▶ As the OPSAD sample and the subset of Online Diaries from the CE sample are not 1:1
 - ▶ Comparisons to the full CE production sample.
 - ▶ Will involve a more detailed analysis of data quality, using multivariate models to introduce controls.
- Final OPSAD Analysis report is forthcoming.

Thank you!

- Gray Jones, Economist
- Email: Jones.graham@bls.gov
- Phone: 202-691-6869

