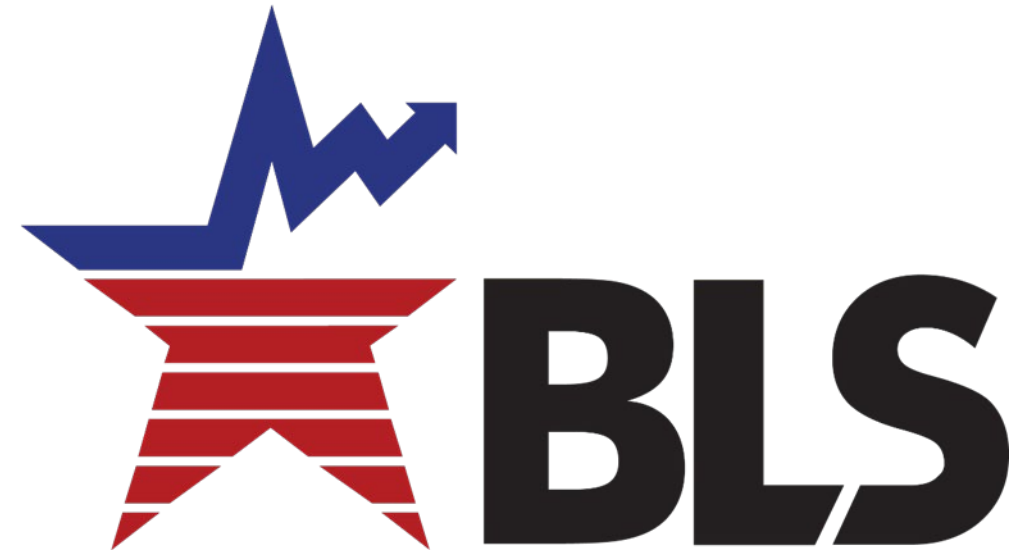


**Assessing the Impact of Device Type on Data
Quality in the Consumer Expenditure (CE)
Online Diary Survey**
Gray Jones, BLS
AAPOR Annual Conference 2024
May 15, 2024

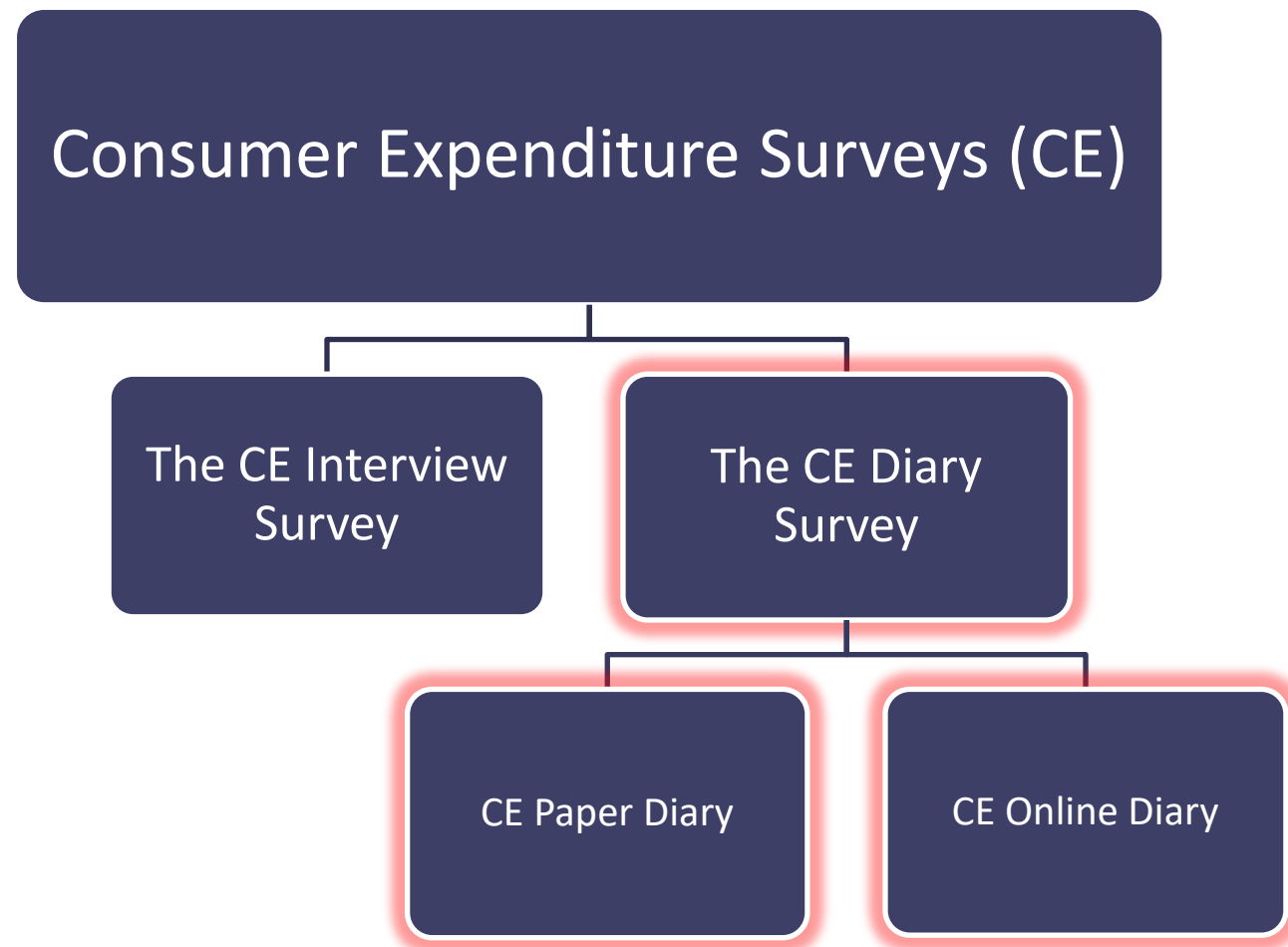


Agenda

- CE Diary Survey Background
- CE Online Diary Background
- Motivation for Analysis
 - Past Diary Paradata Research
- Device Use in the Online Diary
- Analysis Results:
 - Devices and Demographics
 - Reporting Behavior
 - Relationship to Data Quality
- Next Steps for CE Paradata use



CE Diary Survey Background



CE Online Diary Background



CE Online Diary Paradata

- Data generated as a by-product of the survey's a collection process
- Analysis of paradata can lead to...
 - ▶ Improved data collection efficiency
 - ▶ Reduced survey fielding costs
 - ▶ Better understanding of measurement error



Analysis of CE Online Diary Paradata

- BLS has already begun analyzing CE Diary Survey Paradata
 - ▶ Large Scale Feasibility Test Preliminary Report (Krishnamurty et al., 2021)
 - ▶ 2021 Consumer Expenditure Survey Online Diary Paradata Analysis (Jones, G., and P. Krishnamurty)



Motivation for Research

- Past analysis found differences by the device type used.
- Primarily driven by “Mobile Only” device users.

2021 Paradata Analysis Results by Device Type	Desktop Only (n=1,162)	Mobile Only (n=387)	Mixed Use (n=252)
Median Logins per CU	10 logins	6 logins	11 logins
Median Time Spent per CU	27.5 minutes	10.3 minutes	26.5 minutes
Median Expenditure Count	34 entries	27 entries	33 entries
Median Expenditure Total (In Dollars)	\$555.31	\$400.50	\$559.00

Motivation for Research

- These findings led to questions about data quality for mobile users.
 - ▶ Are discrepancies in reporting driven mostly by demographic differences? (e.g., lower income → lower entries and expenditures)
 - ▶ Or are differences driven by mobile device users reporting behavior?



Purpose of Analysis

- Disentangling the relationship between respondent characteristics, device use, and diary reporting behavior.
- Important for optimizing survey fielding procedures and understanding data quality.



Device Type Analysis



Device Type Analysis: Data Use and Set Up

■ Data Used

- ▶ CE Diary Data from 2022q3 and 2022q4

■ Pre-Analysis Steps

- ▶ Generate device type information using CE Online Diary Paradata.
- ▶ Generate expenditure counts and totals (in dollars) using unprocessed expenditure data.
- ▶ Combine these measures with Consumer Unit (CU) level respondent characteristics data associated with complete cases.

Device Type Analysis: Device Definitions

■ Desktop Only User

- ▶ Includes online diary users accessing the instrument with either a desktop, laptop, or tablet, but never a mobile device.

■ Mobile Only User

- ▶ Includes online diary users accessing the instrument with only a mobile device (never a desktop, laptop, or tablet).

■ Mixed Users

- ▶ Includes online diary users who accessed the instrument with a mobile device, as well as a desktop, laptop, or tablet.

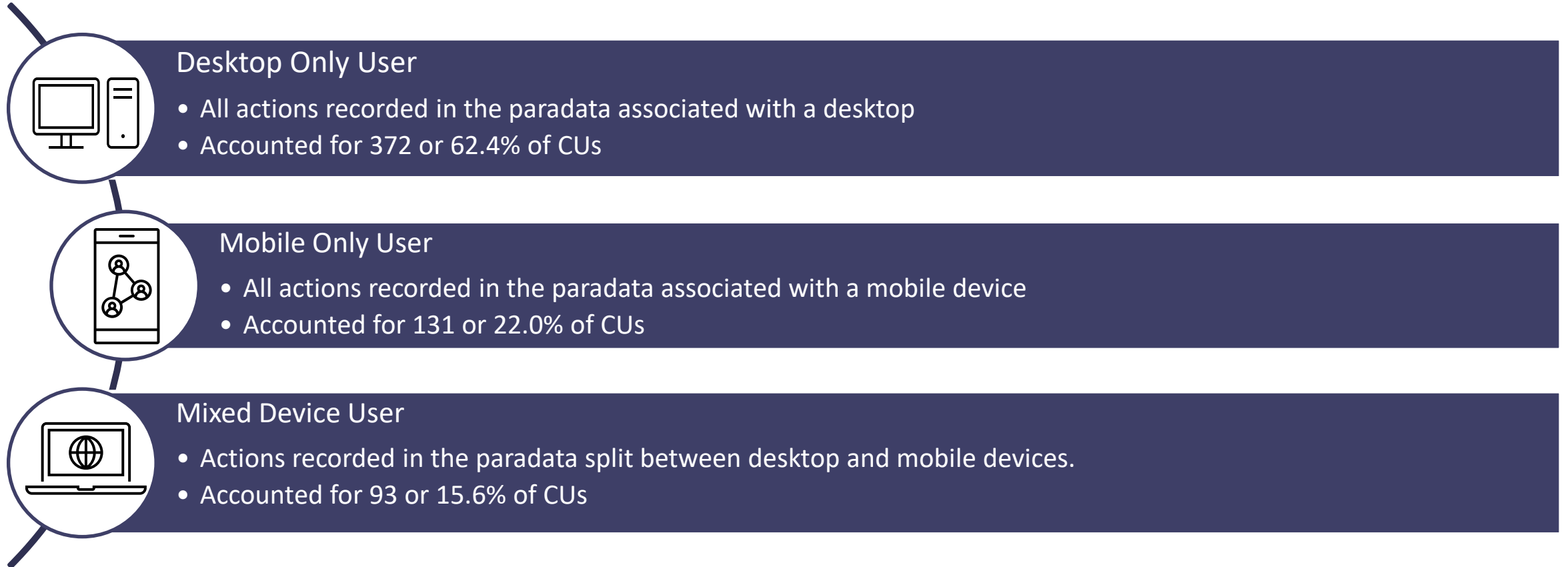
Device Type Analysis

■ Analytical Steps

1. Overview of Device Use
2. Comparison of Nonresponse
3. Comparison of Demographic Characteristics
4. Multivariate analysis



Overview of Device Use



Nonresponse Indicators by Device Type

- In addition to previously mentioned paradata metrics.
- Mobile Only device users have slightly higher rates of **Item Nonresponse & Bundling**.

Nonresponse Indicators	Desktop Only (n=372)	Mobile Only (n=131)	Mixed Use (n=93)
Item Nonresponse (Blank Costs)	0.4%	1.0%	0.3%
Invalid Item Descriptions	1.7%	2.0%	2.8%
Average Bundling Rate	11.7%	13.4%	9.7%

Main Research Questions

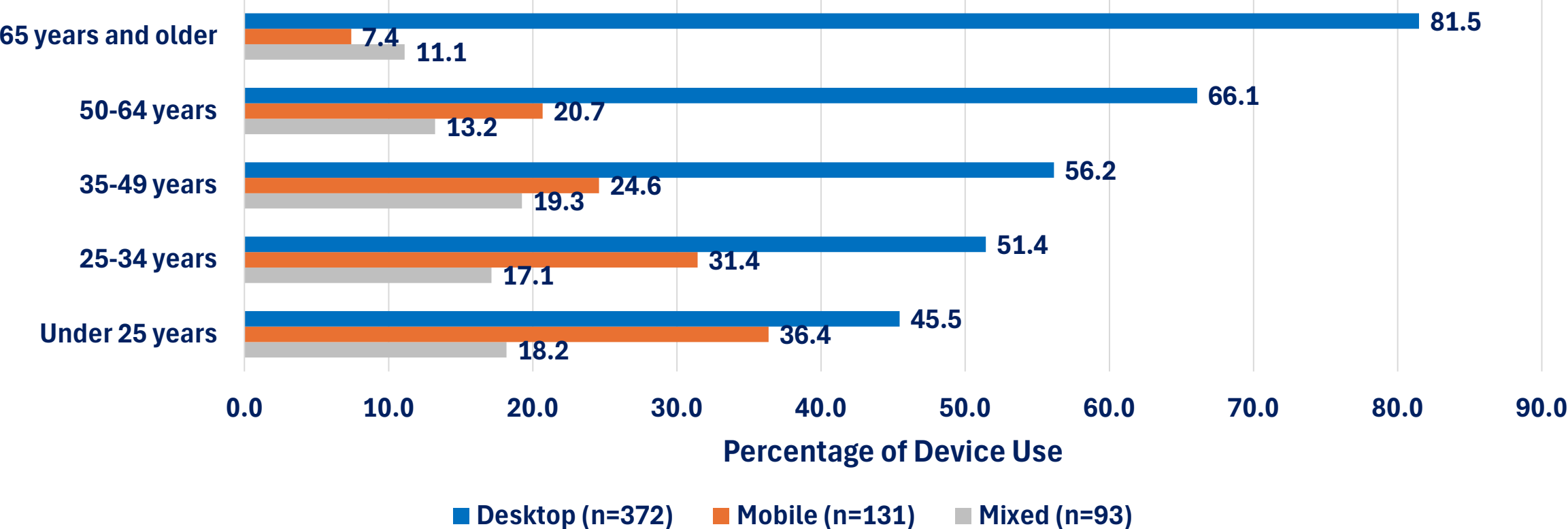
1. Are discrepancies in reporting driven by demographics?
2. Or are differences driven by mobile device users reporting behavior?



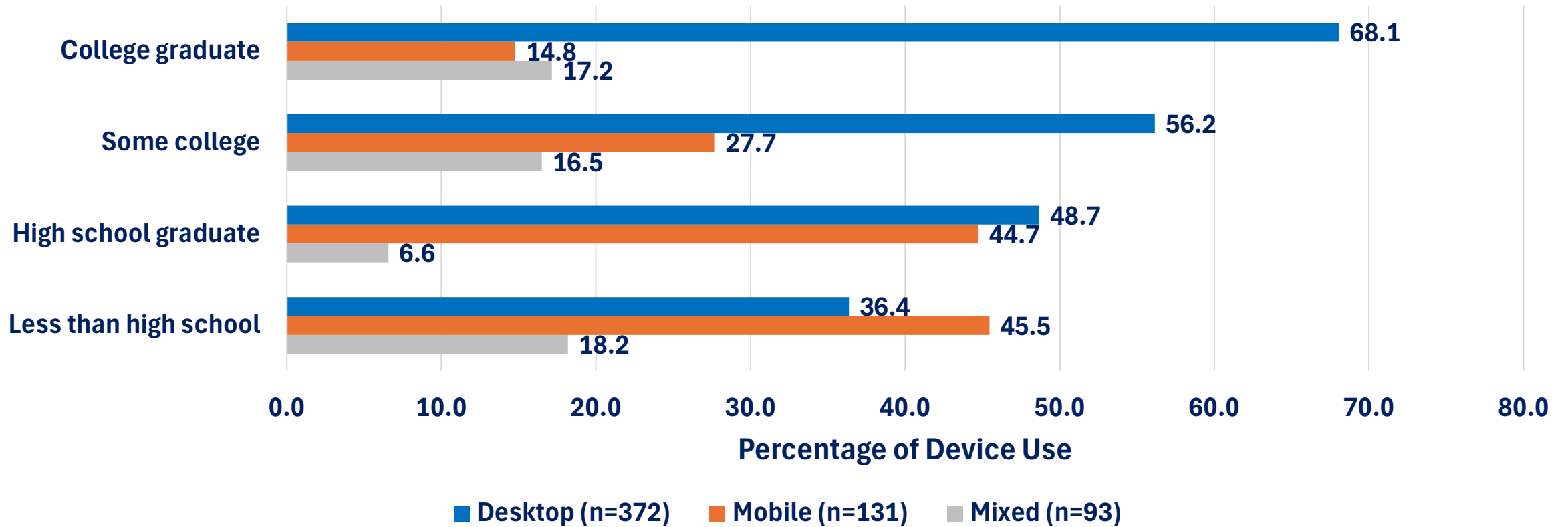
Comparison of Demographic Characteristics



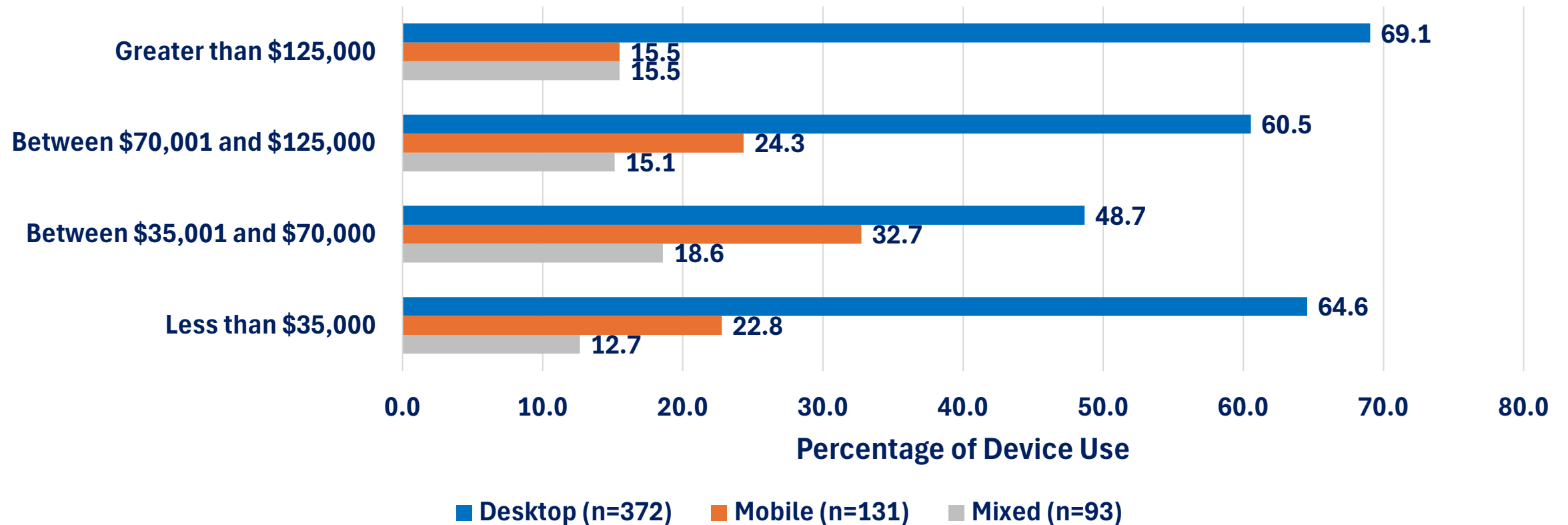
Demographic Comparison: Device Use by Age Group



Demographic Comparison: Device Use by Education Level



Demographic Comparison: Device Use by Pre-Tax Income Level



Multivariate Analysis



Multivariate Analysis

- Regression model specified to control for demographic characteristics.
- **Two Models:**
 - ▶ Full Model
 - ▶ Reduced Model
- **Dependent variable:** Diary Entry Counts



Multivariate Analysis: Full Model

- **Dependent Variable:** Diary Entry Counts
 - ▶ Transformed to the Natural Log of Diary Entry Counts
- **Main Explanatory Variable of Interest:** Mobile Only Device Use
- **Additional Control Variables:** Age, Allocation, Area Type, CU Size, Data Collection Month, Diary Recall, Education, Ethnicity, Gender, Homeowner, Income, Race, and Region.

Multivariate Analysis: Full Model

- Results from the full model show that the impact of “Mobile Only” device use on entry counts was significant.
- Elements of Age, CU Size, Race, and Region were also found to be significant in the full model.
- These findings informed a reduced model.



Multivariate Analysis: Reduced Model

- “Mobile Only” diary use still strongly significant.
- Interpretation of the coefficient implies a negative relationship with expenditure counts.
- Strengthens claim that “Mobile Only” diary use is associated with lower quality expenditure diary reporting.

Variables in Reduced Model	Coefficient	Standard Error
<i>Intercept</i>	3.74	0.148
Online Diary Device: Mixed Use	0.01	0.103
Online Diary Device: Mobile Only***	-0.47	0.093
Race: Other*	-0.24	0.111
Race: Black	-0.24	0.138
Age: Under 25**	-0.57	0.210
Age: 25 - 34*	-0.29	0.126
Age: 35 - 49	-0.22	0.116
Age: 50 - 64	0.03	0.111
CU Size: 1***	-0.47	0.092
CU Size: 4+	0.12	0.100
CU pre-tax income: \$35K-<\$70K	-0.07	0.133
CU pre-tax income: \$70K-<\$125K	0.14	0.130
CU pre-tax income: >\$125K	0.25	0.131
Region: Northeast*	-0.28	0.107
Region: Midwest	-0.01	0.106
Region: West	-0.02	0.104

*p < 0.05, **p < 0.01, ***p < 0.001



End of Device Type Results



Conclusions and Recommendations

- Using **only** a mobile device appears to be associated with sub-optimal diary reporting, even when controlling for other relevant variables.
- The rate of “Mobile Only” diarists should be monitored moving forward to as a data quality metric.
- Consider discouraging “Mobile Only” reporting as a quality assurance measure and pushing those respondents to paper.

Contact Information

Gray Jones
Economist

Consumer Expenditure Surveys

jones.graham@bls.gov

