

BACKGROUND

Introduction

Historically, the Bureau of Labor Statistics (BLS) Consumer Expenditure (CE) diary survey asked respondents to log all expenses for two consecutive one-week periods in a paper form. As part of a CE diary survey redesign, an online version was made available as an alternative option to the paper format of the diary survey in 2020. The current diary placement protocols require interviewers to train and guide respondents on how to properly complete the diary survey.

From November 2021 to January 2022, the BLS conducted a test of a fully self-administered diary in a probability-based online panel (Ipsos KnowledgePanel), where participants had no interviewer interaction. **The purpose of the test was to gauge the impacts of shifting from current diary placement protocols to a fully self-administered mode on data quality and to investigate whether a self-administered online diary in an online panel could be an avenue for testing changes for CE's production diary.** Additionally, the BLS expected that key insights would originate from paradata analyses.

Data

This poster is a preliminary analysis of paradata collected from the **Online Panel Self-Administered Diary (OPSAD)** and **CE Production (CEPROD)** paradata collected over the same time frame that aims to compare user experiences from both surveys.

All figures in this poster are sourced from complete cases, which come out to 3,213 users from OPSAD and 348 users from CE that entered at least one expense in the online diary.

The number of users for OPSAD and CEPROD may be lower depending on the research question being answered. For example, topics that examine something during the "fielding period" do not consider activity that occurred on the placement date for CEPROD, since CE protocols require the start of the fielding period to be the day after placement.

Research Questions

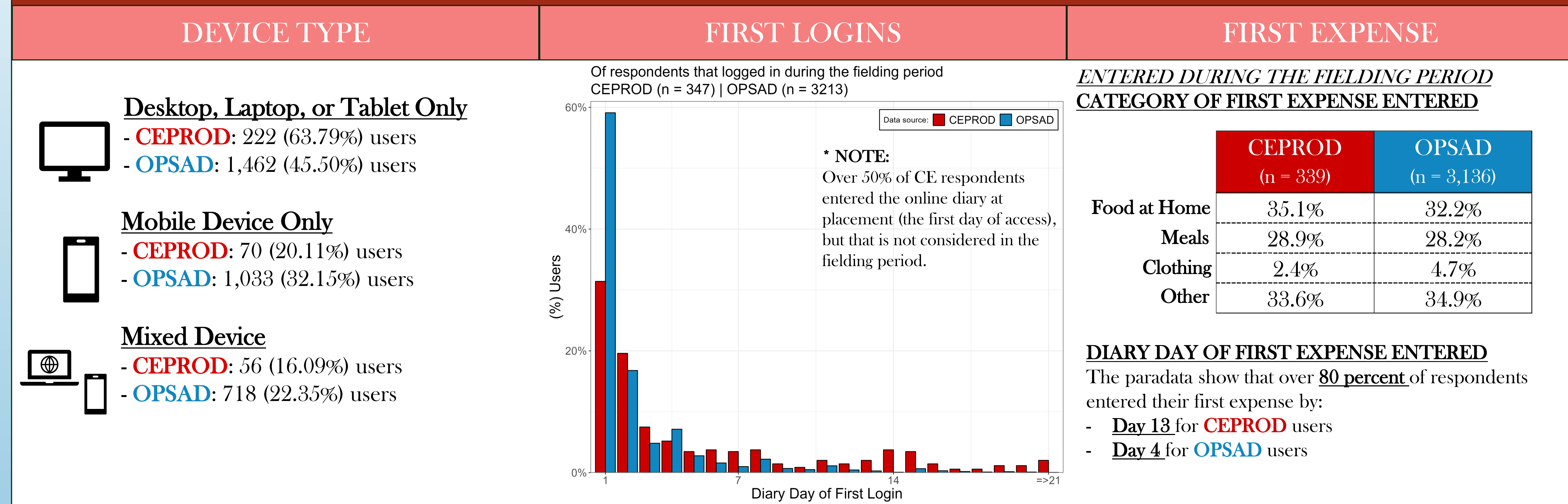
1. How and when did respondents access the diary?

- What kind of device did they use?
- When did respondents first log in during the fielding period?
- For respondents that entered expenses during the fielding period, what was the expense category (Food at Home, Meals Away from Home, Clothing, or Other) and the diary day of respondents' first expense?

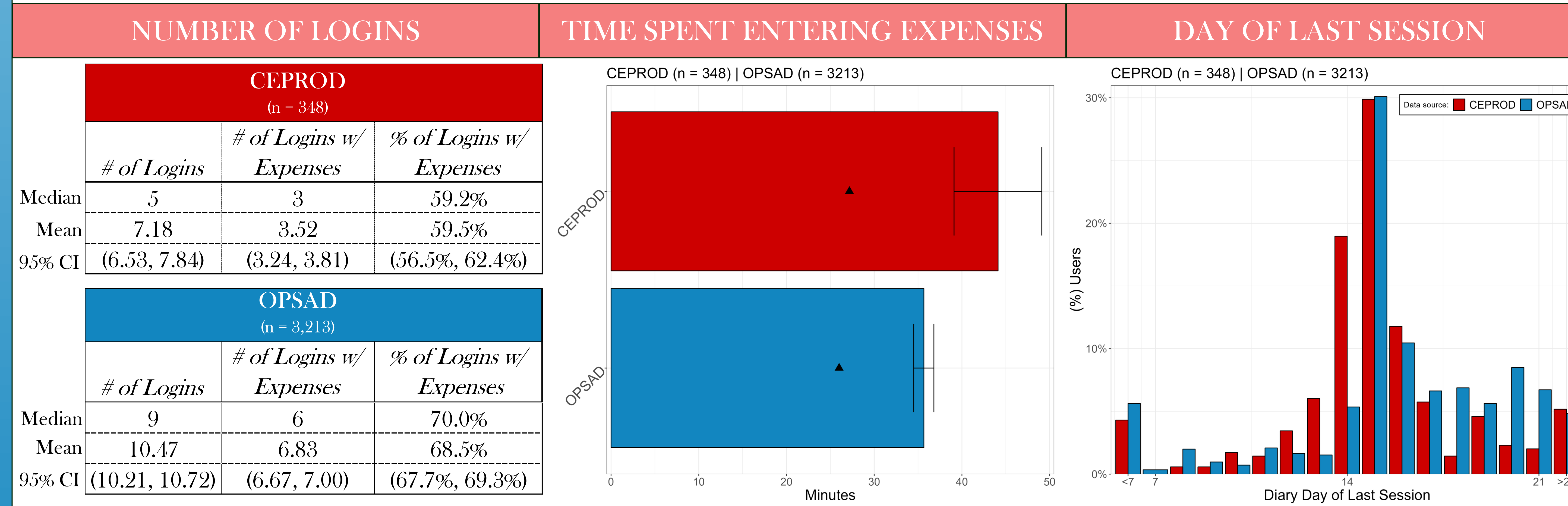
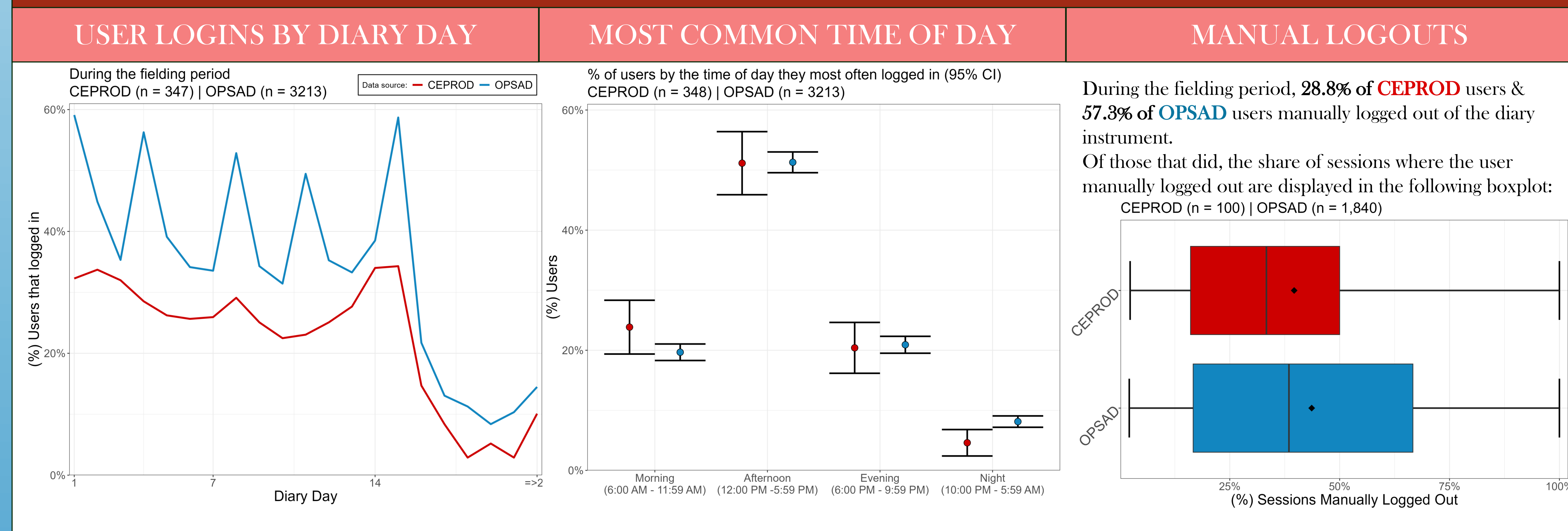
2. During data collection, how did respondents use the diary?

- How many users logged in each diary day of the fielding period?
- What was the most common time of day to enter the diary?
- How often did respondents manually log out during the fielding period?
- How many times did respondents log in?
- How much time did the respondents spend entering expenses?
- What was the last day of the fielding period that respondents logged in?

LOGGING IN TO THE DIARY



USING THE DIARY



CONCLUSIONS

Summary of Paradata Test

Key similarities:

- Similar rates of entering the diary when first accessible to them.
- Time of day respondents logged in most often.
- Those that manually logged out did so at similar rates throughout the fielding period.

Key differences:

- Device use patterns.
- Diary day of first expense entered.
- OPSAD users logged in more often and later into the fielding period.
- CEPROD users spent more time entering expenses in the instrument, despite logging in less.

Limitations of Paradata Test

1. UNEXPECTED BUGS & INCONSISTENCIES

Server issues and feature glitches impact the reliability of paradata in both surveys. For example, both instruments had issues with the automatic logout feature, either resulting in logout action timestamps that occur long after the last action observed, or do not appear at all, which complicated measuring how much time respondents spent in the diary.

2. DATA DESIGN & STRUCTURE

Differences in paradata structure makes comparisons difficult to perform.

OPSAD paradata are easy to analyze on the expense level but are not detailed enough to fully show the time spent in the diary outside of entering expenses, given the issues with the automatic logout feature. Consequently, it is not possible to measure total amount of time spent in the instrument outside of sessions where expenses were entered.

CE paradata have a detailed, but complicated structure with unclear action definitions and redundancies that make it difficult to establish clear patterns. Additionally, there are no expense-level identifiers, making expense level analysis either extremely difficult or potentially unfeasible.

Recommendations for Paradata Research

1. PLAN & DESIGN PARADATA

Clearly define what to measure using paradata beforehand, and plan structure of paradata with clearly defined actions to facilitate analyses. When designing the paradata structure, assume that glitches or bugs will happen. Ensure that the structure demonstrates clear action patterns to better spot errors and problem cases when they do pop up. If a project compares paradata from different sources, try to make the paradata structure more comparable.

2. EXAMINE PARADATA DURING TESTING

Create detailed test cases and rigorously test the instrument. If possible, examine paradata generated from said test cases to establish patterns in paradata, search for errors, and check for possible blind spots or redundancies.

