

Enhancing Data Quality Using Expenditure Records

Safia Abdirizak, Laura Erhard, Yezzi Angi Lee, Brett McBride
Bureau of Labor Statistics
American Association of Public Opinion Research
May 19, 2017



Outline

I. Background- CE Redesign (Gemini Project)

- ▶ Proof of Concept Test

II. Record Use

- ▶ How frequently did respondents consult records?
- ▶ Section-Level Expenditure Comparisons

III. Response Quality

- ▶ Data Quality Measurement
- ▶ Missing Data Rates
- ▶ Extent of Rounded Responses

IV. Summary

V. Next Steps



I. Background



CE Redesign (Gemini Project)

In 2009, the Bureau of Labor Statistics (BLS) Consumer Expenditure Survey (CE) initiated the multi-year Gemini Project for the purpose of researching, developing, and implementing an improved survey design.



Proof of Concept Test

- **Objective:** To assess the feasibility of the redesign by testing a single-sample expenditure survey design consisting of
 - ▶ two personal interviews,
 - ▶ personal diaries,
 - ▶ incentives,
 - ▶ record usage,
 - ▶ and technology usage (online diaries)
- Fielded from July 2015-October 2015 divided across four Census Regions: Atlanta, Chicago, New York, and Denver.
- Final number of 520 completed cases for a response rate of 50 percent.

Test Design

- **Visit 1:**

- ▶ Recall Interview

- **One week of diary keeping and records collection**

- **Visit 2:**

- ▶ Records Interview

- **Incentives including,**

- ▶ \$2 token incentive with the advanced mailing,
- ▶ \$20 after Visit 1, if respondent completed all sections of Recall Interview,
- ▶ \$20 for each eligible household member that entered at least 1 expenditure in diary,
- ▶ and \$20 debit card mailed to the respondent *after completing* the records interview with an *additional* \$20 debit card included if 1+ records were used.

Note: Our analysis compared the Proof of Concept Test sample (“test” group) with Matched Production sample responses to corresponding sections in the current survey (“control” group).

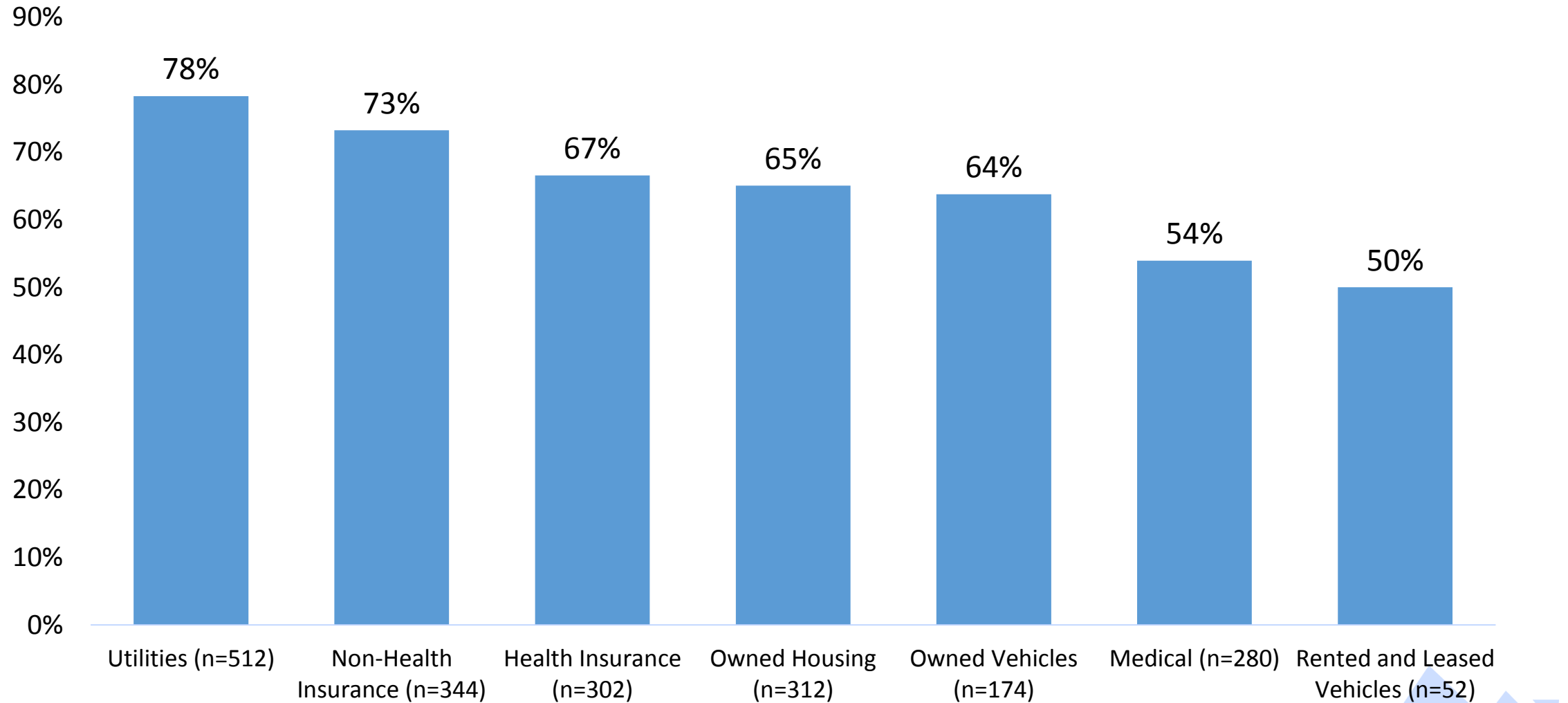


II. Record Use



How frequently did respondents consult records?

Percent Frequency of Record Use (among CUs with 1+ expenditure)



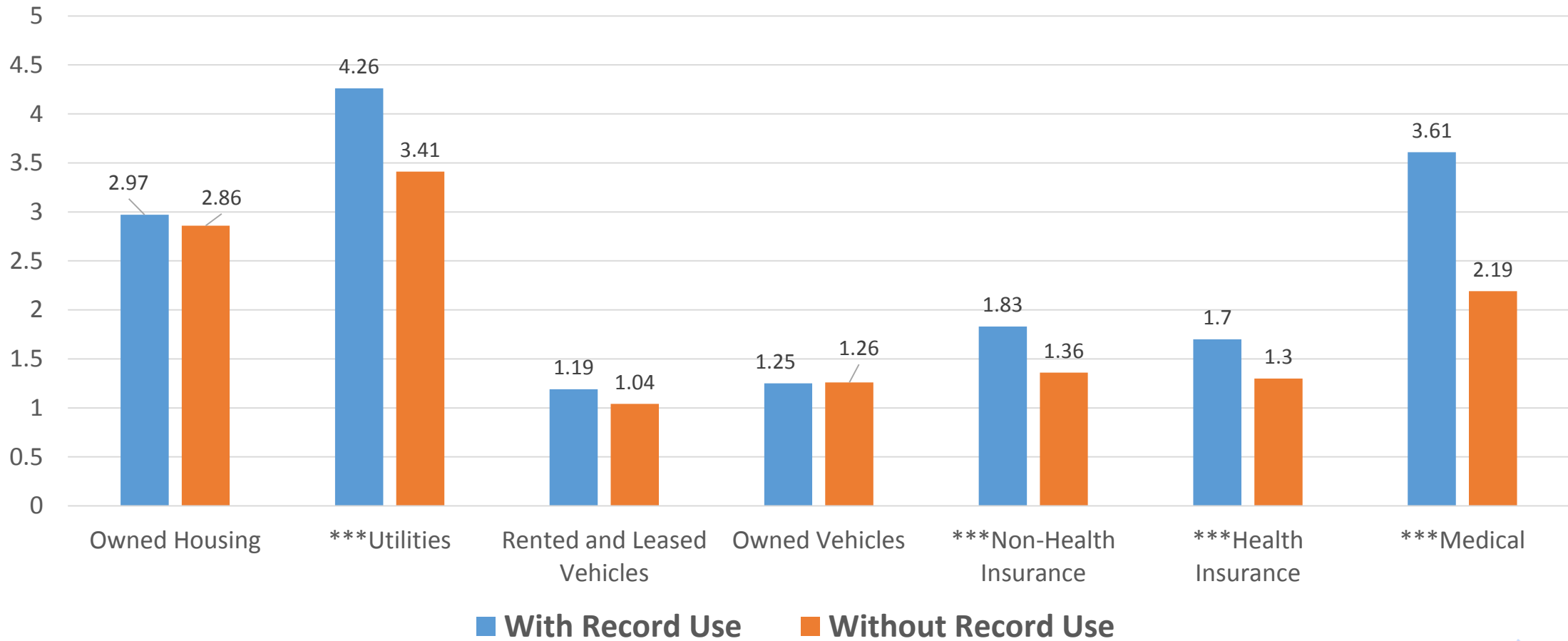
Section-Level Expenditure Comparisons

Did record use cue additional expenditure reports, within the test sample?

Average Count of Entries by Record Use

(among CUs with 1+ section expenditure)

Test



*p<.05; **p<.01; ***p<0.001

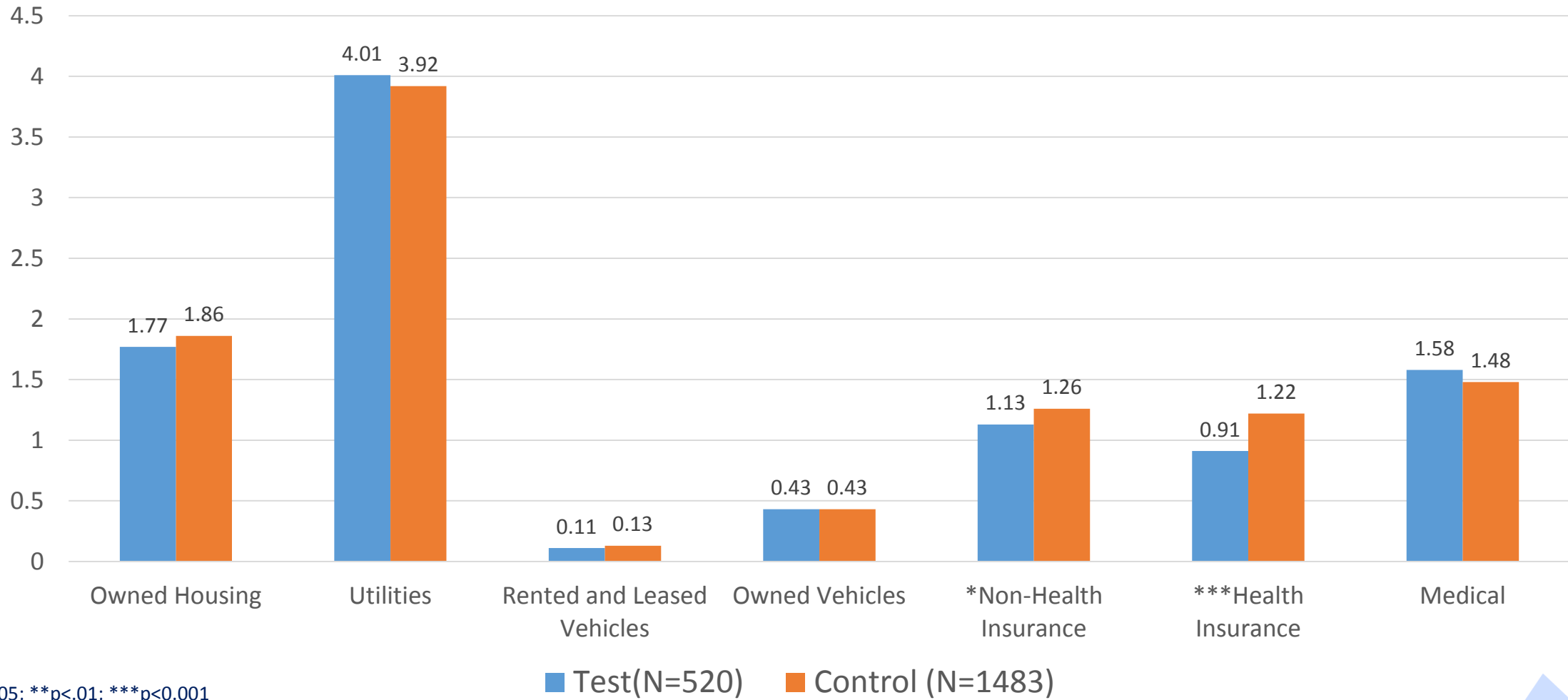


Section-Level Expenditure Comparisons

Did record use cue additional expenditure reports, overall by sample?



Average Count of Entries by Sample



*p<.05; **p<.01; ***p<0.001



Section Level Expenditure Comparisons

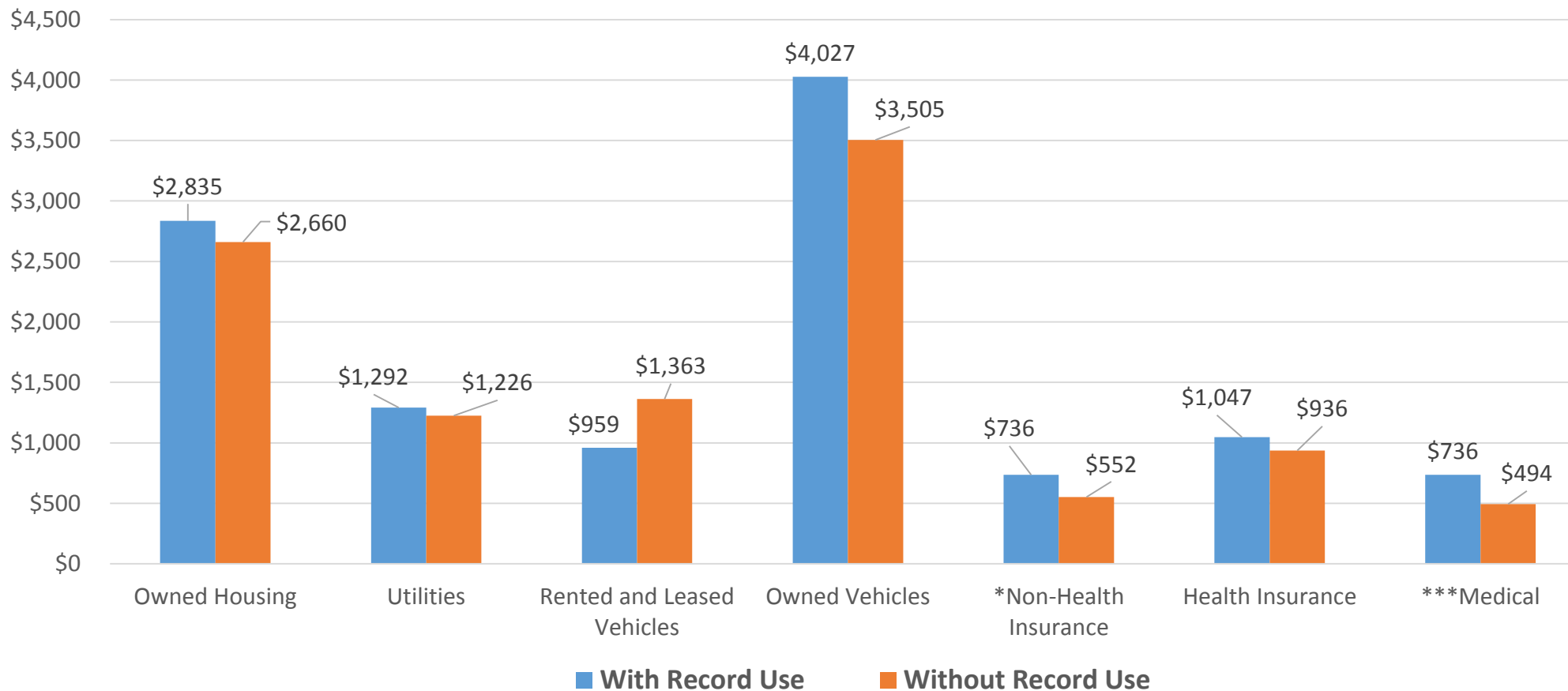
Did respondents report increased expenditure values when referencing records, within the test sample?



Average Expenditure Totals by Record Use

(among CUs with 1+ section expenditure)

Test



*p<.05; **p<.01; ***p<.001

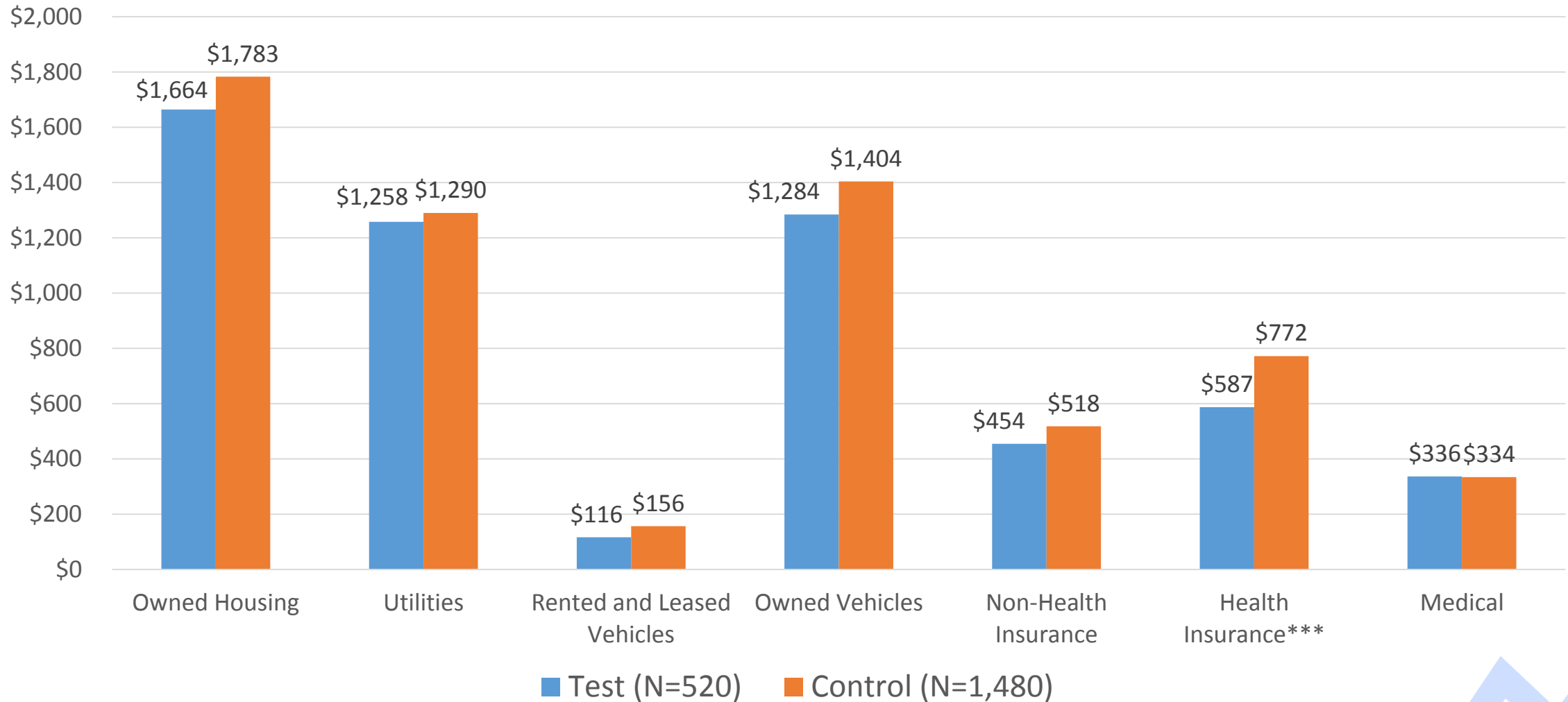


Section Level Expenditure Comparisons

Did respondents report increased expenditure values when referencing records, overall by sample?



Average Expenditure Totals by Sample



III. Response Quality



Data Quality Measurement

We explored the extent of rounding in the records interview and evaluated item non-response through analyses of respondent-provided invalid blanks in both the test and control samples.

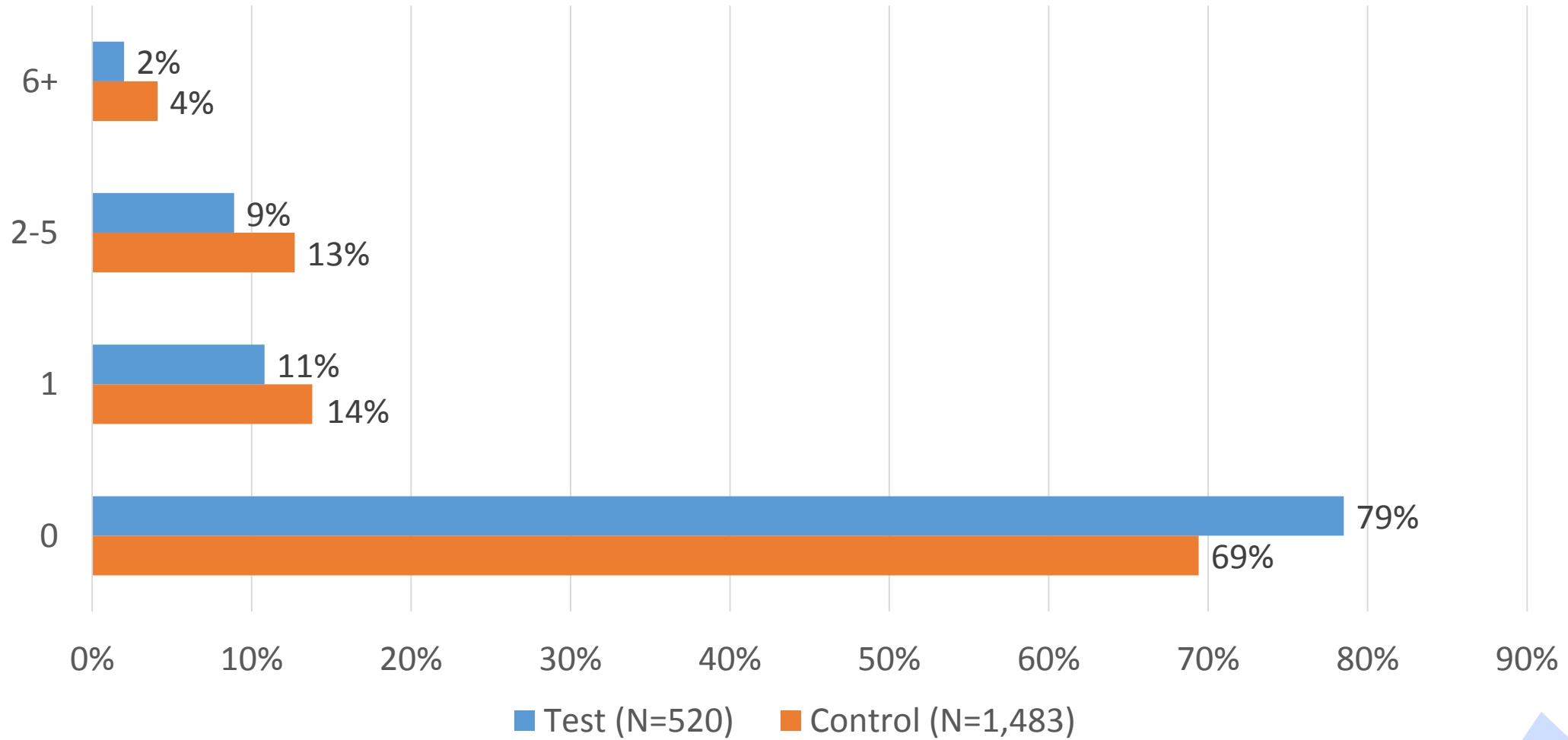


Missing Data

Was there less item non-response, in the form of don't knows?



Distribution of Number of Don't Knows by Sample

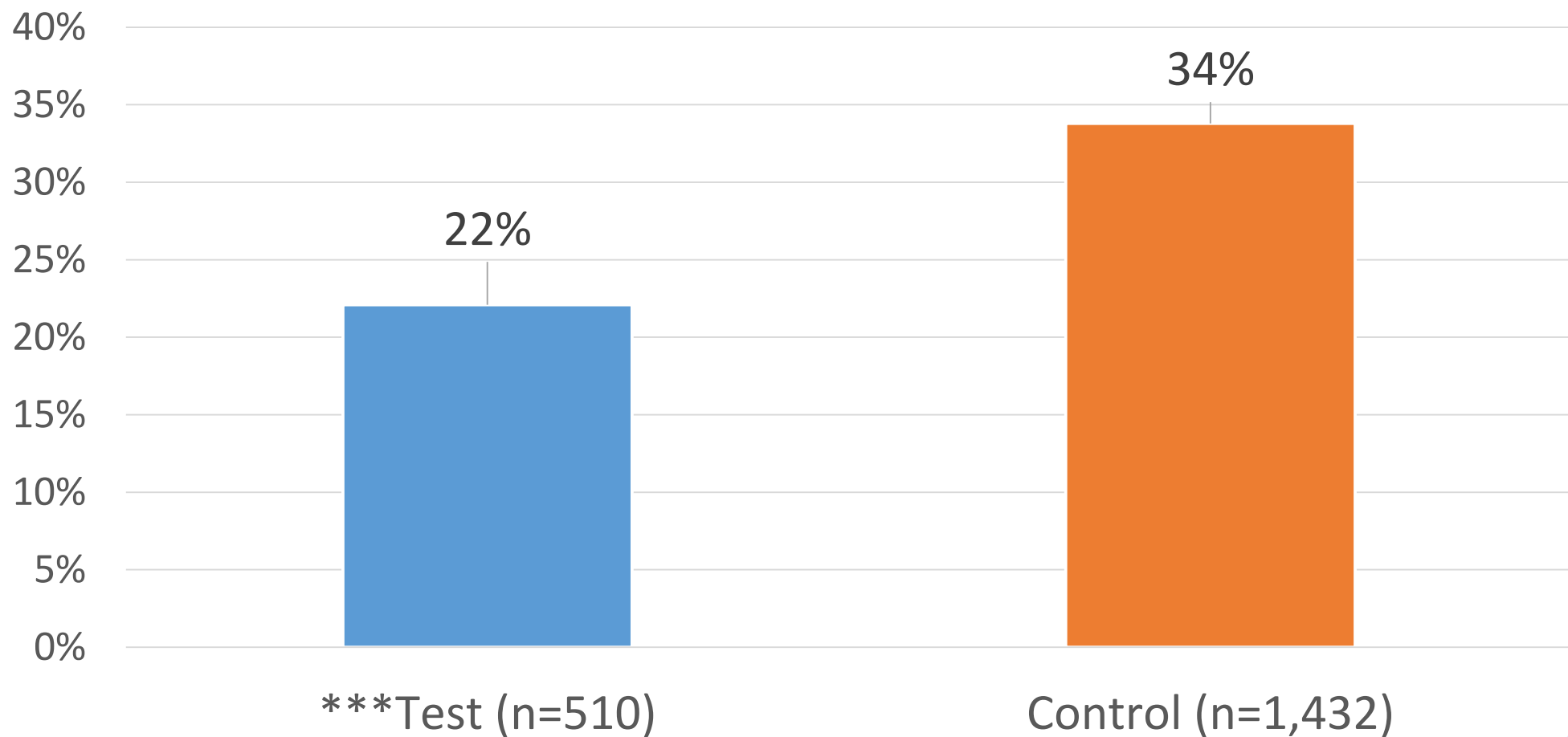


Extent of Rounded Responses

With the use of records, was there less evidence of rounding to multiples of \$25?



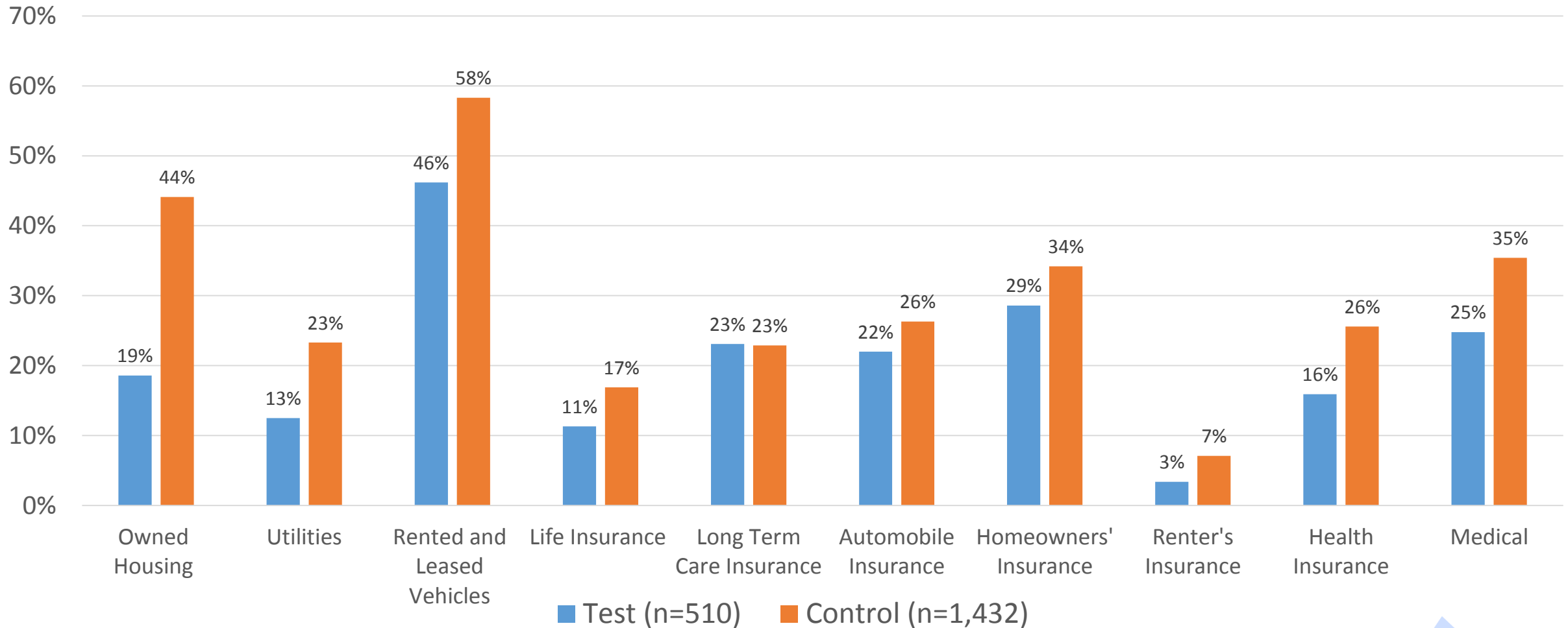
Average Percent of Rounded Expenditure Responses



*p<.05; **p<.01; ***p<.001

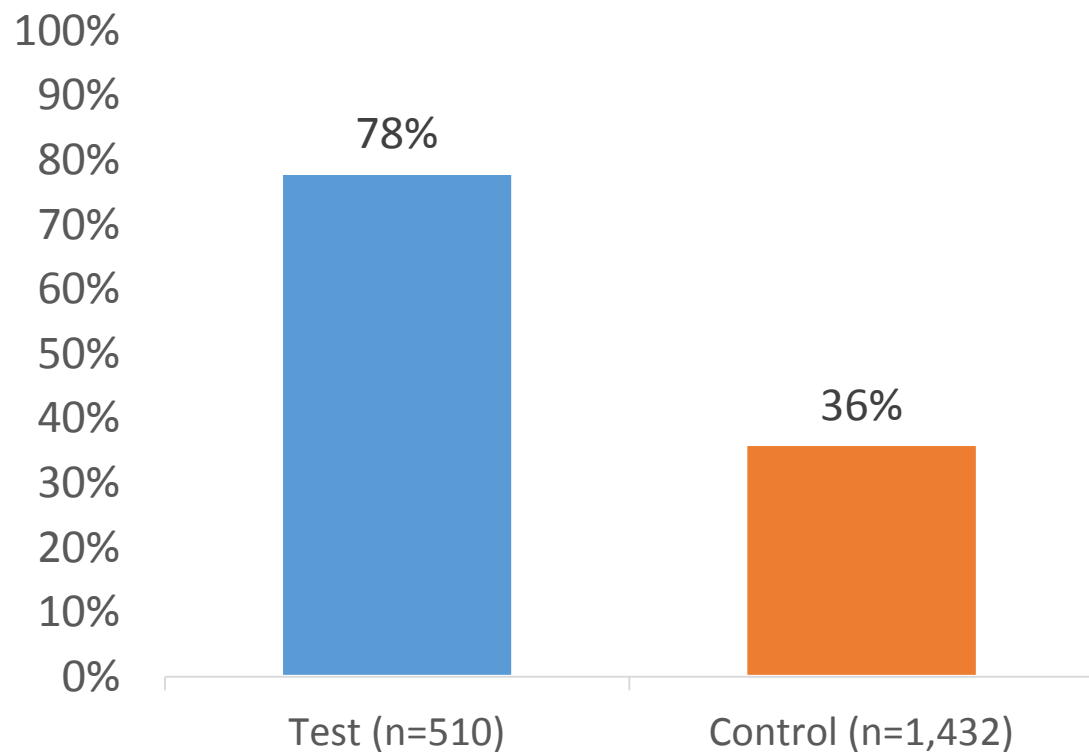


Average Percent of Rounded Expenditure Responses by Question

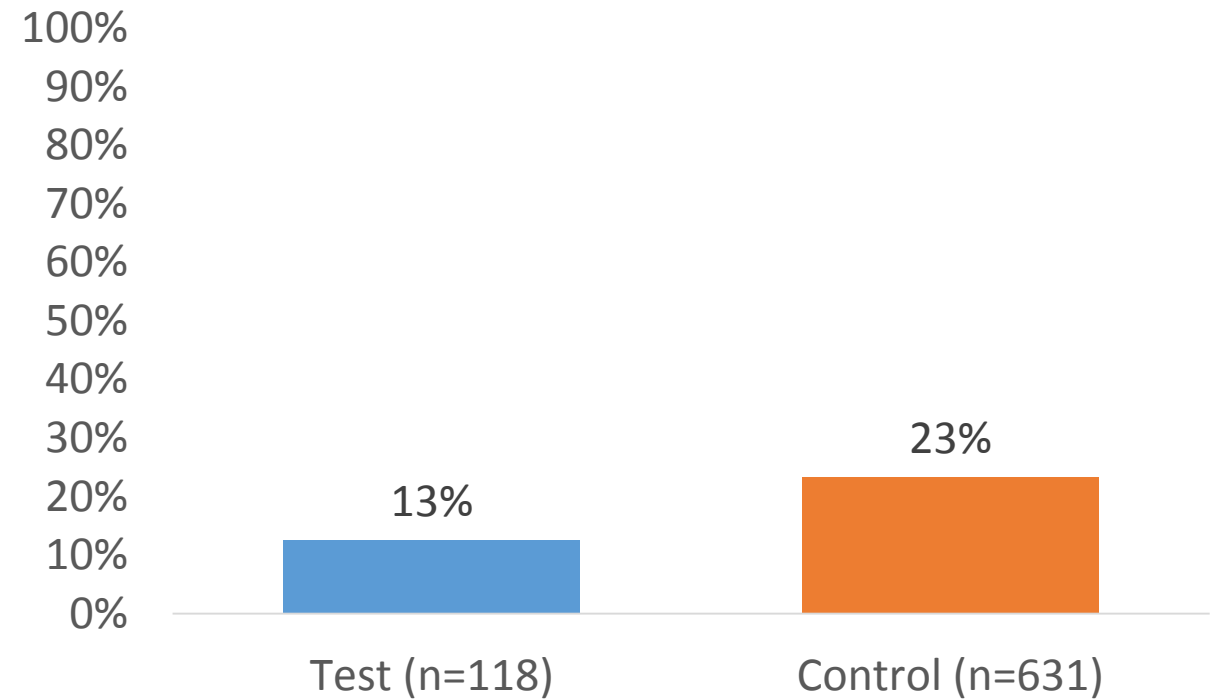


Record Use and Rounding - Utilities Example

Percentage Frequency of Record Use
(among CUs with 1+ Section Expenditures)



Comparison of CU average percent of rounded expenditures



IV. Summary



Summary

- 1. Did record use cue additional expenditure reports?**
 - Yes, all sections (except Owned Vehicles) had higher entry averages when CUs used records.
- 2. Did respondents report increased expenditure values when referencing records?**
 - Though not consistently significant, record use was associated with increases in expenditure values in both the control and test sample.
- 3. Was there less item non-response?**
 - Yes, respondents in the test sample provided fewer missing values for expenditure variables than respondents in the control sample.
- 4. Was there less evidence of rounding to multiples of \$25?**
 - Yes, for questions in the records interview the test sample had higher record use with lower levels of rounding compared to the control sample.



V. Next Steps



Next Steps

- Explore incentives data once available
- Further investigate section level expenditure comparisons
- Continued development of records interview protocol
- Large Scale Feasibility Test (tentatively 2020)
 - ▶ Larger sample sizes (2,000 completes)



Contact Information

Safia Abdirizak

Economist

Division of Consumer Expenditure Surveys

www.bls.gov/cex

202-691-5137

abdirizak.safia@bls.gov

