

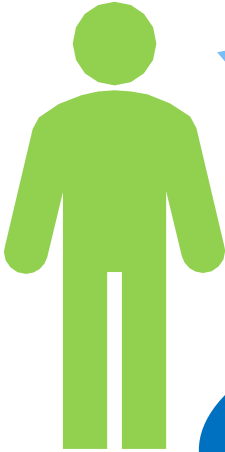


# Studying Changes in an Ongoing In-Person Survey Using a Web Survey

Stephanie Eckman

Research supported by ASA / NSF / BLS Fellowship





Has anyone in this household had any expenses for:  
Oil change, lubrication or oil filter?

In what month?  
What was the total cost?  
Did this expense include labor?

Has anyone in this household had any expenses for:  
Tire purchases or mounting?

Yes

May  
\$49.99  
No

Yes?



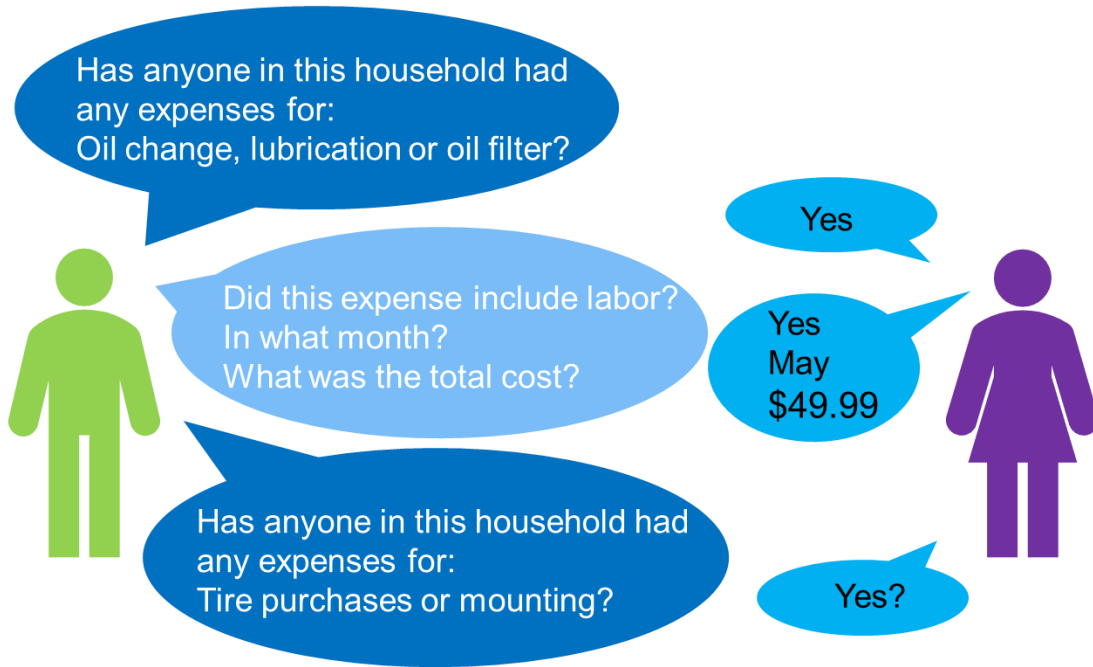
# National Academy Review

- This sequence is repeated dozens of times during each interview and may affect respondent behavior.
- It seems likely that respondents **learn quickly**...that the interview will last longer if they answer “yes” to these screening questions.
- Fifty percent of field representatives said that it happened frequently or very frequently.

Do Rs engage in **motivated misreporting**  
in **Wave 1** of the CE?

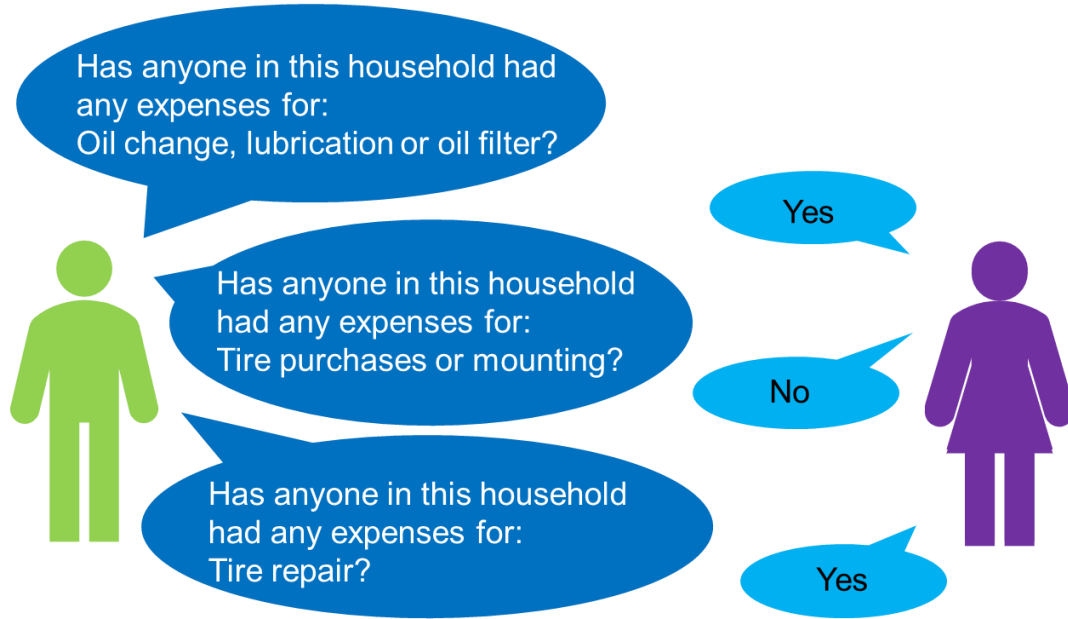


# Interleafed Format



- Filter
  - Follow up
  - Follow up
- Filter
  - Follow up
  - Follow up
- Filter
  - Follow up
  - Follow up

# Grouped Format



- Filter
- Filter
- Filter
- Filter
  - Follow up
  - Follow up
  - Follow up
  - Follow up
  - Follow up

# Previous research on MM

- CE-like questions:

- Kreuter et al (2011) The Effects of Asking Filter Questions in Interleaved versus Grouped Format. SMR
- Eckman et al (2014) Assessing the Mechanisms of Misreporting to Filter Questions in Surveys. POQ
- Daikeler et al (2019) Motivated Misreporting in Smartphone Surveys. SSCR
- Bach et al (2019) Misreporting Among Reluctant Respondents. JSSAM

**Full list:** [http://www.stepheckman.com/project/motivated\\_misreporting](http://www.stepheckman.com/project/motivated_misreporting)

- CE: underreporting in later waves

- Bosley et al (1999), Shields & To (2005), Yan & Copeland (2010), Bach & Eckman (2019)


- CE: MM in Wave 1

- McBride (2013)

**Barring experiment within CE,  
how can we estimate underreporting?**

# How to Measure MM in CE

- CE asks filters only in interleaved format
- Web survey in both formats

	X	Interleaved	Grouped
<b>Consumer Expenditure</b>	•	•	
<b>Nonprob</b>	•	•	
<b>Web Survey</b>	•		•



# Similar Imputation Studies

- Rendall et al (2013)
- Powers et al (2005)  
Christensen et al (2006)  
Kolenikov & Kennedy (2014)  
Park et al (2016)
- Kim et al (2020)

# Similar Imputation Studies

- **Rendall et al (2013)**
- Powers et al (2005)  
Christensen et al (2006)  
Kolenikov & Kennedy (2014)  
Park et al (2016)
- Kim et al (2020)

Survey	$X_1$	$X_2$	Y
Survey 1	●	○	●
Survey 2	●	●	●

Model  $Y|X_1, X_2$

Variables are *jointly observed*

*Cross-Survey imputation*

# Similar Imputation Studies

- Rendall et al (2013)
- **Powers et al (2005)**  
**Christensen et al (2006)**  
**Kolenikov & Kennedy (2014)**  
**Park et al (2016)**
- Kim et al (2020)

Mode	X	$Y_{\text{web}}$	$Y_{\text{phone}}$
Web	●	●	
Phone	●	○	●

Impute responses in *better* mode

# Similar Imputation Studies


- Rendall et al (2013)
- Powers et al (2005)  
Christensen et al (2006)  
Kolenikov & Kennedy (2014)  
Park et al (2016)
- **Kim et al (2020)**

Survey	X	Y
Probability	•	○
Nonprobability	•	•

Impute variable of interest on  
probability survey

# How to Measure MM in CE

- Eckman (2020) design:

	X	Interleafed	Grouped
<b>Consumer Expenditure</b>	•	•	
<b>Nonprob Web Survey</b>	•	•	•

# Web Survey Questionnaire

- Section 1: Demographics
- Section 2: Household characteristics

Section	# Filters	# Follow Ups
Utilities	5	4
Clothing	6	5
Insurance	5	5

} Order of sections randomized

} Order within sections forwards/backwards

- Section 6: Income

# Imputation Approach

- Multiple imputation of Yes/No to filter question
  - 6 models
  - 5 predictions from each model
- Predictors:
  - Format (grouped or interleaved)
  - Survey (CE or web)
  - Item
  - Section order
  - Respondent & HH characteristics

Pseudo-R2: 0.20-0.22

AUC: 0.80

# Compare Background Variables

HH members\*

HU has pool

HU has off-street parking

HU has porch

HU has apartment

HU has central air

HU has window air

HU has solar panels

HU single family

HU bedrooms

HU bathrooms

HH has wage income

HH has self-employ. inc.

HH income

R owns home

R age

R white

R black

R asian

R married

R never married

R hispanic

R education

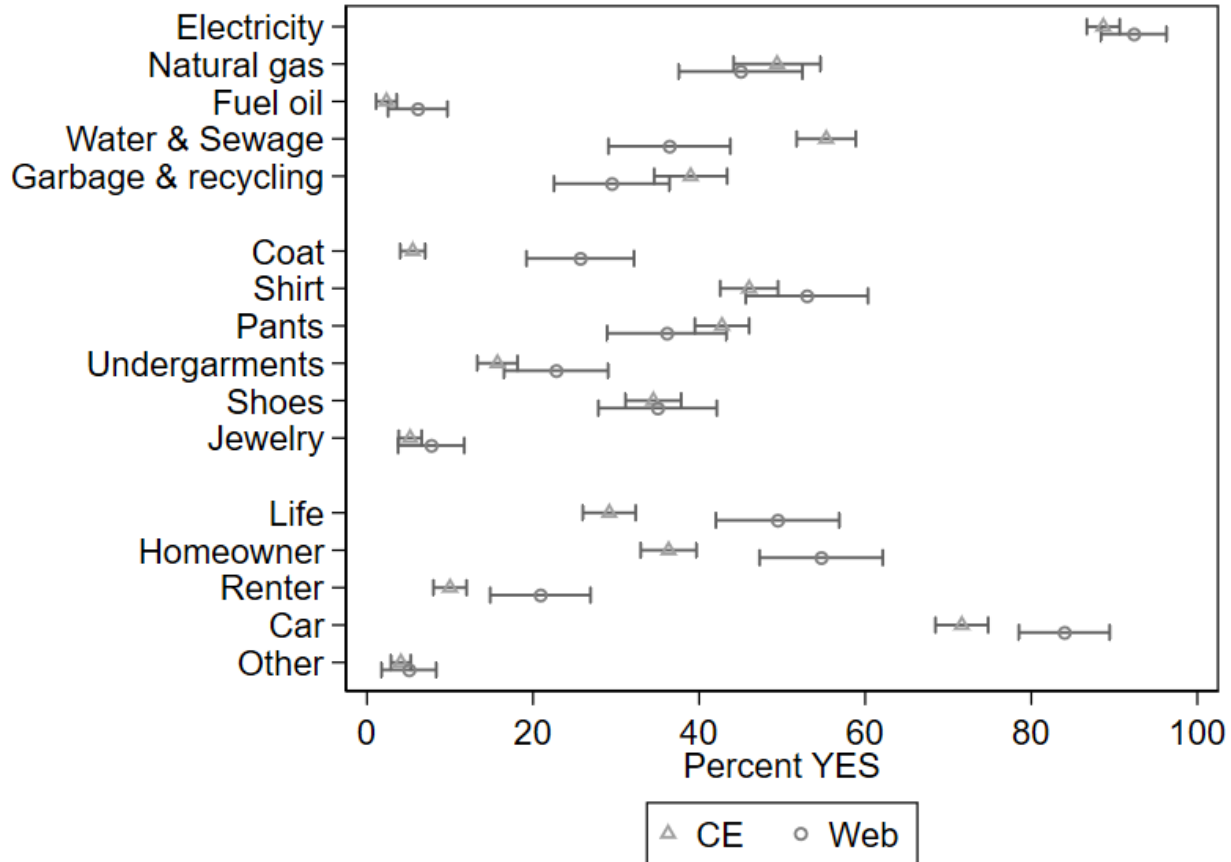
23 variables

20 with significantly different distributions  
CE vs Web

Good representation in each category



# Compare Purchases



Only web survey respondents in interleaved format included

# Results

	Interleafed % Yes	Grouped % Yes	Difference % points
<b>Consumer Expenditure</b>	33.5	38.5	5.1 %points (1.5)
<b>Web Survey</b>	36.8	42.9	6.1 %points (0.80)

---

Utilities: 0.2 %points  
Clothing: 9.9 %points\*  
Insurance: 5.1 %points\*

# Take Aways

- Evidence that motivated misreporting occurs in Wave 1
- Effect of misreporting on CPI unknown
- Imputation approach needs more theoretical support
- Approach could be used to assess hypothetical changes to ongoing surveys when experimentation not possible



# Thank You

**Stephanie Eckman**

Fellow, Survey Research

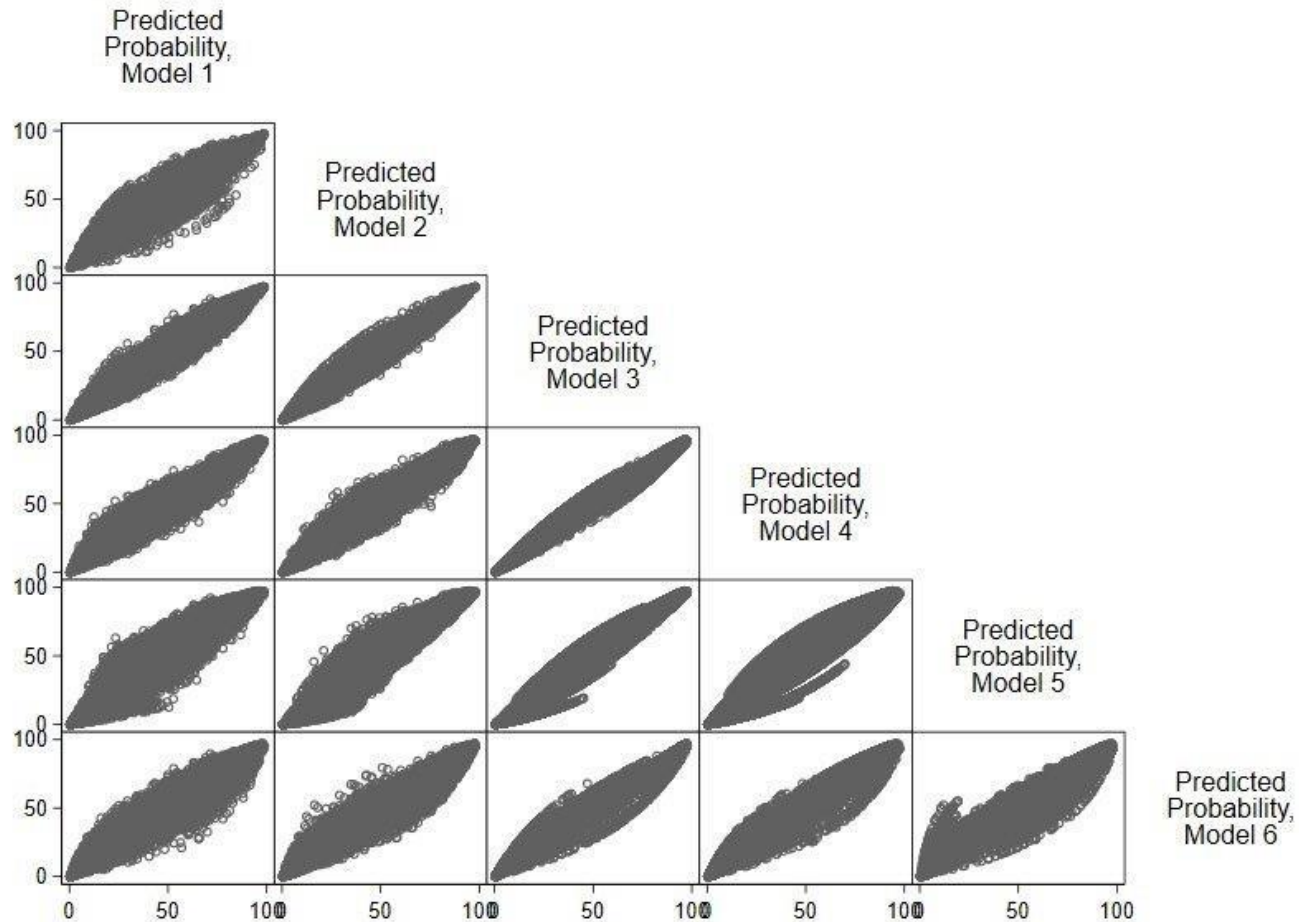
RTI International

[seckman@rti.org](mailto:seckman@rti.org)

<http://stepheckman.com>

@stephnie

# Comparison of Imputation Models



Variable	Model					
	1	2	3	4	5	6
Grouped (vs. Interleafed)	a	x	x	a	x	x
CE (vs. web)	x	a	x	x	a	x
Item (1-16)	x	x	x	b	b	a
Section Order	x	x	x	x	x	x
HH members*	b	b	x	x	x	b
HU has pool	b	b	x	x	x	b
HU has off-street parking	b	b	x	x	x	b
HU has porch	b	b	x	x	x	b
HU has apartment	b	b	x	x	x	b
HU has central air	b	b	x	x	x	b
HU has window air	b	b	x	x	x	b
HU has solar panels	b	b	x	x	x	b
HU single family	b	x	x	x	x	x
HU bedrooms	b	b	x	x	x	b
HU bathrooms	b	b	x	x	x	b
HH has wage income	b	b	x	x	x	b
HH has self-employ. inc.	b	x	x	x	x	x
HH income	b	x	x	x	x	x
R owns home	b	b	x	x	x	b
R age	b	b	x	x	x	b
R white	x	x	x	x	x	x
R black	x	x	x	x	x	x
R asian	x	x	x	x	x	x
R married	b	b	x	x	x	b
R never married	b	b	x	x	x	b
R hispanic	b	b	x	x	x	b
R education	b	b	x	x	x	b

Case base for each model is 50,848 filters

Each *a* variable interacted with each *b* variable