

Towards Determining an Optimal Contact Attempt Threshold for a Large-Scale Personal Visit Survey

Adam Safir & Lucilla Tan

Division of Consumer Expenditure Survey
Bureau of Labor Statistics

70th Annual AAPOR Conference

May 16, 2015



Overview

■ Topics

- ▶ Motivation
- ▶ Methodology
- ▶ Findings
- ▶ Limitations
- ▶ Recommendations

■ Acknowledgements

- ▶ Jeffrey M. Gonzalez (Office of Survey Methods Research) & Brett McBride (Division of Consumer Expenditure Survey).

■ Disclaimer

- ▶ The opinions expressed in this paper are those of the authors and do not reflect official policy of the Bureau of Labor Statistics.

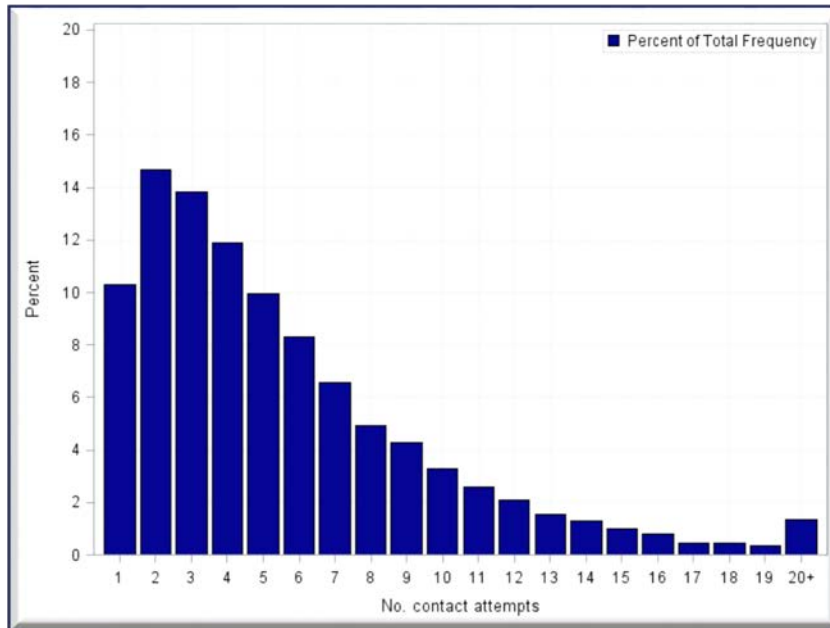
Motivation

- Conventional wisdom:
 - ▶ Maximize response rates to improve data quality
 - ▶ But, this interest is most connected to sampling theory, i.e., allowing for the computation of CIs and *sampling error*
- From a data quality standpoint, maximizing RRs leads to:
 - ▶ Benefits, reducing *nonresponse error* and *measurement error*
 - ▶ Costs, cases on the right hand side of the contact attempt distribution are relatively expensive
 - ▶ Max (ROI) → Max (benefit, data quality), Min (cost, effort)
- *Safir and Tan (2009)* recommended a threshold of 7 contact attempts in the trade-off between data quality and cost
 - ▶ The current study is a retrospective analysis w/more recent data, re-examining the 2009 recommendation prior to field testing

Contact Attempt Distribution

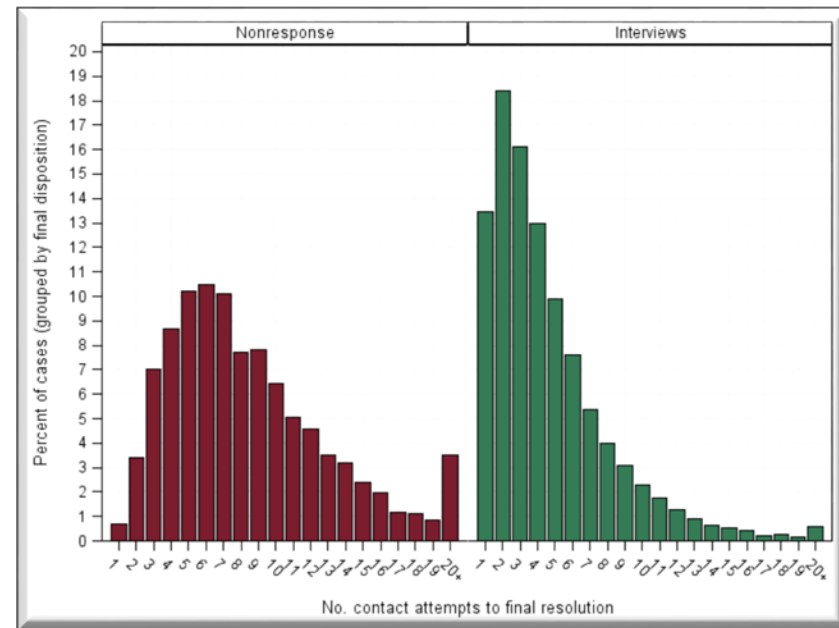
Overall

- A total of 100,775 contact attempts
- mode=2, median=4

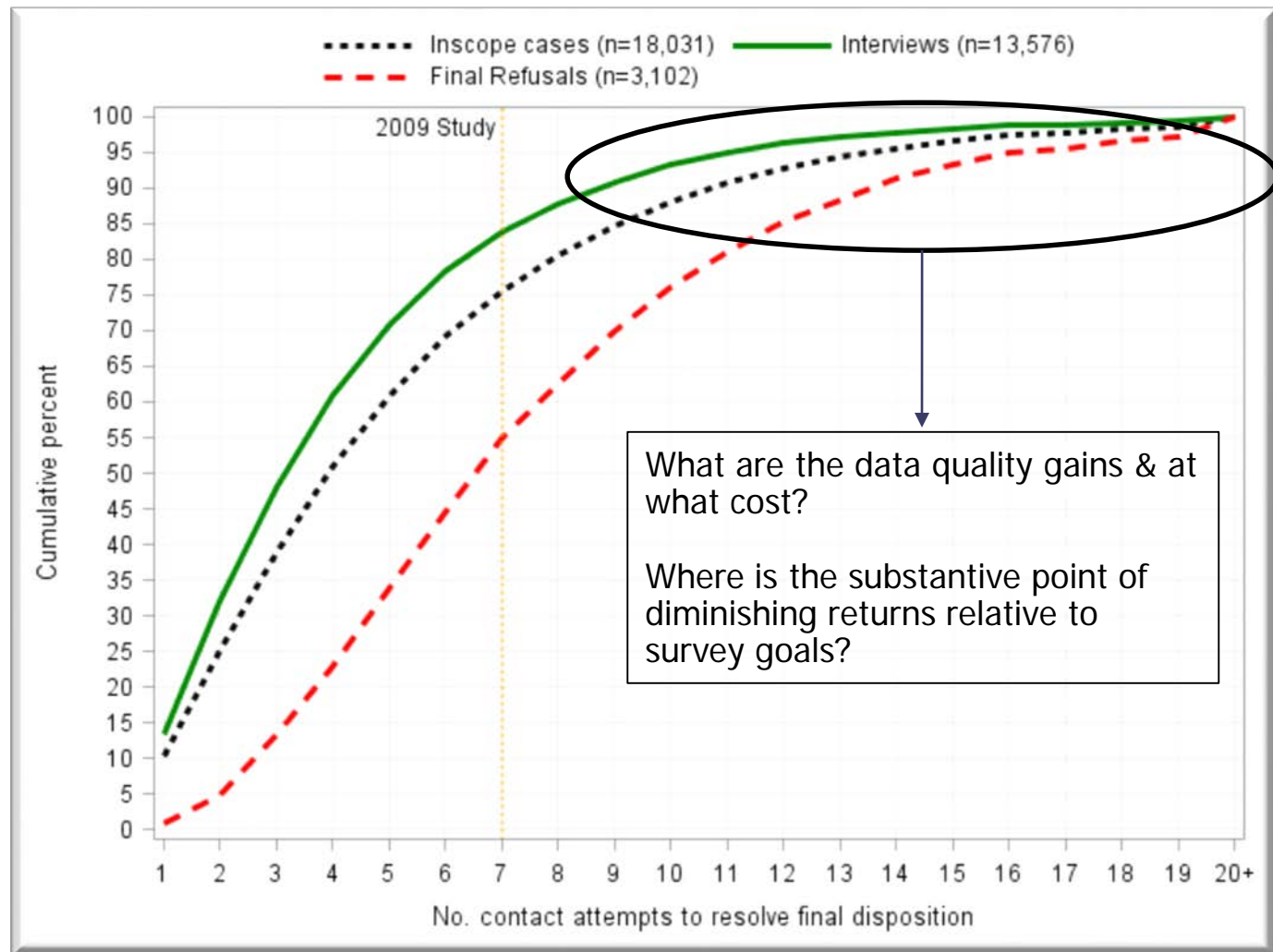


By final disposition

- **Nonresponse** mode=6, median=7
- **Interviews**: mode=2, median=4



Cumulative Resolution by Contact Attempt



Methodology

■ Sample

- ▶ Consumer Expenditure Interview Survey (more info: <http://www.bls.gov/cex/>)
- ▶ Data collected from April 2012 through March 2014
- ▶ Limited to sample units eligible for Wave 1 survey, with at least 1 contact reported in the sample unit's contact attempt history (n=18,031)

■ Comparison Groups

- ▶ Group 1: 1-7 contact attempts (n=13,631; interviewed cases=11,370)
- ▶ Group 2: 8+ contact attempts (n=4,400; interviewed cases=2,197)

■ Evaluation Measures

- ▶ Response rates
- ▶ Cost/effort (no. attempts, visit attempts, multiple interviewers)
- ▶ Sample characteristics
- ▶ Doorstep concerns
- ▶ Response composition (Subgroups RRs; R-indicators)
- ▶ Reporting quality (recall aids; endorsement rate of filter Qs)
- ▶ Expenditure reporting

Contact History Instrument (CHI)

CHI

♦ **CONCERN / BEHAVIOR / RELUCTANCE**

♦ Select the categories that describe respondent concerns, behaviors, or reluctance during this contact attempt.

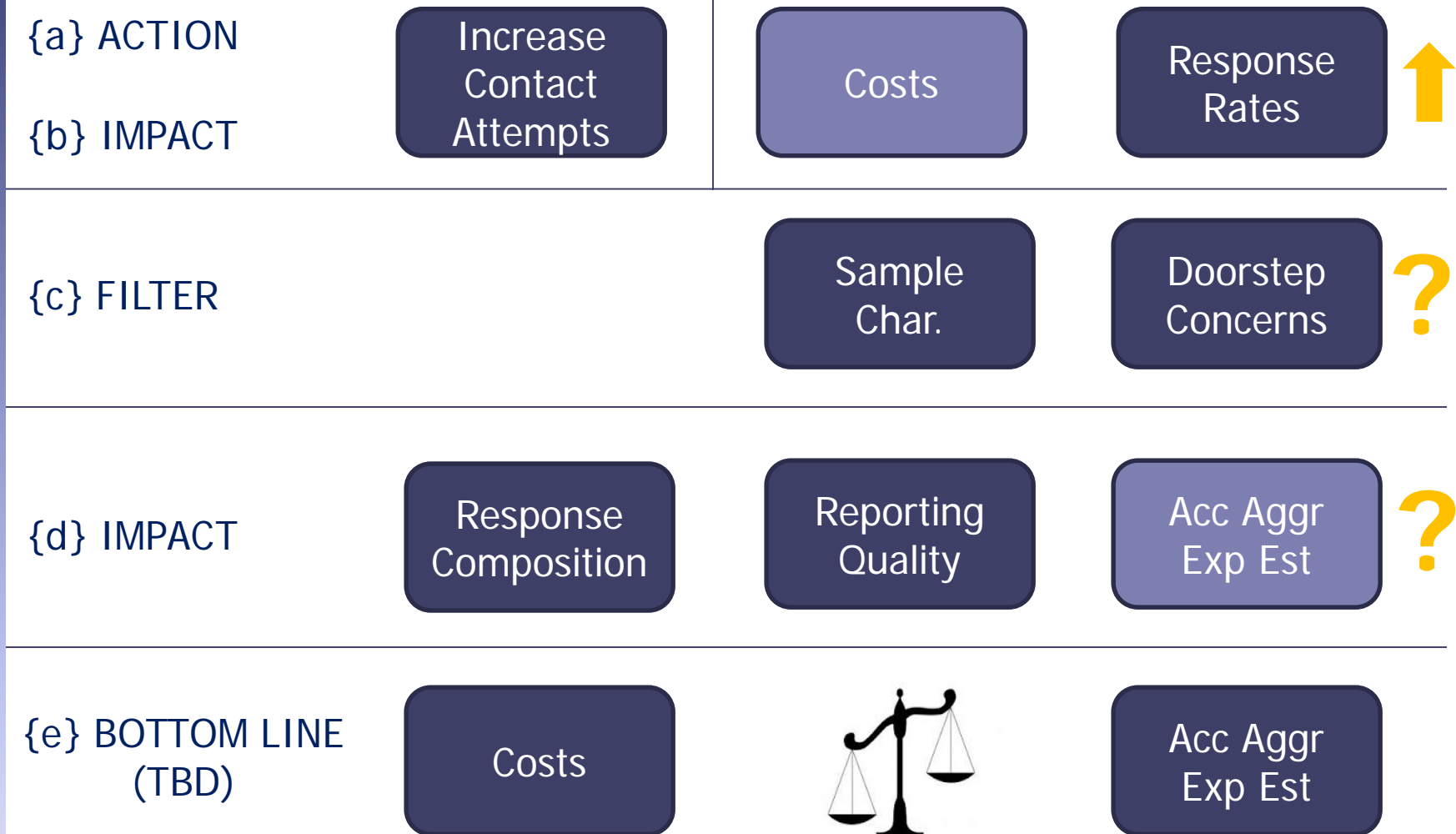
♦ Enter all that apply, separate with commas.

<input type="checkbox"/> 1. Not interested / Does not want to be bothered	<input type="checkbox"/> 12. Hostile or threatens FR
<input type="checkbox"/> 2. Too busy	<input type="checkbox"/> 13. Other household members tell respondent not to participate
<input type="checkbox"/> 3. Interview takes too much time	<input type="checkbox"/> 14. Talk only to specific household member
<input type="checkbox"/> 4. Breaks appointments (puts off FR indefinitely)	<input type="checkbox"/> 15. Family issues
<input type="checkbox"/> 5. Scheduling difficulties	<input type="checkbox"/> 16. Respondent requests same FR as last time
<input type="checkbox"/> 6. Survey is voluntary	<input type="checkbox"/> 17. Gave that information last time
<input type="checkbox"/> 7. Privacy concerns	<input type="checkbox"/> 18. Asked too many personal questions last time
<input type="checkbox"/> 8. Anti-government concerns	<input type="checkbox"/> 19. Too many interviews
<input type="checkbox"/> 9. Does not understand survey / Asks questions about the survey	<input type="checkbox"/> 20. Last interview took too long
<input type="checkbox"/> 10. Survey content does not apply (retired, healthy, no crimes to report)	<input type="checkbox"/> 21. Intends to quit survey
<input type="checkbox"/> 11. Hang-up / slams door on FR	<input type="checkbox"/> 22. No concerns
	<input type="checkbox"/> 23. Other - specify

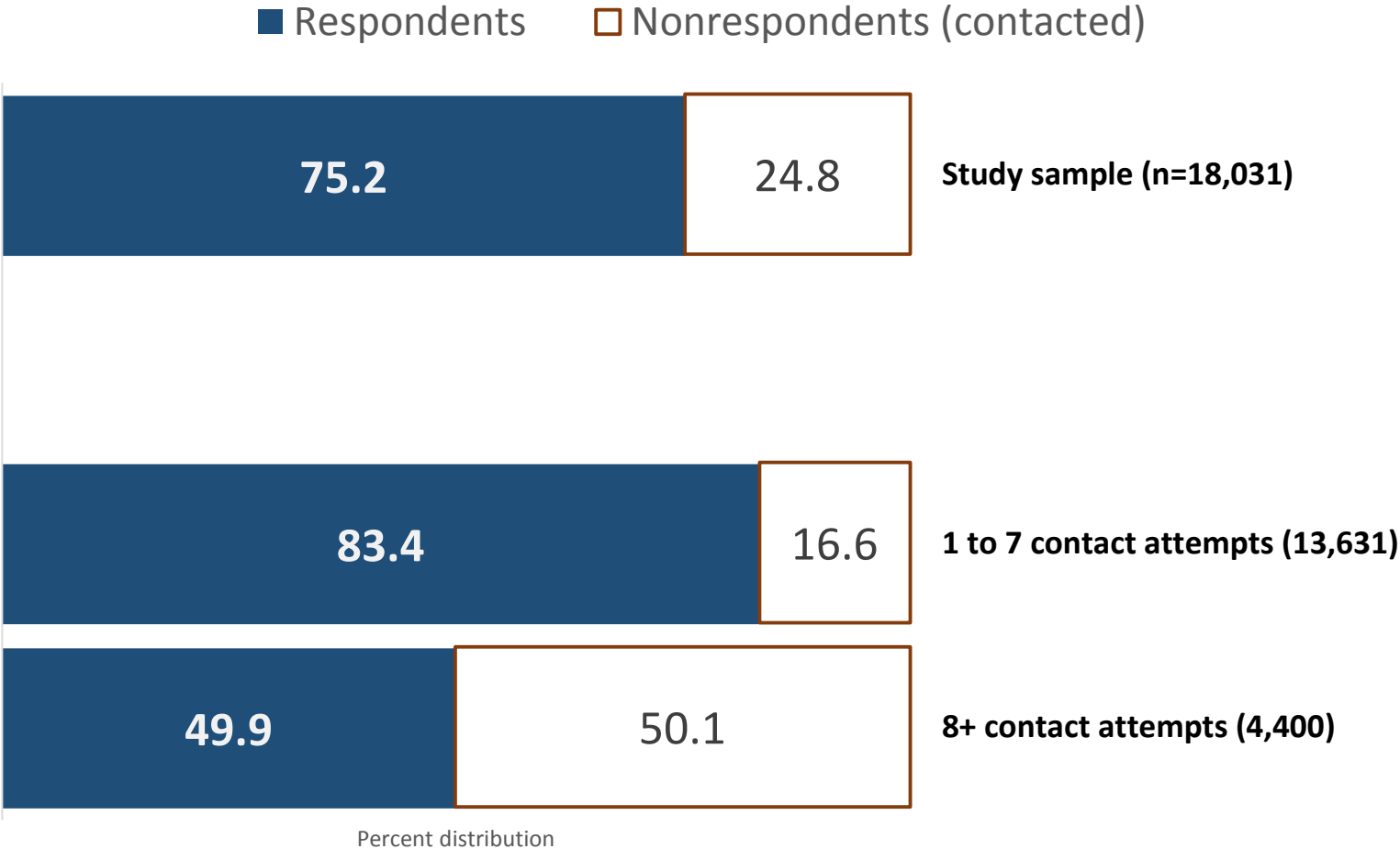


ID # of CHI doorstep concern items grouped to form theme	Doorstep concern theme (used in analysis)
1, 11, 12	1. Not interested / hostility
2, 3, 4, 5	2. Time
6, 7, 8, 9, 10	3. Survey content / privacy
13, 14, 15	4. Gatekeeping
16, 17, 18, 19, 20, 21, 23	5. Other / prior wave

Assessing the Impact of Increased Contact Attempts



Finding 1a: Response Rates

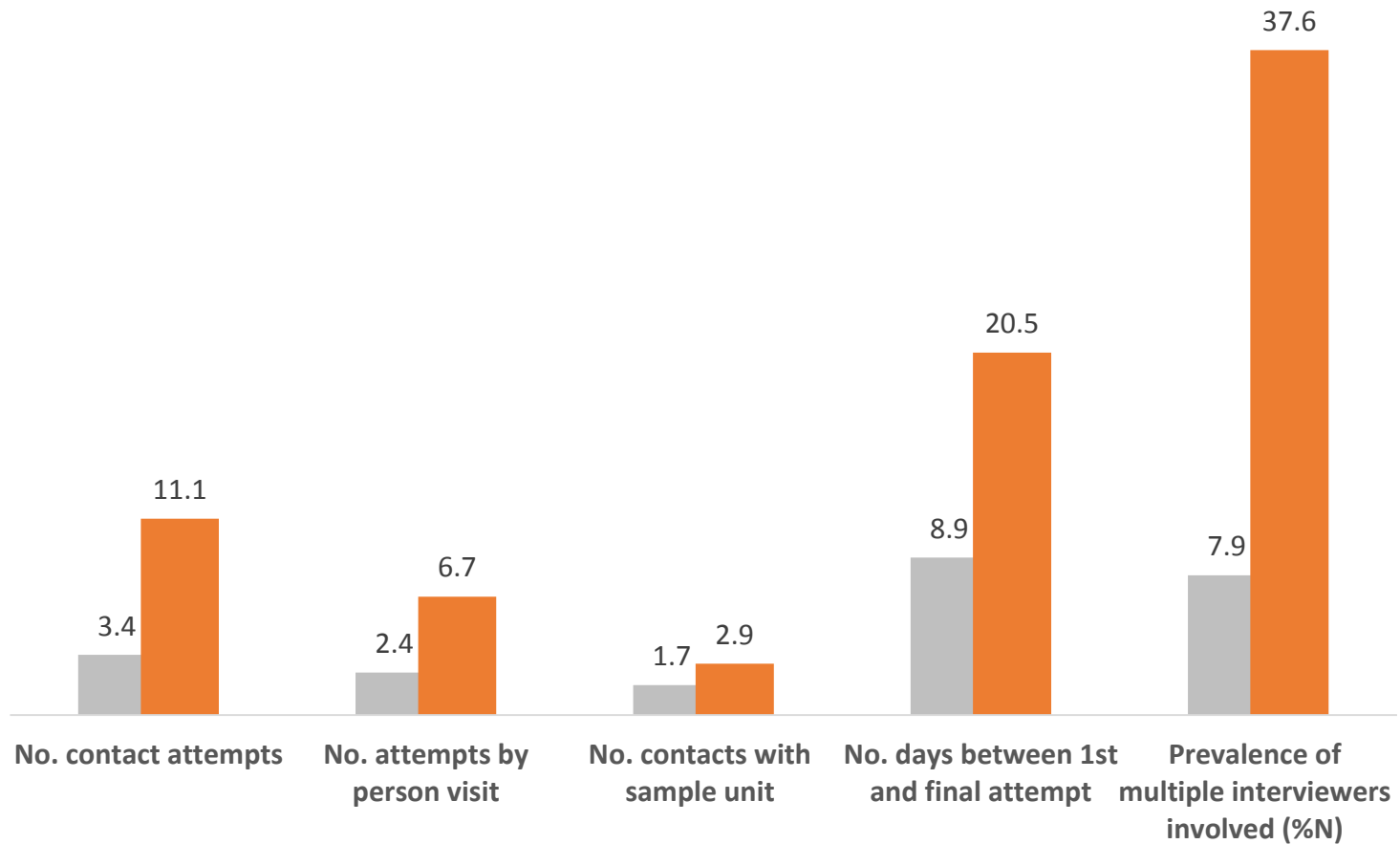


Finding 1b: Response Rates

- By the 7th contact attempt:
 - 78 percent of the sample resolved
 - 84 percent of the interviews completed
- Response rate of:
 - 83.4 percent for the 1-7 contact attempt group
 - 49.9 percent for the 8+ group
- Cumulative rate of in-scope cases resolved as interviews:
 - **63.1 percent after the 7th contact attempt**
 - **75.2 percent after 20+ contact attempts**

Finding 2. Cost (Effort)

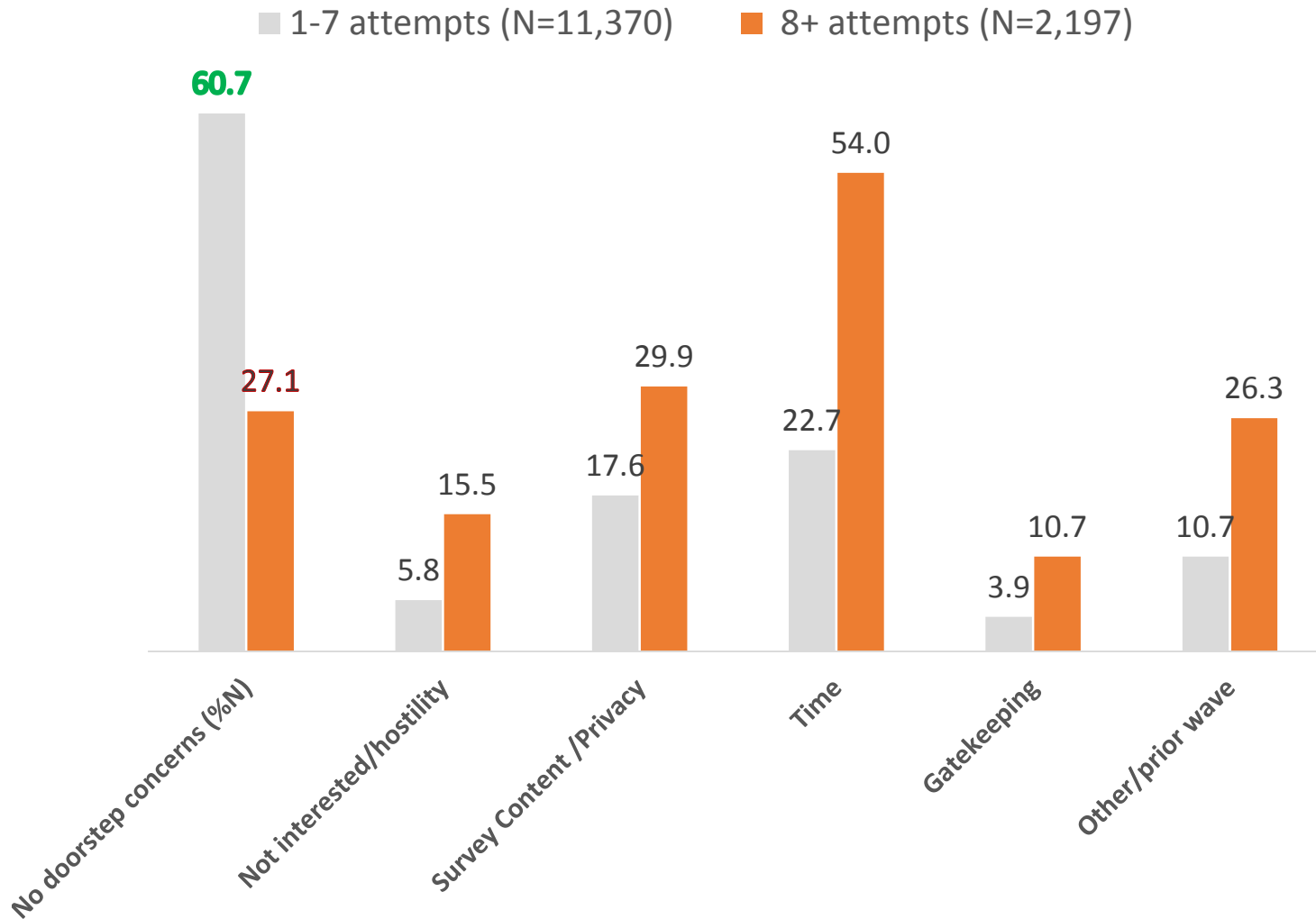
■ 1-7 attempts (N=11,370) ■ 8+ attempts (N=2,197)



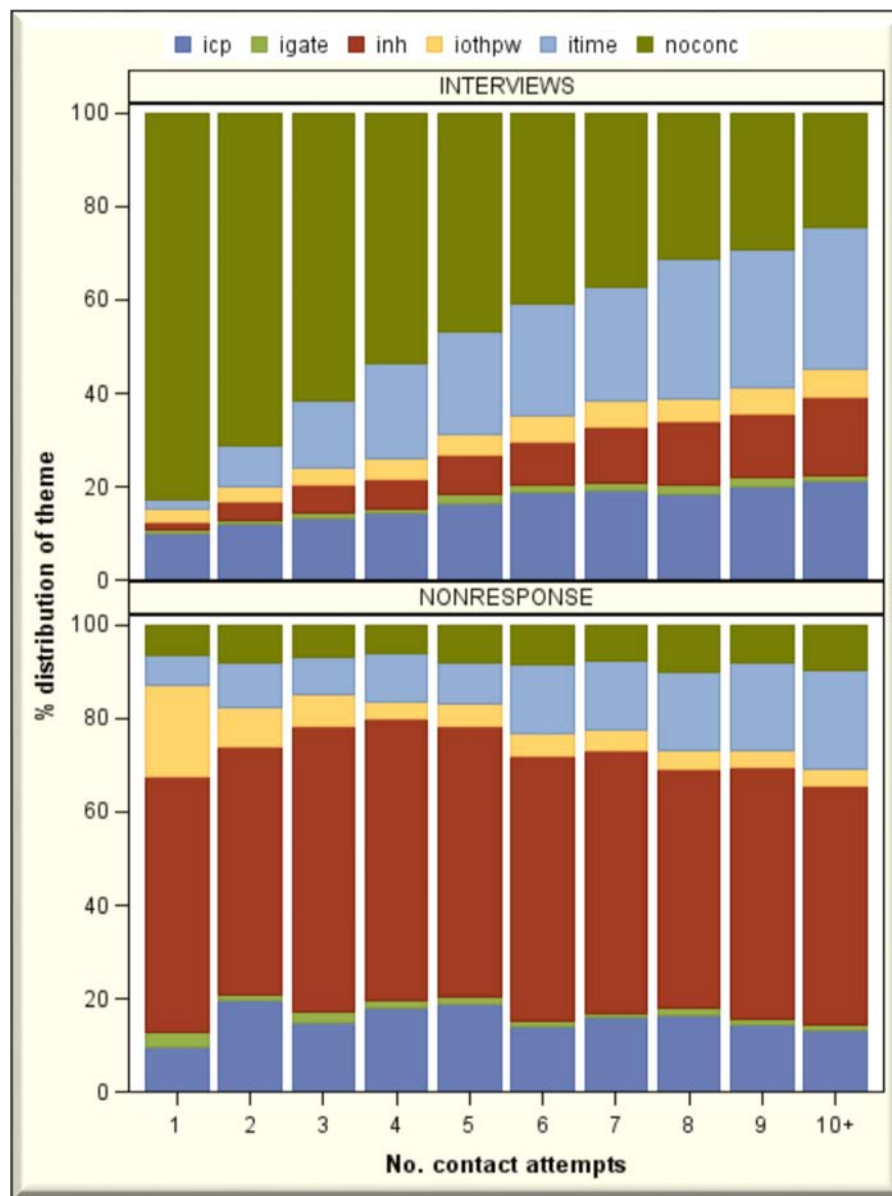
Finding 3. Sample Characteristics

	1-7 attmpts (n=11,370)	8+ attmpts (n=2,197)
Race: white (reference person; %N)	81.4	77.1
Education attainment (reference person; %N)		
Up to high school graduation	37.4	35.4
College graduate	31.5	33.1
Consumer Unit (CU) characteristics		
Size (%N)		
1 person	29.4	29.4
2-3 persons	48.2	45.5
4+ persons	22.4	25.1
No. persons aged 18 and under	0.58	0.71
No. persons aged 65 and older	0.39	0.21
No. earners	0.76	0.82
Dwelling in MSA (%N)	86.6	89.7
Owned home (%N)	64.3	59.0

Finding 4a. Doorstep Concern Theme by Group



Finding 4b. Doorstep Concern Theme by Contact Attempt & Final Disposition



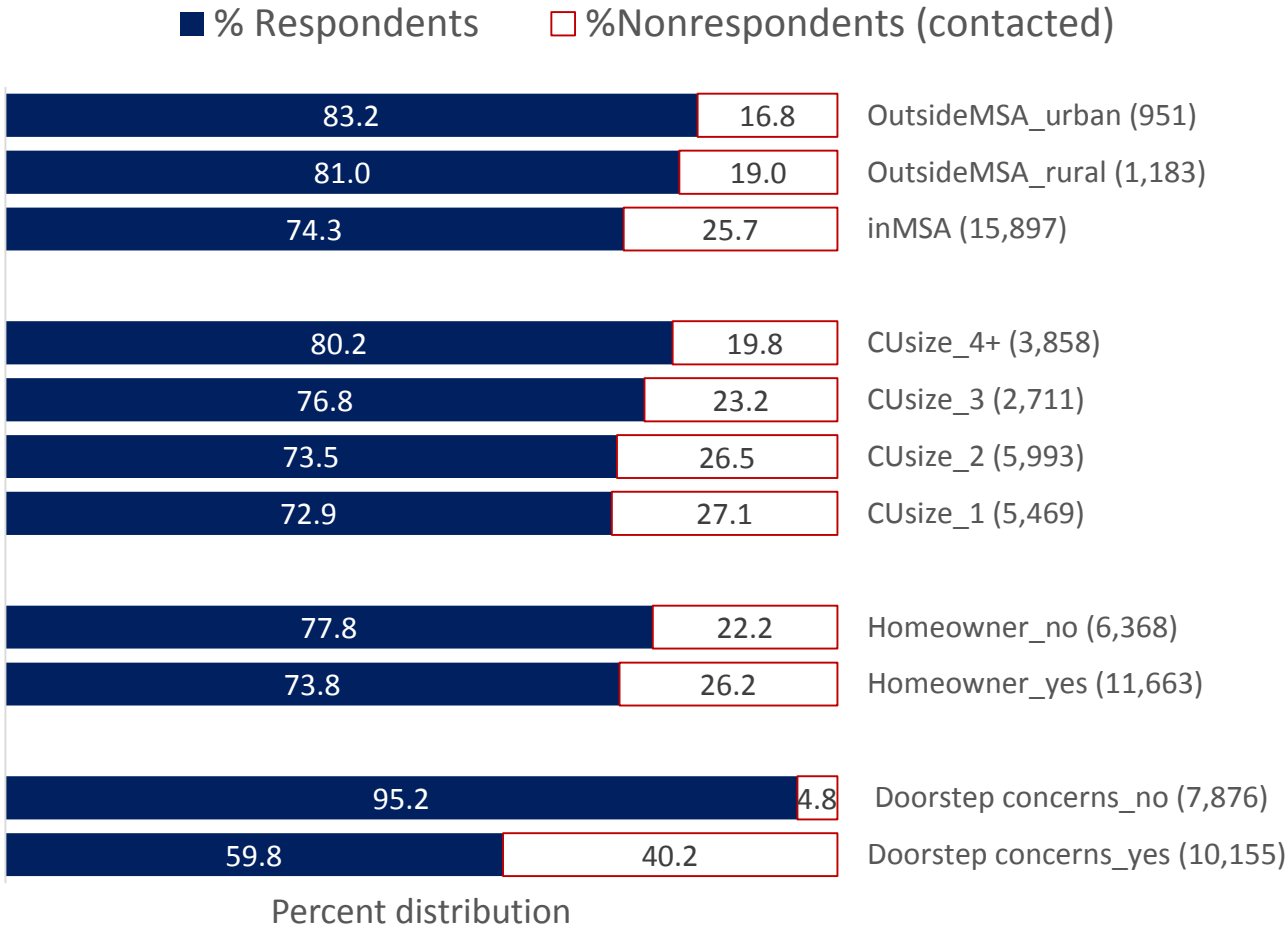
- No concerns
- Time
- Other
- Not interested / hostility
- Gatekeeping
- Survey content/ privacy

* At most 1 doorstep concern

Findings 5a. Response Comp Subgroup RRs

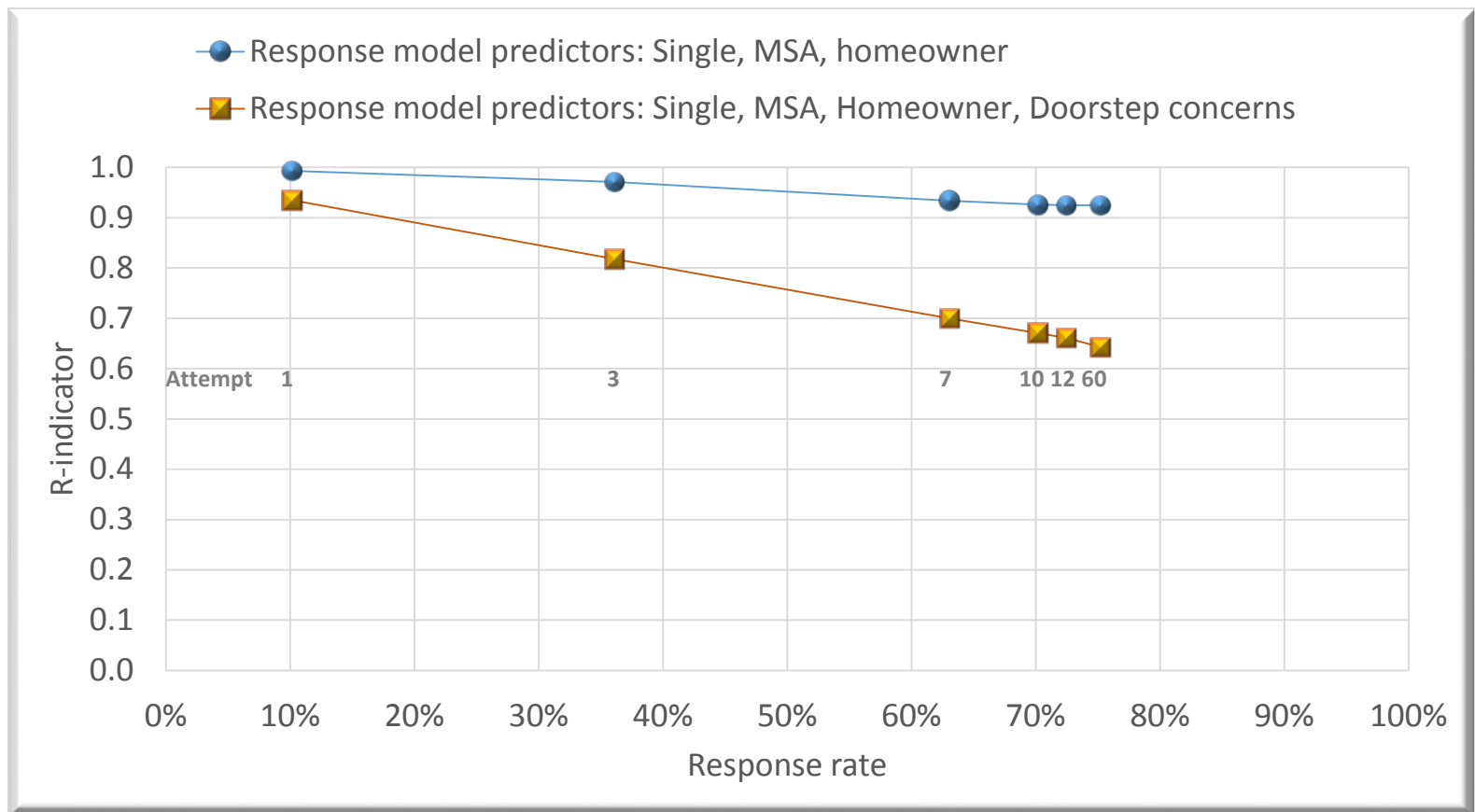
Predictors in response propensity models for R-indicators

Question: When evaluating variation in response propensity within each subgroup, what is the effect of increasing response rates on subgroup response propensities?



Finding 5b. R-Indicators & RRs from 2 Response Propensity Models

An R-indicator is a measure of the variation in response propensity among subgroups of interest. A strong indicator of potential nonresponse bias is a large variation in response propensity within a subgroup, as well as the extent to which that subgroup (e.g., HH tenure) is related to a key variable of interest (e.g., expenditures)



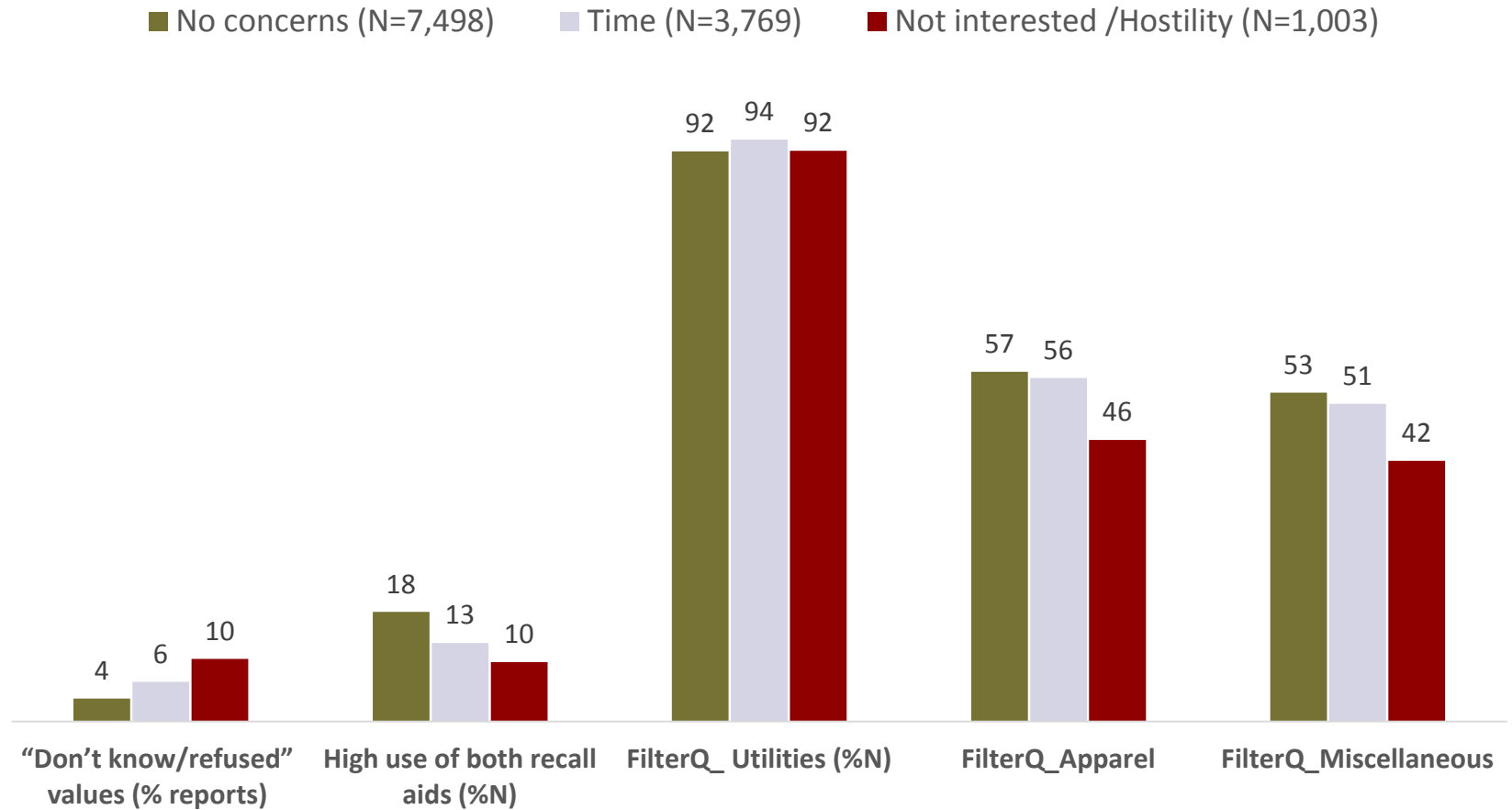
Finding 6a. Reporting Quality

Lower Reporting Quality Indicators in the 8+ Attempt Group

- Higher proportion of “don’t know/refused” expenditure reports
- Lower prevalence in the use of recall aids
- Similar or lower endorsement rates of filter questions to selected expenditure categories
- Lower prevalence in the conduct of the interview entirely by personal visit
- Similar or higher rate of “combined items” reporting among the high attempt respondents

Finding 6b. Reporting Quality

Doorstep Concern Themes & Reporting Quality Indicators

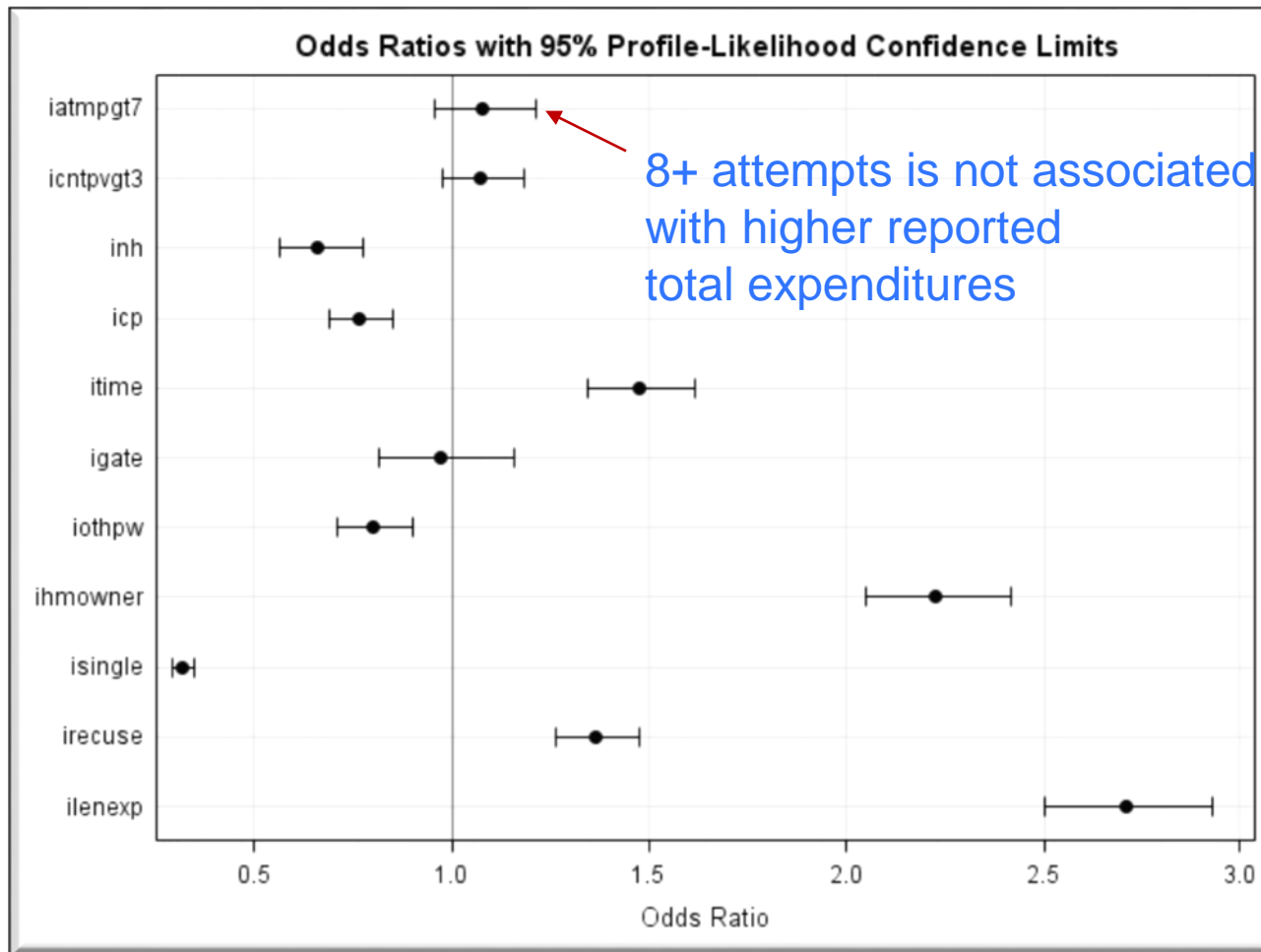


Finding 7. Impact of Higher Attempts on Expenditure Reporting

Logistic regression

Dependent variable: indicator for total expenditures > median value

Predictors (indicator variables): >7 attempts, >3 visit attempts, doorstep concerns (inh, icp, itime, igate, iothpw), and controls (homeowner, single member hh, recall aids used >50%, survey duration > median)



In Summary

Evaluated the impact of increased contact attempts on:

- Response rates
- Cost
- Sample characteristics
- Doorstep concerns
- Response composition
- Reporting quality
- Expenditure reporting

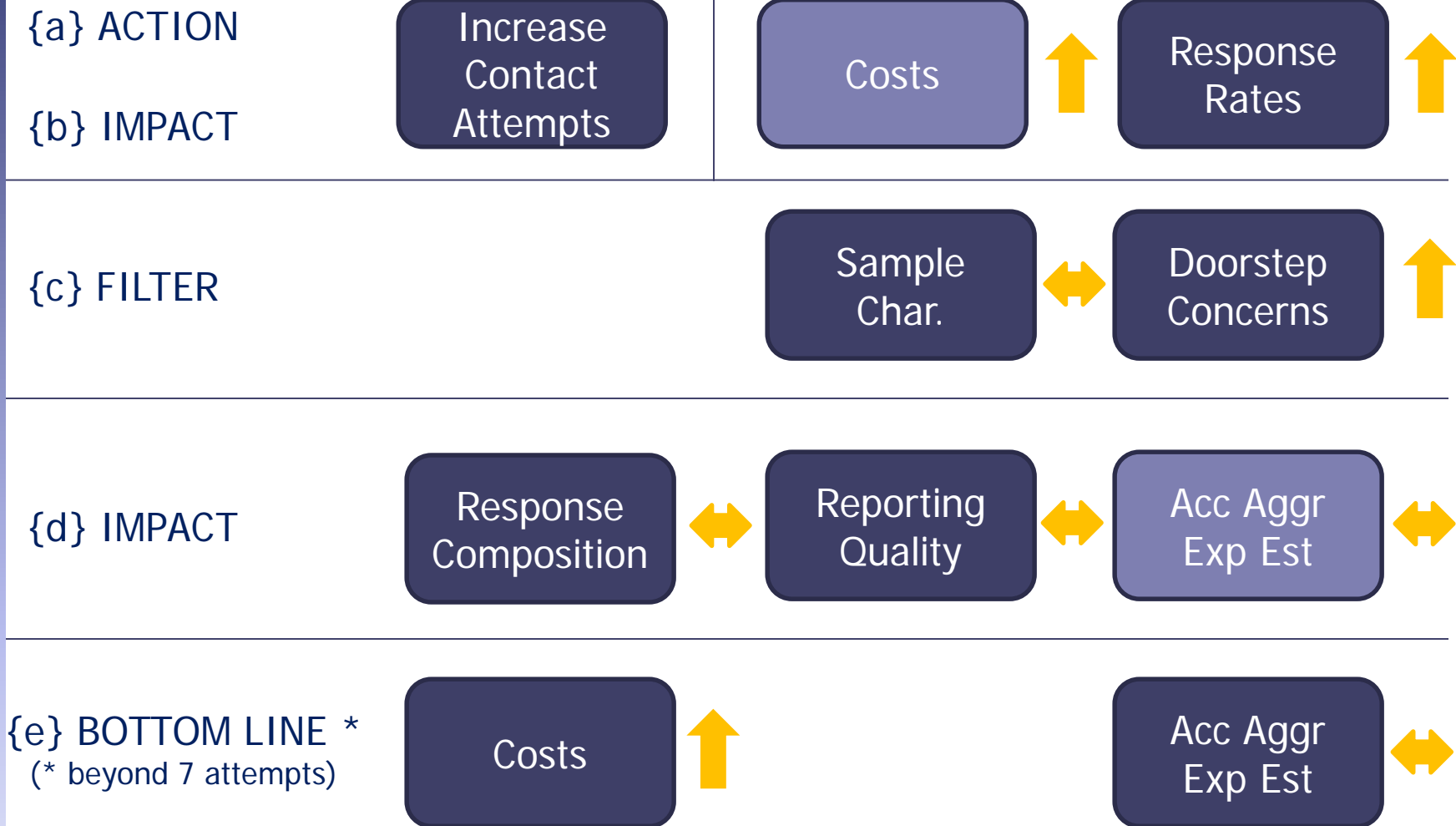
Key Findings

Found that increased contact attempts, beyond the 7-contact attempt threshold:

- did improve **response rates** (12.1 percentage points), but
- was **costly**
- did not substantively impact **sample characteristics**
- increased prop Rs exhibiting high levels of **doorstep concerns**
- did not improve **response composition**
- did not improve **reporting quality** *
- did not improve **expenditure reporting**

* In particular, the “not interested/hostile” doorstep concern theme has been shown to have a strong association with key variables of interest across a number of analyses.

Assessing the Impact of Increased Contact Attempts



Limitations

- Due to the criterion that **only sample units with whom interviewers recorded at least one contact in the CHI** was included in the study sample, in theory, the nonrespondents in this study were restricted to nonresponse due to refusals.
- We lacked **direct measures of reporting quality**. Ideally, the use of recall aids, especially records, would be a very useful indicator. However, the extent of the usage of records and information booklet questions are asked of the interviewers at the end of the survey and based solely on the interviewer's assessment and recall.
- With the paucity of socio-demographic information on the sample frame, we relied on the **CHI** for information that would be available for both respondents and nonrespondents related to their contact attempt characteristics and perceived sample units' pre-survey doorstep concerns. However, the CHI relies solely on the interviewer recording every contact attempt entry and his / her subjective assessment of the contacted sample unit's reaction to the survey request and when that information is recorded.
- In addition, the meaning of the **doorstep concern themes** used in this analyses have not been formally tested cognitively or otherwise for their practical meaning. A better understanding of the correct interpretation of these themes is necessary to inform the design of interventions that would be based on them.

Recommendations

Recommendation #1

- From a total survey error perspective, considering the trade-off between the cost of procuring higher survey response rates versus the benefit of significant data collection cost savings, w/o evidence that the higher response rates are associated with decreasing the potential for nonresponse bias or measurement error, recommend that CE field test a 7-contact attempt threshold.
 - ▶ Also possibility of differentially implementing a strict threshold of 7 contact attempts for hostile respondents, based on CHI.
 - ▶ Bottom line: CHI data have strong associations with key variables of interest, and should be used (in a responsive design, or otherwise).

Recommendation #2

- Reinvest resources in developing a standardized process for calculating data collection costs. Each new effort to calculate data collection costs results in discrepancies, and underscores the impact of information gaps and the need for more complete data.

Contact Information

Adam Safir

**Bureau of Labor Statistics
Division of Consumer Expenditure Survey
Branch of Research and Program Development**

**202-691-5175
safir.adam@bls.gov**

