

An Embedded Experiment for Targeted Nonresponse Follow-Up in Establishment Surveys

Stephen J. Kaputa¹, **Katherine J. Thompson**¹ and Jennifer L. Beck²

¹Economic Statistical Methods Division, U.S. Census Bureau

²National Center for Science and Engineering Statistics, National Science Foundation



U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
[census.gov](https://www.census.gov)

Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied. (CBDRB-FY20-ESMD001-001)

Background

Background

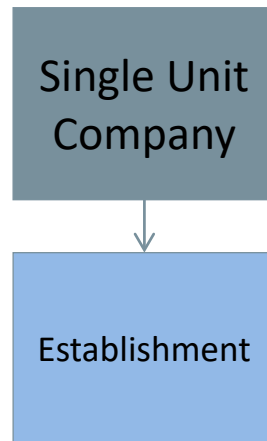
- Census Bureau conducted a series of embedded field experiments on data collection features in preparation for the 2017 Economic Census
 - Ongoing annual business surveys
 - Adaptive nonresponse follow-up (NRFU) as one potential collection strategy
- NRFU protocols considered
 - Targeted NRFU with optimal allocation (aka “Targeted Allocation”)
 - Nonrespondent subsampling

Annual Survey of Manufactures (ASM)

- Alternative to Economic Census in off-census years (manufacturing sector)
 - Similar electronic questionnaire
 - Similar editing/imputation procedures
 - Same sampling unit (establishment)
- Different sampling design
 - Economic Census: Stratified SRS-WOR (sample smallest units)
 - ASM: Stratified PPS-WOR

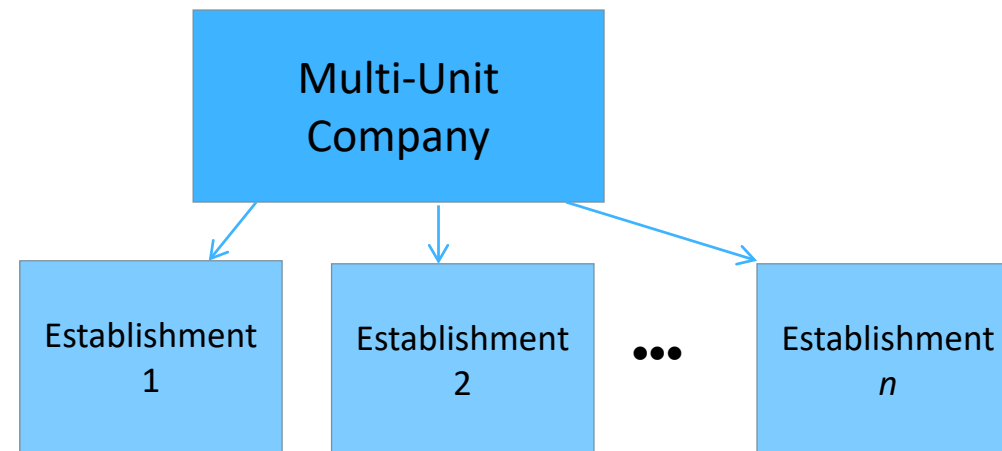
Business Organization Structures

Single Unit (SU)



- One primary industry

Multi Unit (MU)



- Can operate in more than one industry

ASM Adaptive Design Research Projects

Methods/Simulation Studies

Field Tests/Embedded Experiments

Adaptive Design Simulation Study (2014)

- Develop optimized allocation methods for nonrespondent subsampling
- Found that subsampling nonrespondents without changing the contact strategy may have minimal benefits

Thompson, K.J., Kaputa, S., and Bechtel, L. 2018. Strategies for Subsampling Nonrespondents for Economic Programs. *Survey Methodology*, 44 (1), pp. 75-99.

ASM Adaptive Design Research Projects

Methods/Simulation Studies

Field Tests/Embedded Experiments

Adaptive Design Simulation Study (2014)

- Develop optimized allocation methods for nonrespondent subsampling
- Found that subsampling nonrespondents without changing the contact strategy may have minimal benefits

Adaptive Design Simulation Study (2015)

- Enhanced proposed methods for nonrespondent subsampling
- Promising results incorporating distance and balance metrics into objective function

Kaputa, S and Thompson, K.J. 2018. Adaptive Design Strategies for Nonresponse Follow-Up in Economic Surveys. *JOS*, 34(2), pp. 445–462. DOI: <https://doi.org/10.2478/jos-2018-0020>.

ASM Adaptive Design Research Projects

Methods/Simulation Studies

Adaptive Design Simulation Study (2014)

- Develop optimized allocation methods for nonrespondent subsampling
- Found that subsampling nonrespondents without changing the contact strategy may have minimal benefits

Adaptive Design Simulation Study (2015)

- Enhanced proposed methods for nonrespondent subsampling
- Promising results incorporating distance and balance metrics into objective function

Field Tests/Embedded Experiments

ASM Field Test (2015)

- Tested three contact strategies for “hard to reach” single unit establishments
- Recommended certified letter for recontact

Thompson, K.J. and Kaputa, S. 2017. Investigating Adaptive Nonresponse Follow-up Strategies for Small Businesses through Embedded Experiments. *JOS*, 33 (3), pp. 1–23.

ASM Adaptive Design Research Projects

Methods/Simulation Studies

Adaptive Design Simulation Study (2014)

- Develop optimized allocation methods for nonrespondent subsampling
- Found that subsampling nonrespondents without changing the contact strategy may have minimal benefits

Adaptive Design Simulation Study (2015)

- Enhanced proposed methods for nonrespondent subsampling
- Promising results incorporating distance and balance metrics into objective function

Field Tests/Embedded Experiments

2014 ASM Field Test (conducted in 2015)

- Tested three contact strategies for “hard to reach” single unit establishments
- Recommended certified letter for recontact

2015 ASM Field Test (conducted 2016) – Presented Today

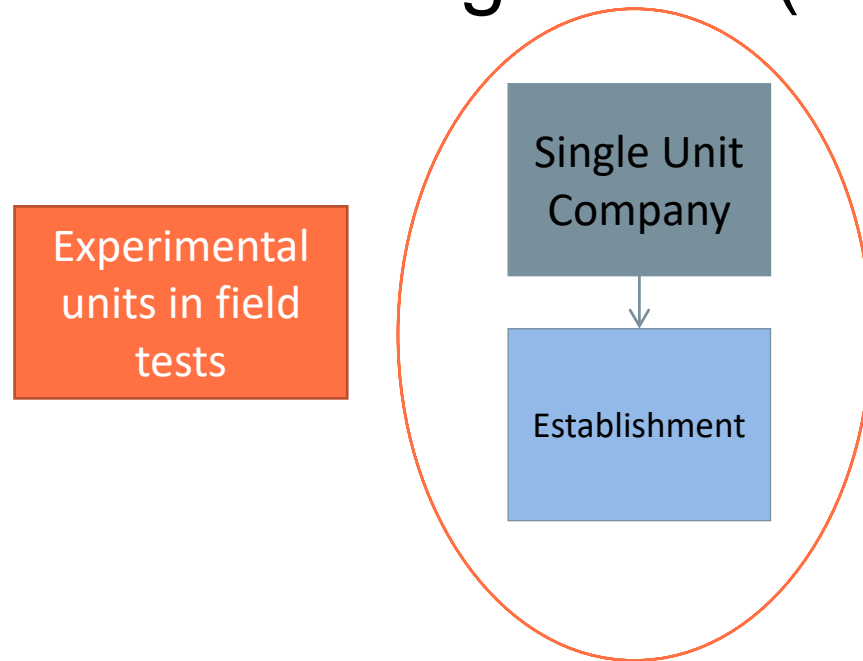
- Original Goal: Combine nonrespondent subsampling with certified letters (most effective)
- Problem: Certified letters are now standard NRFU (outcome from 2015 test)
- **New Goal: Compare across-the-board certified letter NRFU to...targeted allocation and nonrespondent subsampling**

Kaputa, S., Thompson, K.J., and Beck, J.L. 2019. An Embedded Experiment for Targeted Non-Response Follow-Up in Establishment Surveys. *Journal of the Royal Statistical Society Series A*, 182(4), pp. 1371–1391. DOI:10.1111/rssa.12448

2015 Field Test Design

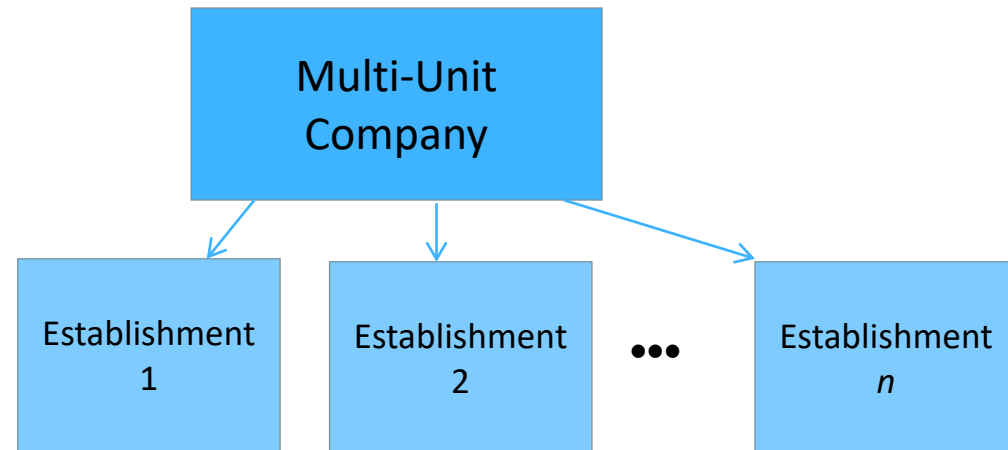
Business Organization Structures

Single Unit (SU)



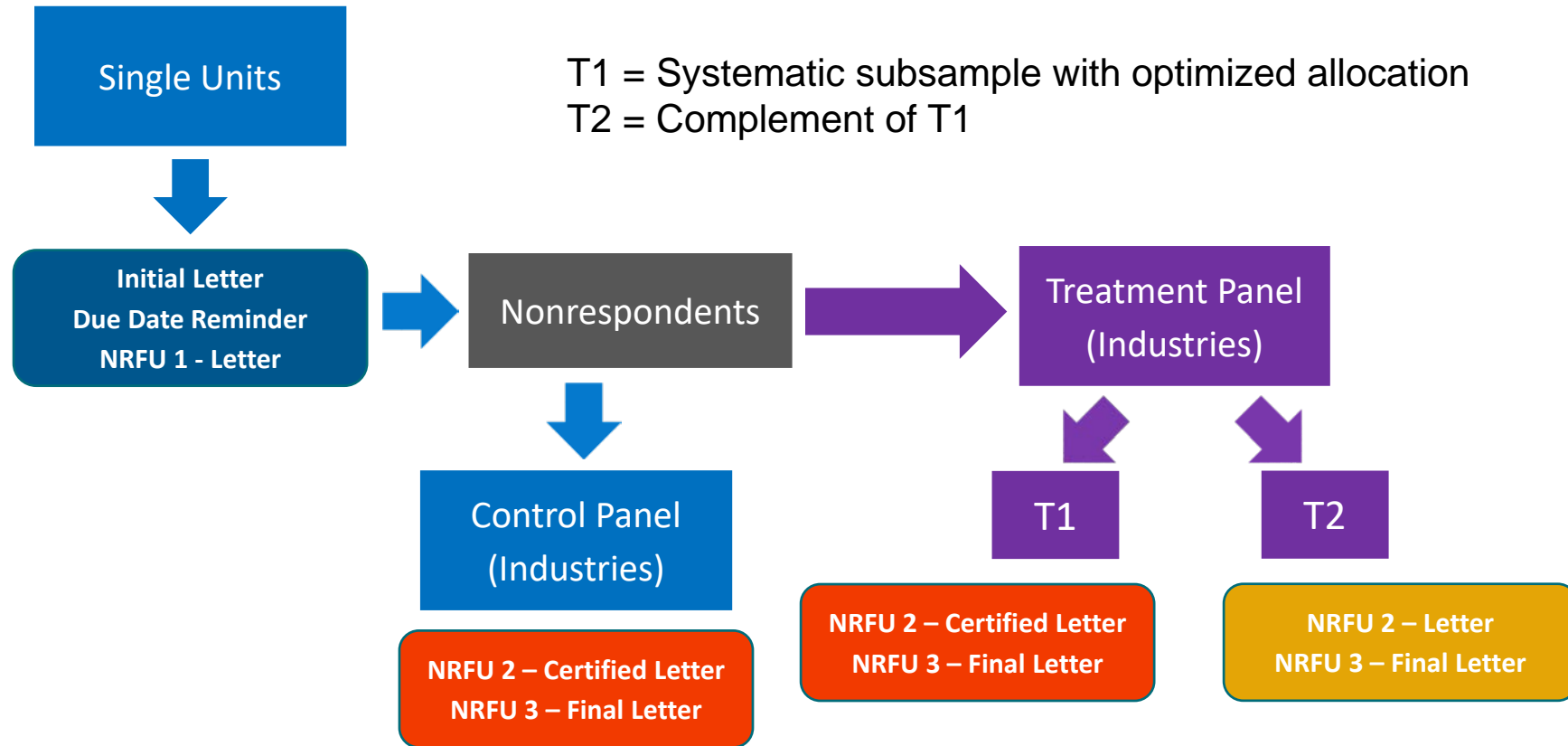
- One primary industry

Multi Unit (MU)

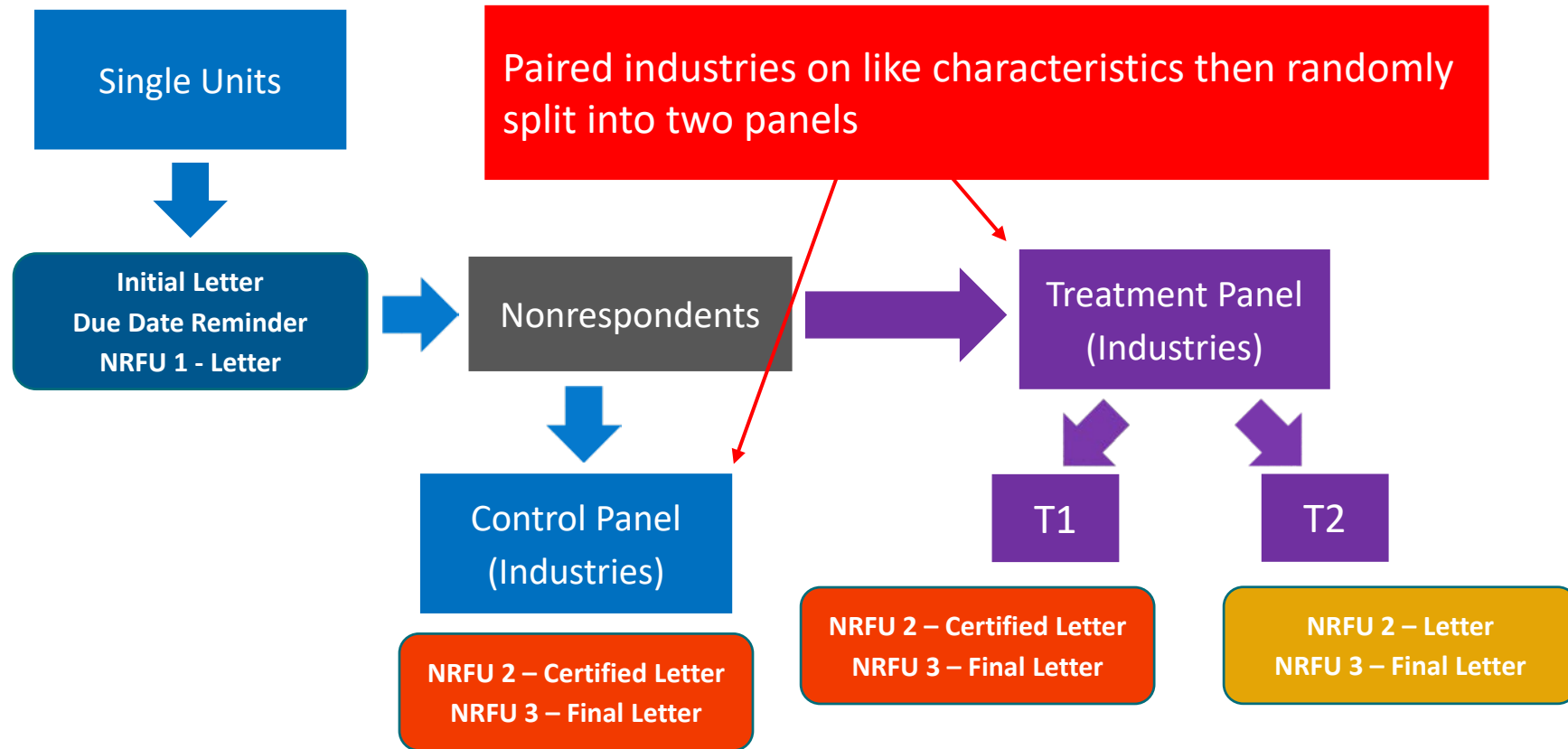


- Can operate in more than one industry

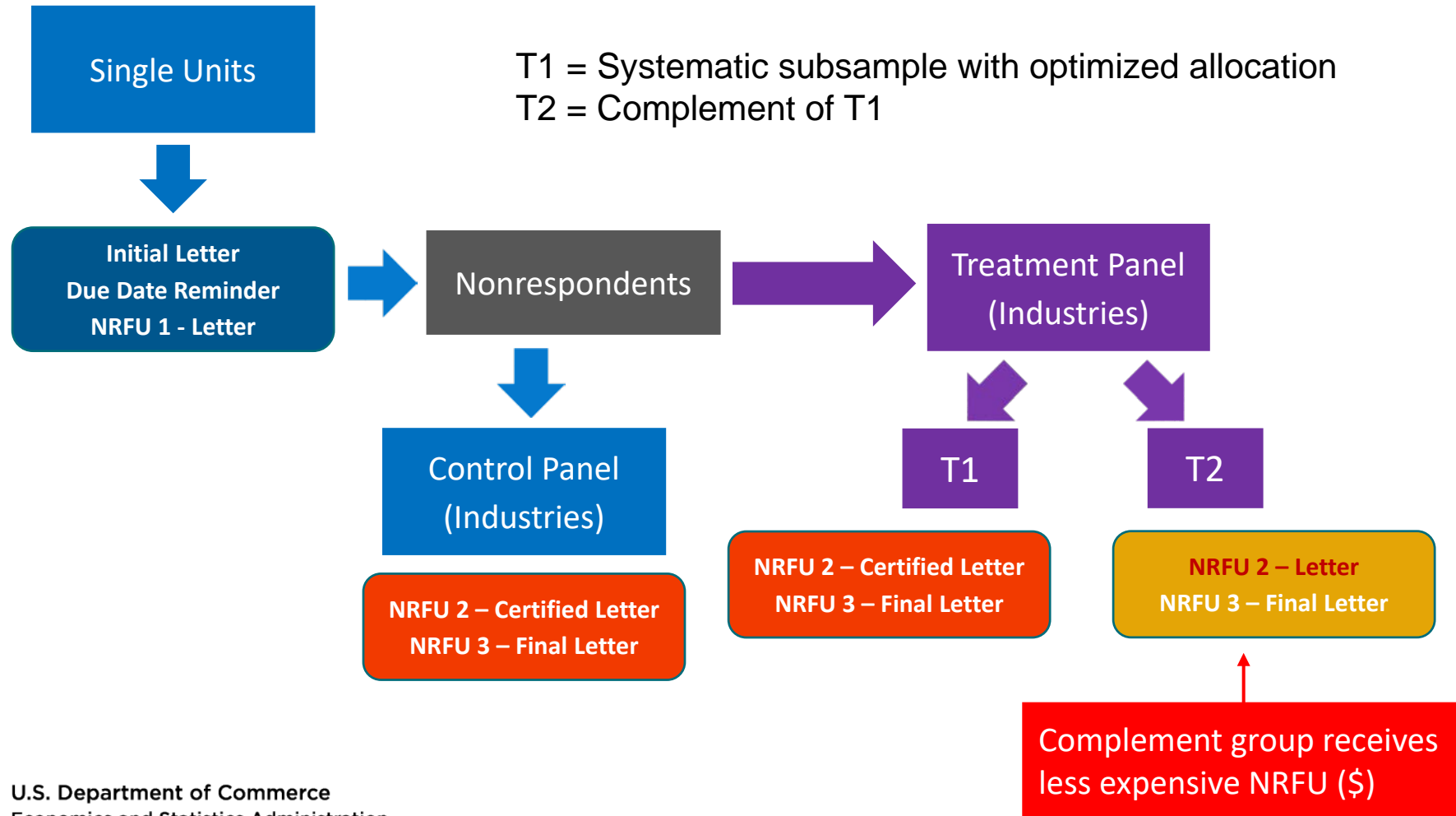
2015 ASM Field Experiment



2015 ASM Field Experiment



2015 ASM Field Experiment



Industry Allocations

- Industries are allocated certified letters based on their final predicted response rates given industry-level subsampling rates
- The goal is to equalize response rates across industries
 - The allocation is formulated as a quadratic program
 - Wait for next presentation to learn all about it

Tested NRFU Treatments (Single Units)

Control Panel

Domain 1	Respondents	\$\$\$ NRFU
Domain 2	Respondents	\$\$\$ NRFU
...		
Domain <i>h</i>	Respondents	\$\$\$ NRFU

Treatment Panel – ALL nonresponding units are contacted

- Controlled selection of certified (\$\$\$ NRFU) and noncertified mailings (\$ NRFU)

Domain 1	Respondents	\$\$\$ NRFU	\$ NRFU
Domain 2	Respondents	\$\$\$ NRFU	\$ NRFU
...			
Domain <i>h</i>	Respondents	\$\$\$ NRFU	\$ NRFU

Simulated NRFU Treatment (Single Units)

Treatment Panel – ALL nonresponding units are contacted

Domain 1	Respondents	\$\$\$ NRFU	\$ NRFU
Domain 2	Respondents	\$\$\$ NRFU	\$ NRFU
...			
Domain <i>h</i>	Respondents	\$\$\$ NRFU	\$ NRFU

Nonrespondent Subsampling – Only \$\$\$ NRFU units receive contact

- “Ignore” responses from \$ units (simulated and pessimistic results)

Domain 1	Respondents	\$\$\$ NRFU	
Domain 2	Respondents	\$\$\$ NRFU	
...			
Domain <i>h</i>	Respondents	\$\$\$ NRFU	

Evaluation methods

Single Unit Evaluation Measures

Measure	Description	Range	Level
(Proxy) Unit Response Rate	Proportion of responding establishments to sampled establishments (unweighted)	0% - 100%	Panel
Quantity Response Rates	Proportion of tabulated item value obtained from reported data (weighted)	0% - 100%	Panel by item
Source of Data Item	Proportion of responding units that retain reported data after processing	0% - 100%	Panel by item
Fraction of Missing Information (FMI)	Measure of effects of potential nonresponse bias on collected items	0 - 1	Panel by item

Fraction of Missing Information (FMI)*

$$\sigma_{Total}^2 = \sigma_{Within}^2 + \sigma_{Between}^2$$

↑
Variance of μ_y had
all units responded

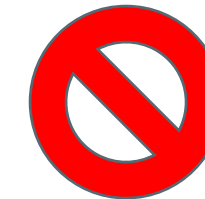
↙
Variance due to ~~imputation method~~ imbalance
in the respondent sample

$$FMI(Y) = \frac{\sigma_{Within}^2}{\sigma_{Total}^2}$$

$$FMI(Y) \rightarrow 0$$



$$FMI(Y) \rightarrow 1$$



Obtaining the FMI

$$\sigma_{Total}^2 = \sigma_{Within}^2 + \sigma_{Between}^2$$

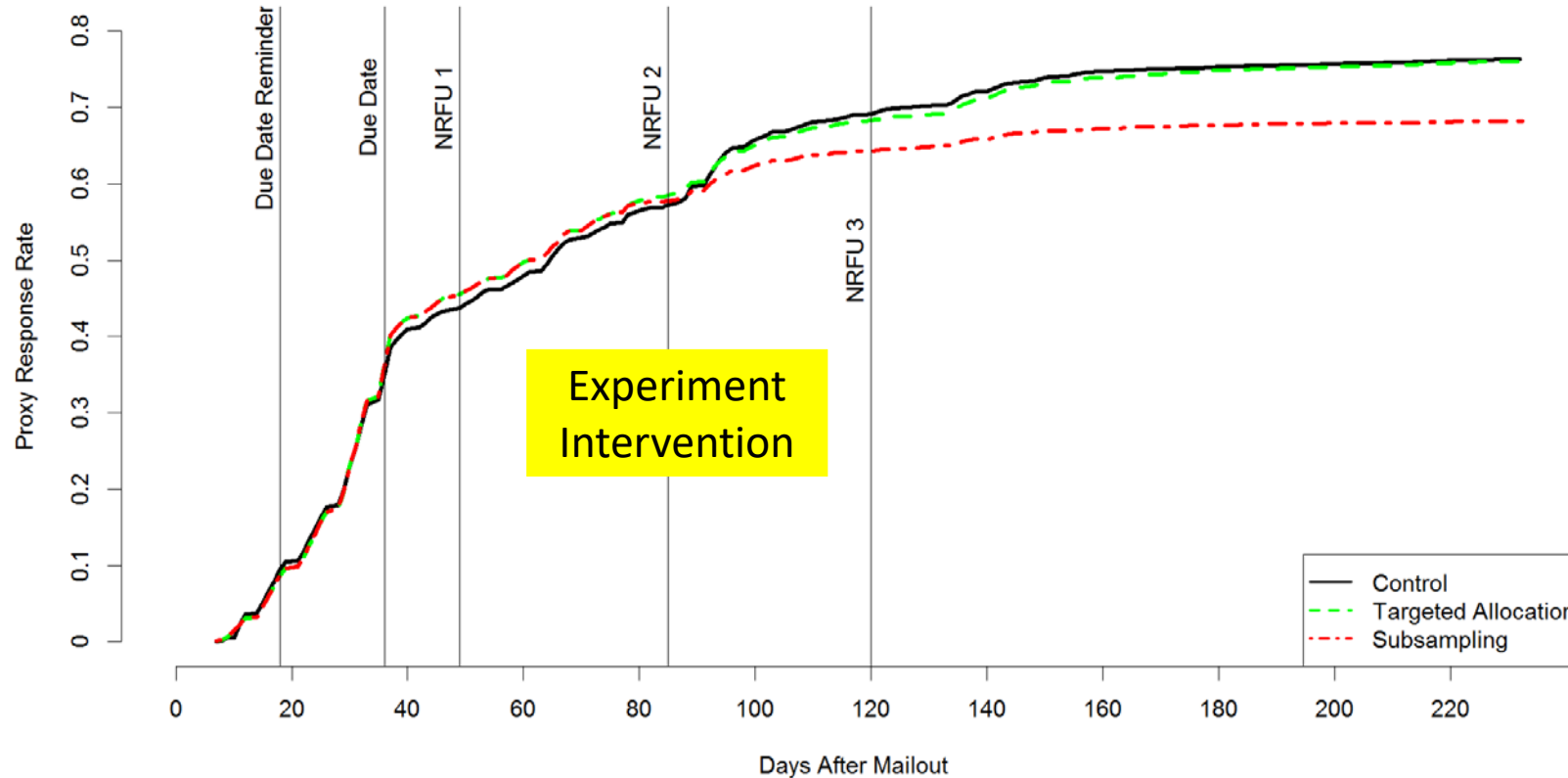
↑
Variance of μ_y had
all units responded

↙
Variance due to ~~imputation method~~
imbalance in the respondent sample

- Multiple Imputation
- Proxy Pattern-Mixture Model
 - Gamma PPM Model (Andridge and Thompson 2015)
 - Predict outcome variable from frame measure of size (Proxy)
 - Obtain different models for respondents and nonrespondents (Pattern-Mixture model)
 - Measures sensitivity by range of response mechanisms
 - MAR: missing at random – “Best Case”
 - NMAR: not missing at random – “Worst Case”

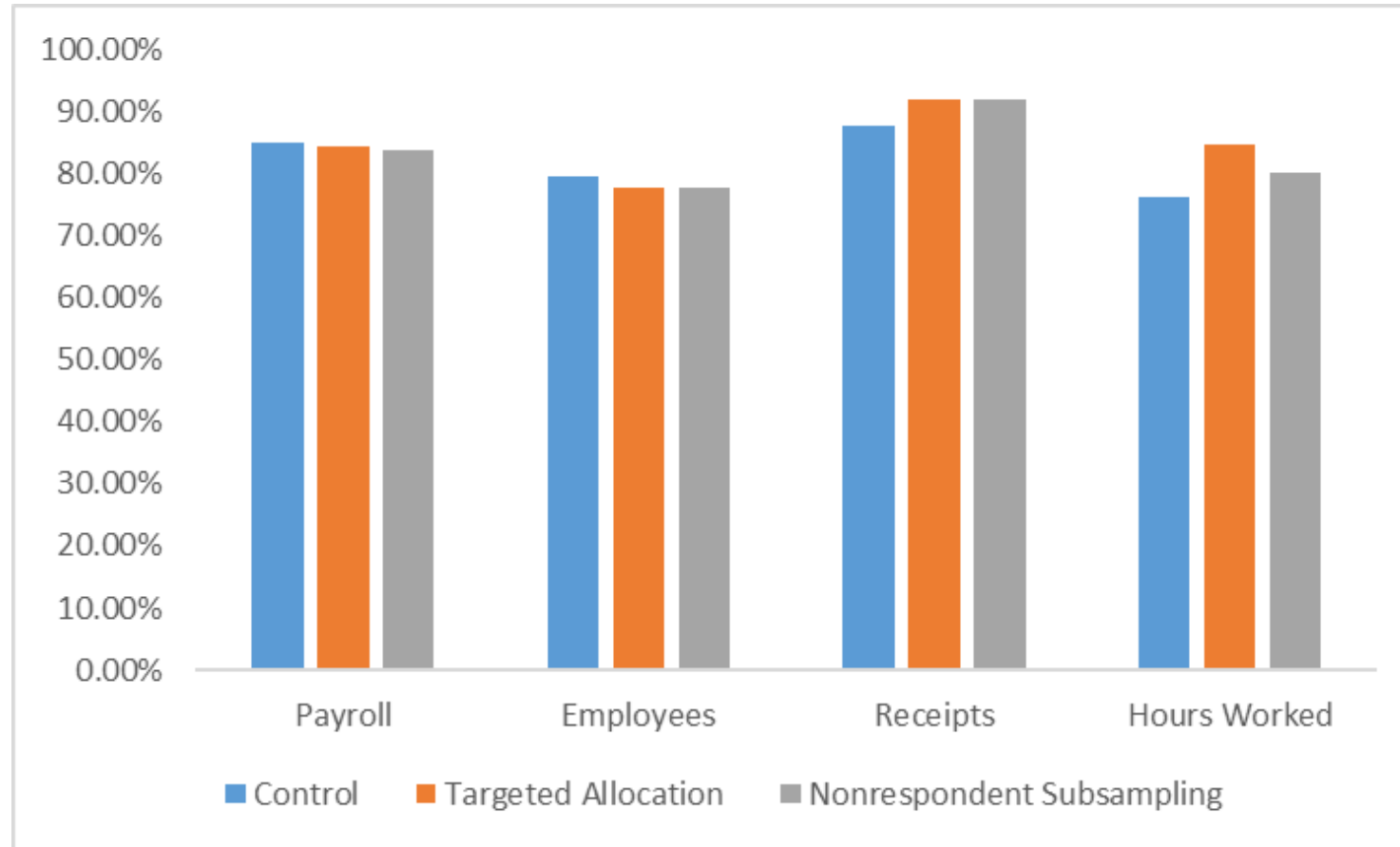
SINGLE UNIT Results

Unit Response Rate



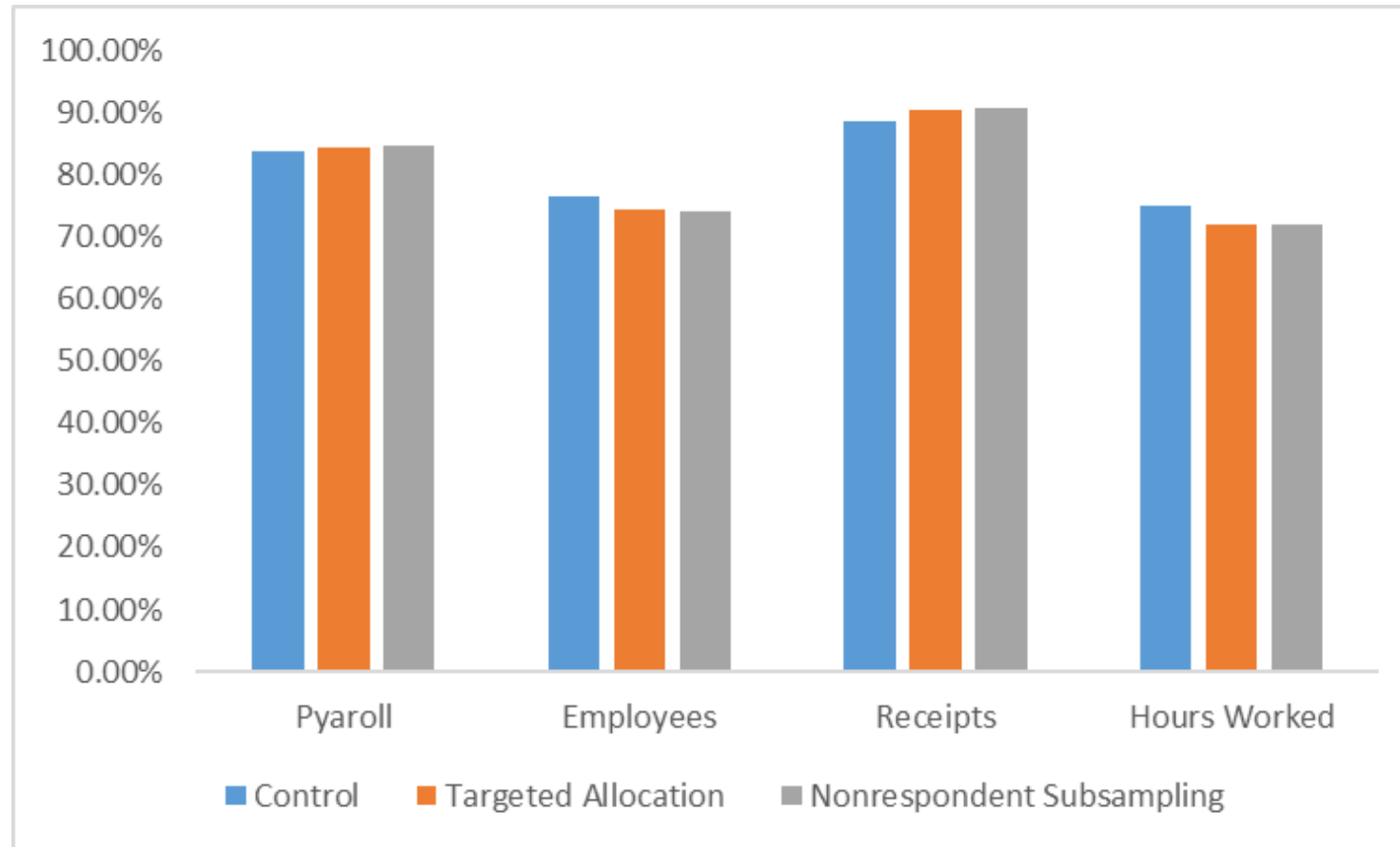
Quantity Response Rates

Proportion of tabulated item value obtained from reported data

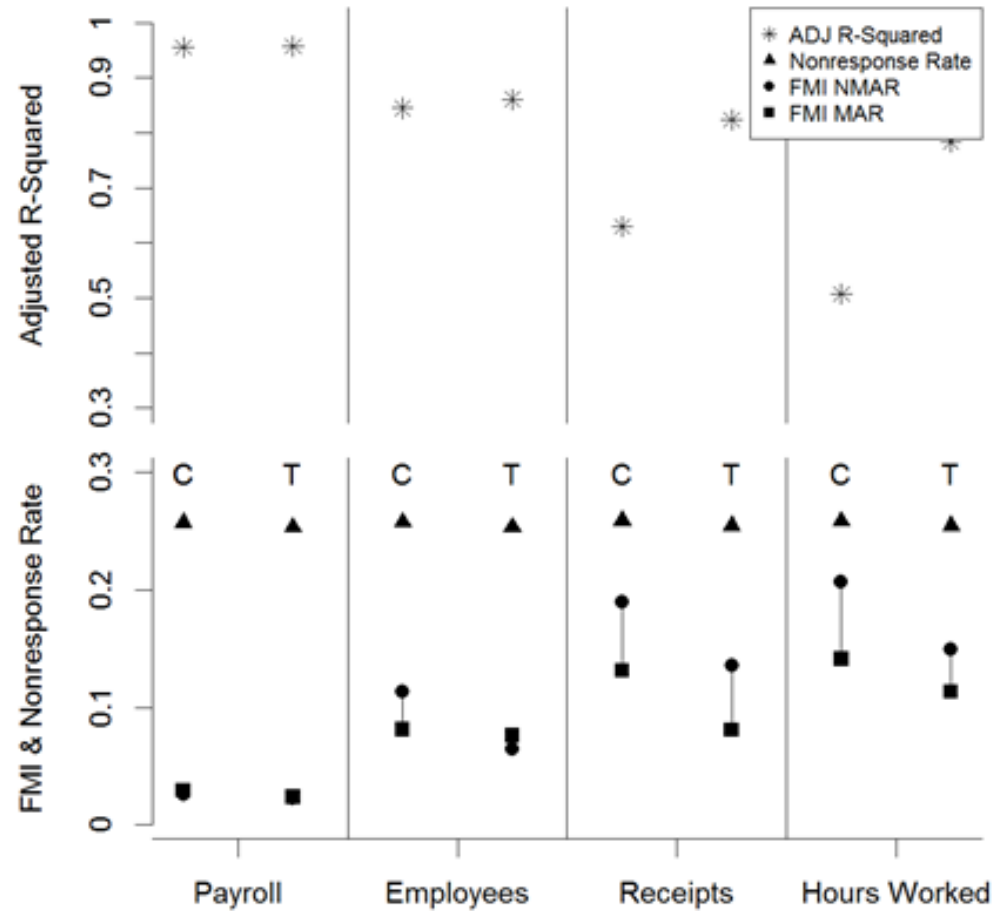


Source of Data Item

Proportion of responding units that retain reported data after processing

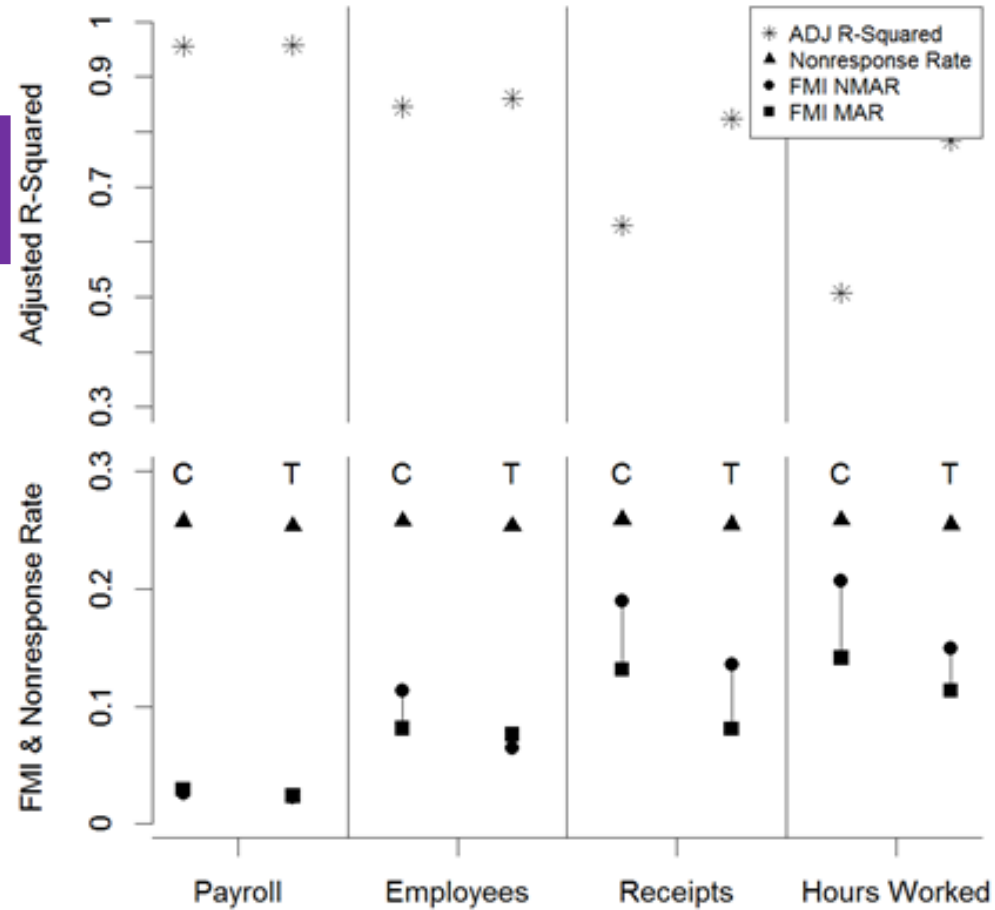


FMI: Control (C) versus Targeted Allocation (T)



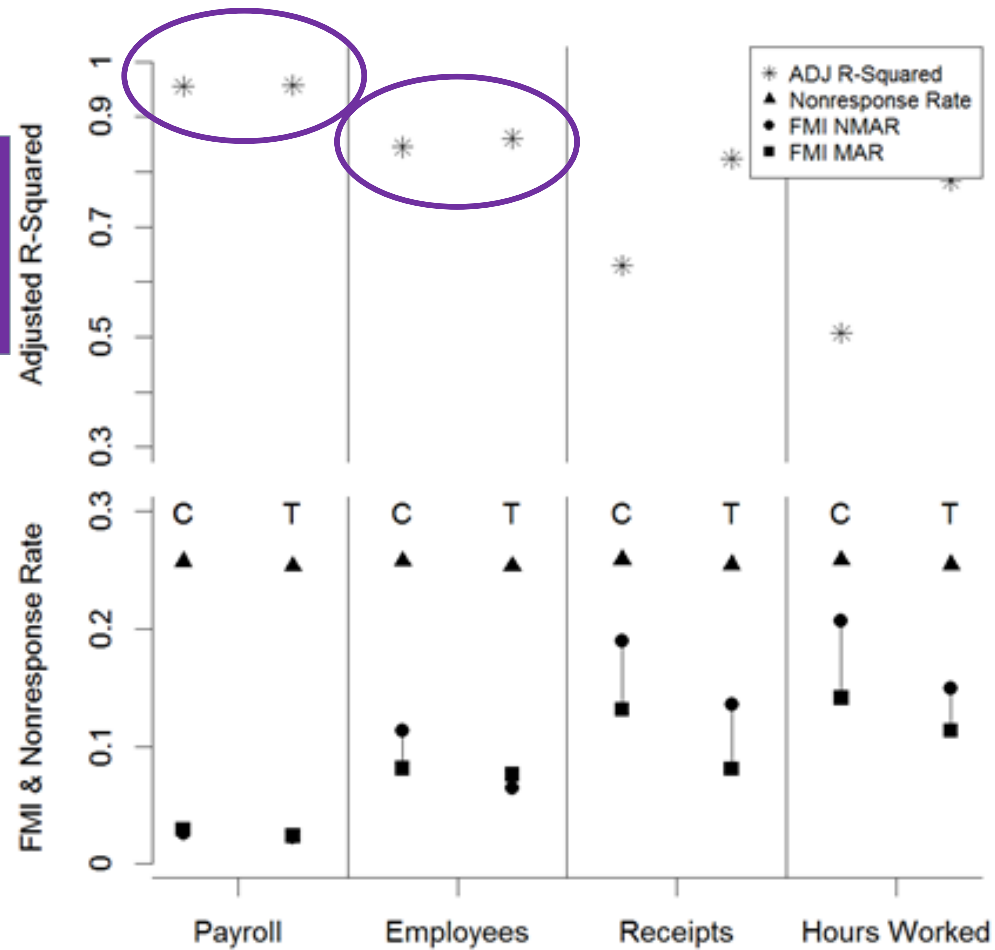
FMI:Control (C) versus Targeted Allocation (T)

Strength of Proxy

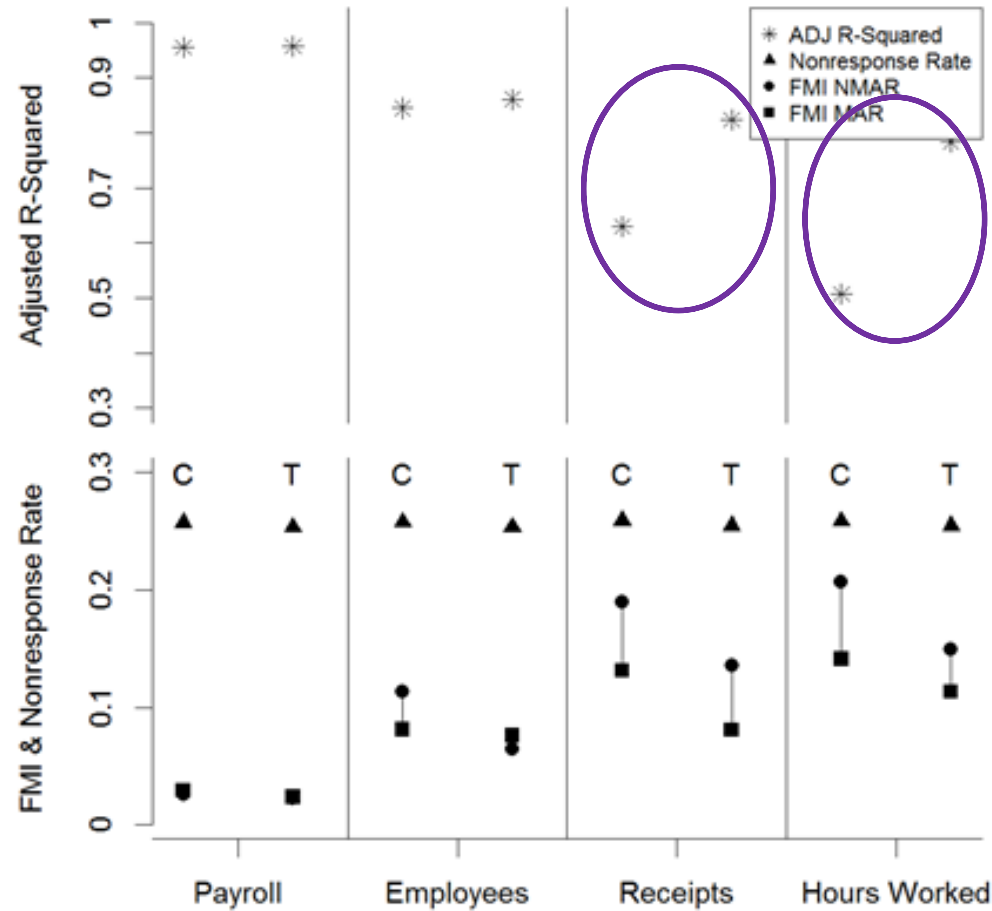


FMI: Control (C) versus Targeted Allocation (T)

Strength of proxy is about the same in both panels (C & T)

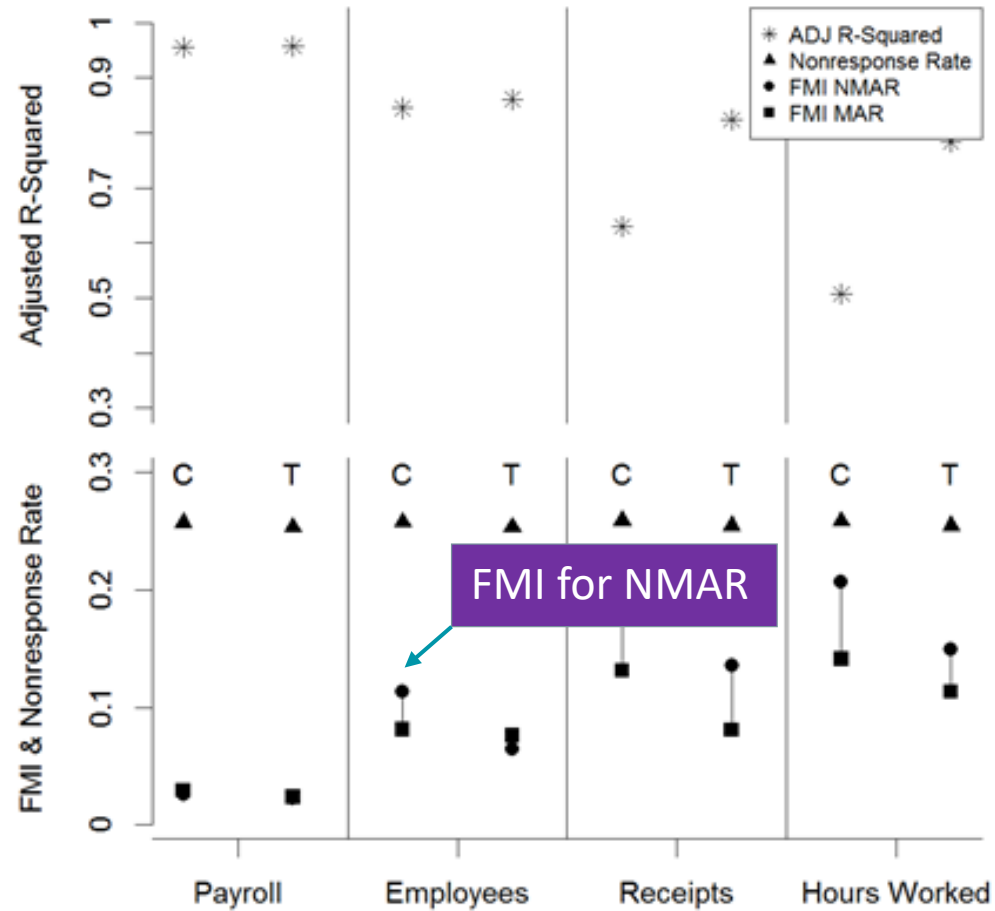


FMI: Control (C) versus Targeted Allocation (T)

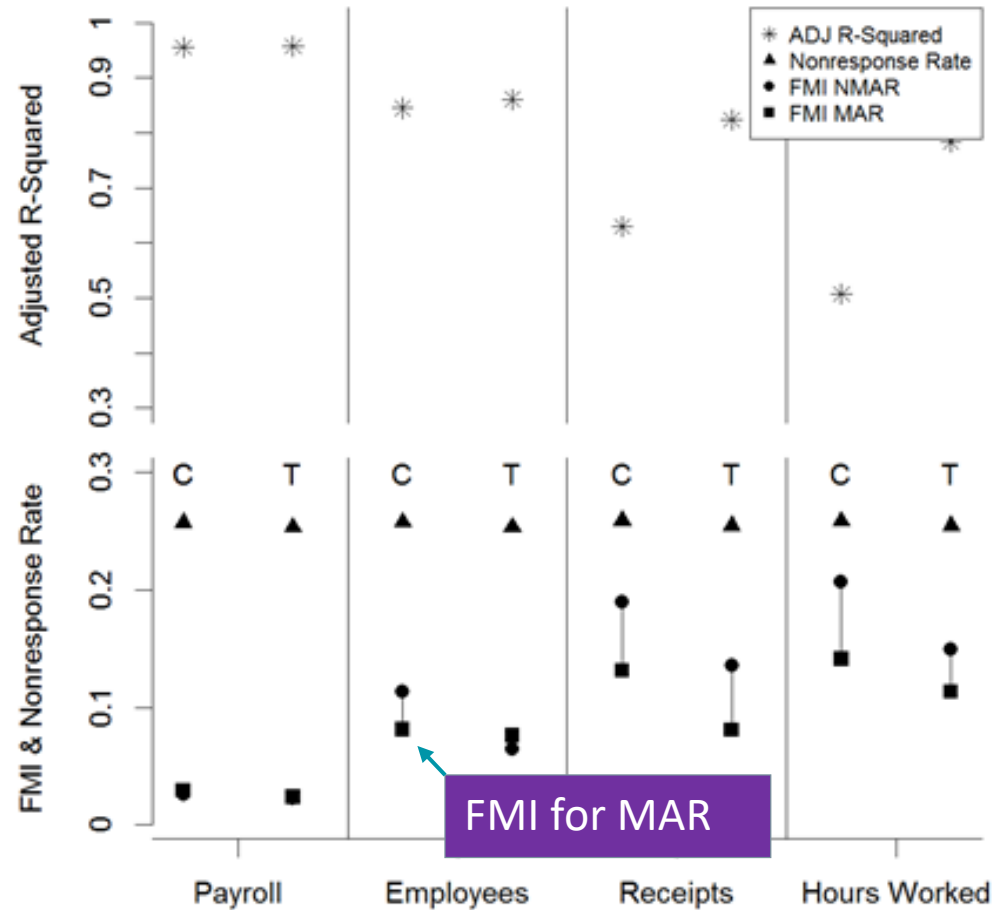


Strength of proxy is stronger in treatment panel

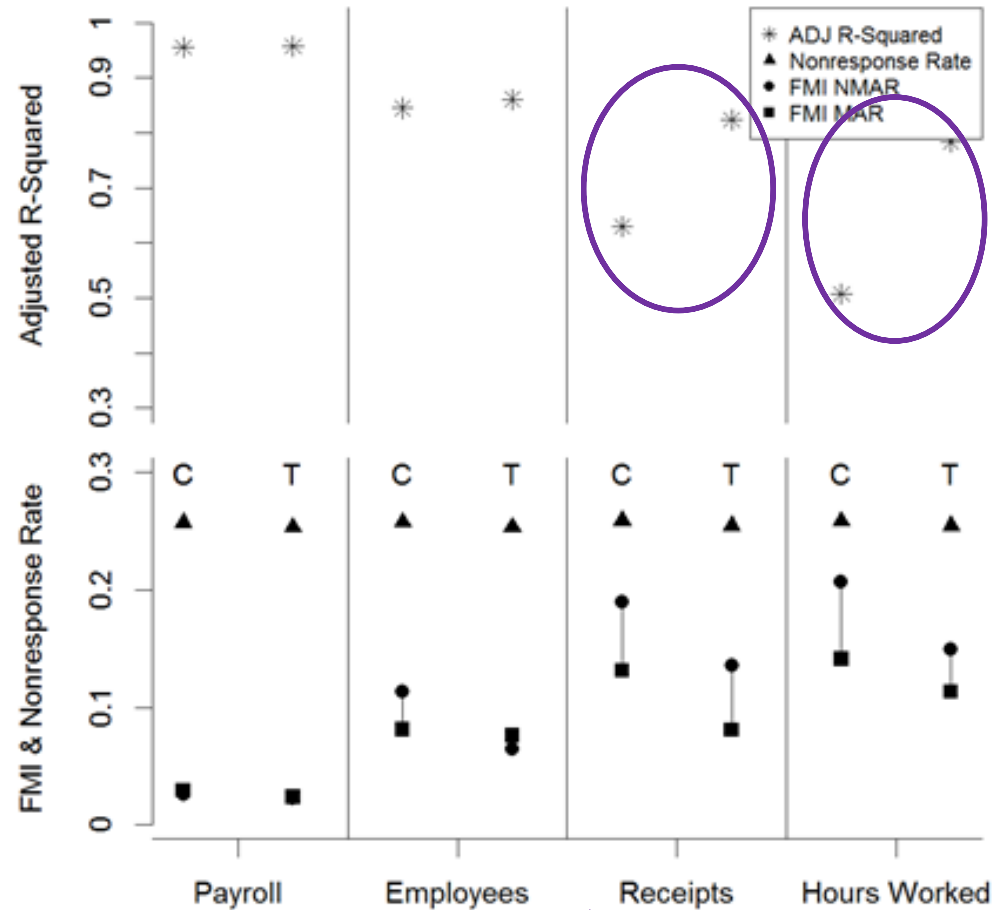
FMI: Control (C) versus Targeted Allocation (T)



FMI: Control (C) versus Targeted Allocation (T)



FMI: Control (C) versus Targeted Allocation (T)



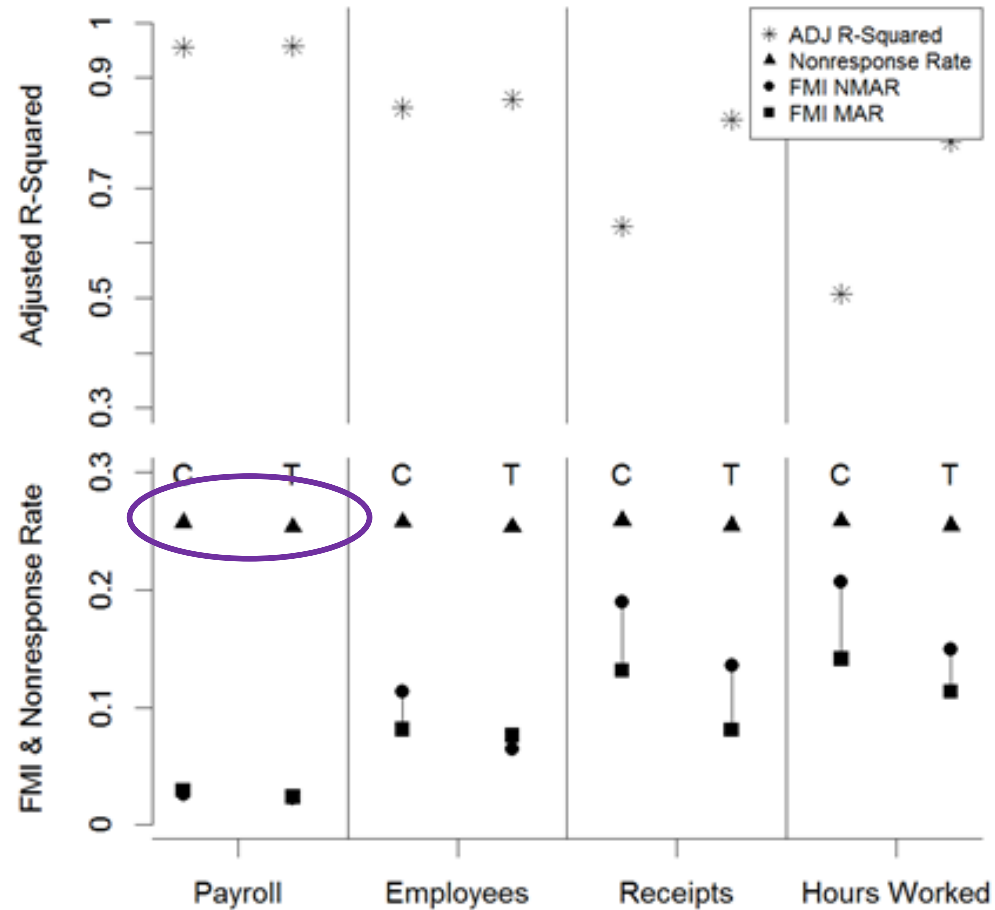
Strength of proxy is stronger in treatment panel

Effect of varying strength of proxy shown in spread between control and target FMIs

FMI: Control (C) versus Targeted Allocation (T)

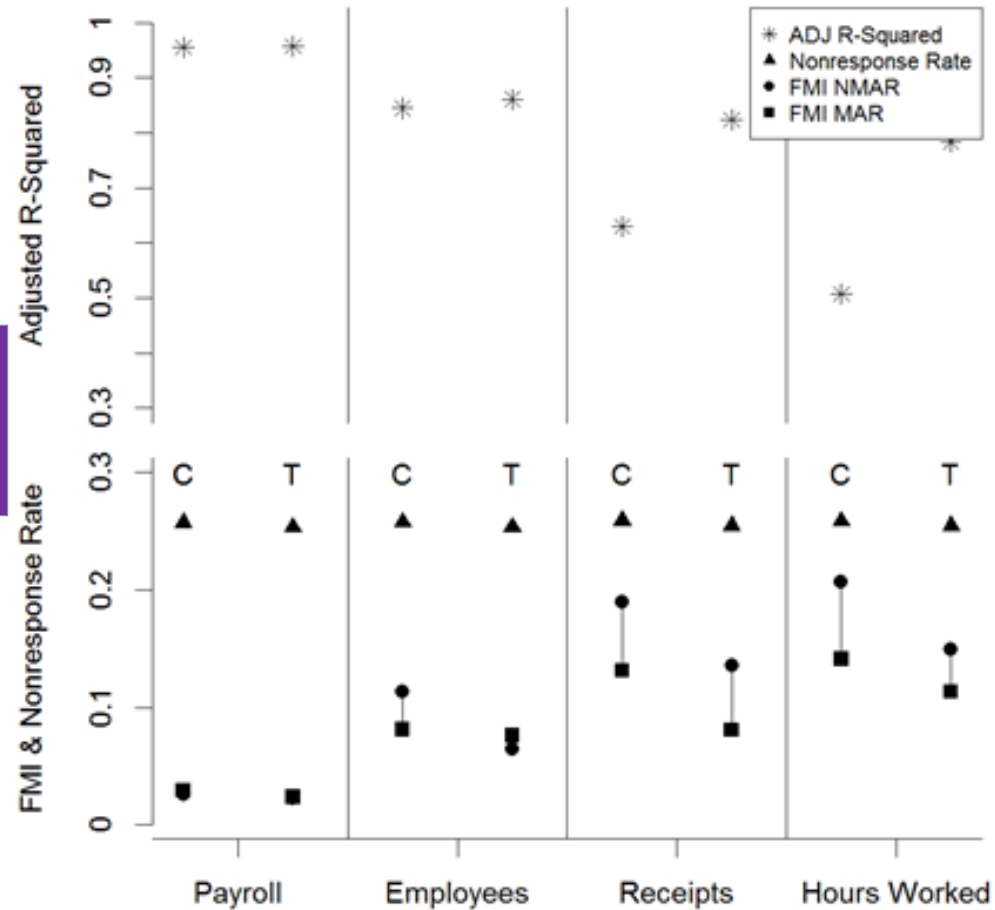
Unit Nonresponse Rate (NRR)

If $FMI < NRR$ then response set is said to be “balanced”



FMI: Control (C) versus Targeted Allocation (T)

The targeted allocation FMI is less than the control FMI for all variables



Conclusion

- Targeted allocation procedure effective
 - Maintains unit and item response rates
 - No detrimental effects on data quality
 - Reduced cost compared to the uniform follow-up
- Targeted allocation procedure of single units was implemented in the 2016 ASM
- Implementation for the 2017 Economic Census discussed next

References

- Andridge, R.R. and K.J. Thompson. 2015(B). Assessing nonresponse bias in a business survey: proxy pattern-mixture analysis for skewed data. *Annals of Applied Statistics*, 9(4), pp. 2237–2265.
- Kaputa, S and Thompson, K.J. 2018. Adaptive design strategies for nonresponse follow-up in economic surveys. *JOS*, 34(2), pp. 445–462. DOI: <https://doi.org/10.2478/jos-2018-0020>.
- Thompson, K.J. and Kaputa, S. 2017. Investigating adaptive nonresponse follow-up strategies for small businesses through embedded experiments. *JOS*, 33 (3), pp. 1–23.
- Thompson, K.J., Kaputa, S., and Bechtel, L. 2018. Strategies for subsampling nonrespondents for economic programs. *Survey Methodology*, 44 (1), pp. 75-99.
- Kaputa, S., Thompson, K.J., and Beck, J.L. 2019. An embedded experiment for targeted non-response follow-up in establishment surveys. *Journal of the Royal Statistical Society Series A*, 182(4), pp. 1371–1391. DOI:10.1111/rssa.12448.
- Wagner, J. 2010. The fraction of missing information as a tool for monitoring the quality of survey data. *Public Opinion Quarterly* ,74, pp. 223–243.

Thanks!!!

Katherine.J.Thompson@census.gov