Improving the Efficiency of Address-Based Frames with the USPS No-Stat File

Bonnie Shook-Sa

2013 AAPOR Conference

Boston, MA

May 18, 2013

Overview

- Address-Based Sampling for In-Person Surveys
- The problem of rural undercoverage
- USPS No-Stat File
  - Background / Prior research
  - Hybrid frame analysis
  - Conclusions / Limitations / Future Research

RTI INTERNATIONAL

Address-Based Sampling for In-Person Surveys

- Address-Based Sampling (ABS) frames derived from USPS Computerized Delivery Sequence (CDS) File
- Alternative to traditional field enumeration (FE)
  - Cost savings
  - More timely
- Household coverage in urban areas is reasonable, but coverage in rural areas remains problematic[1]
- Urban (~95%); Rural (~72%)[2]

[1] Dohrmann et al. 2007; Iannacchione et al. 2007; O'Muicheartaigh et al. 2007

[2] Shook-Sa and Currivan, 2011

RTI INTERNATIONAL

Primary Sources of Undercoverage In-Person Surveys

1. Unlocatable Mailing Addresses
2. Geocoding Error
3. Addresses excluded from CDS file

USPS No-Stat File

- The USPS No-Stat file is a supplemental list to the CDS file containing 7.0 M residential, locatable addresses
  - Active Addresses (~9%)
  - Non-Active Addresses (~90%)
  - Overlap with the CDS file (~1%)
- Addresses maintained on the No-Stat file for USPS administrative purposes
- Made available to companies with CDS licenses starting in 2009

RTI INTERNATIONAL

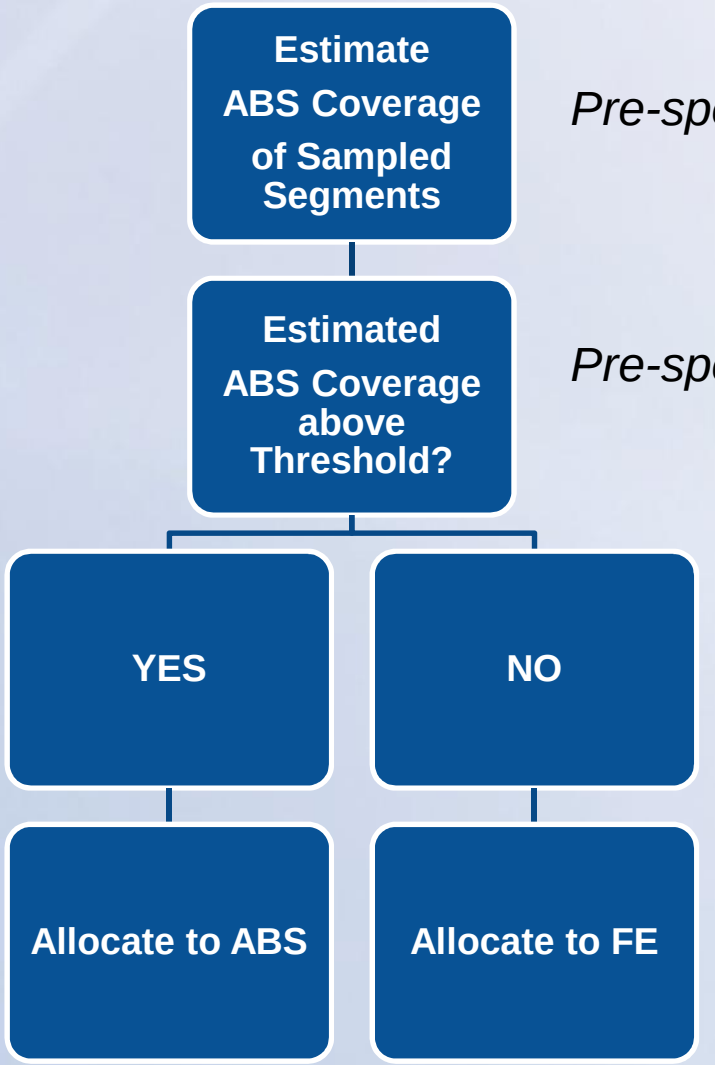
USPS No-Stat File: Prior Research Findings[1]

- Occupancy Rates
  - 21% No-Stat addresses (1.3 M)
  - 90% of CDS addresses
  - 87% combined CDS and No-Stat frame
- Addresses are clustered in primarily rural areas
- Provides modest coverage gains in rural areas (~4%)
- Hybrid frame still needed

[1]Shook-Sa et al. 2012

RTI INTERNATIONAL

Hybrid ABS/Field Enumerated Frames



Pre-specified coverage estimator

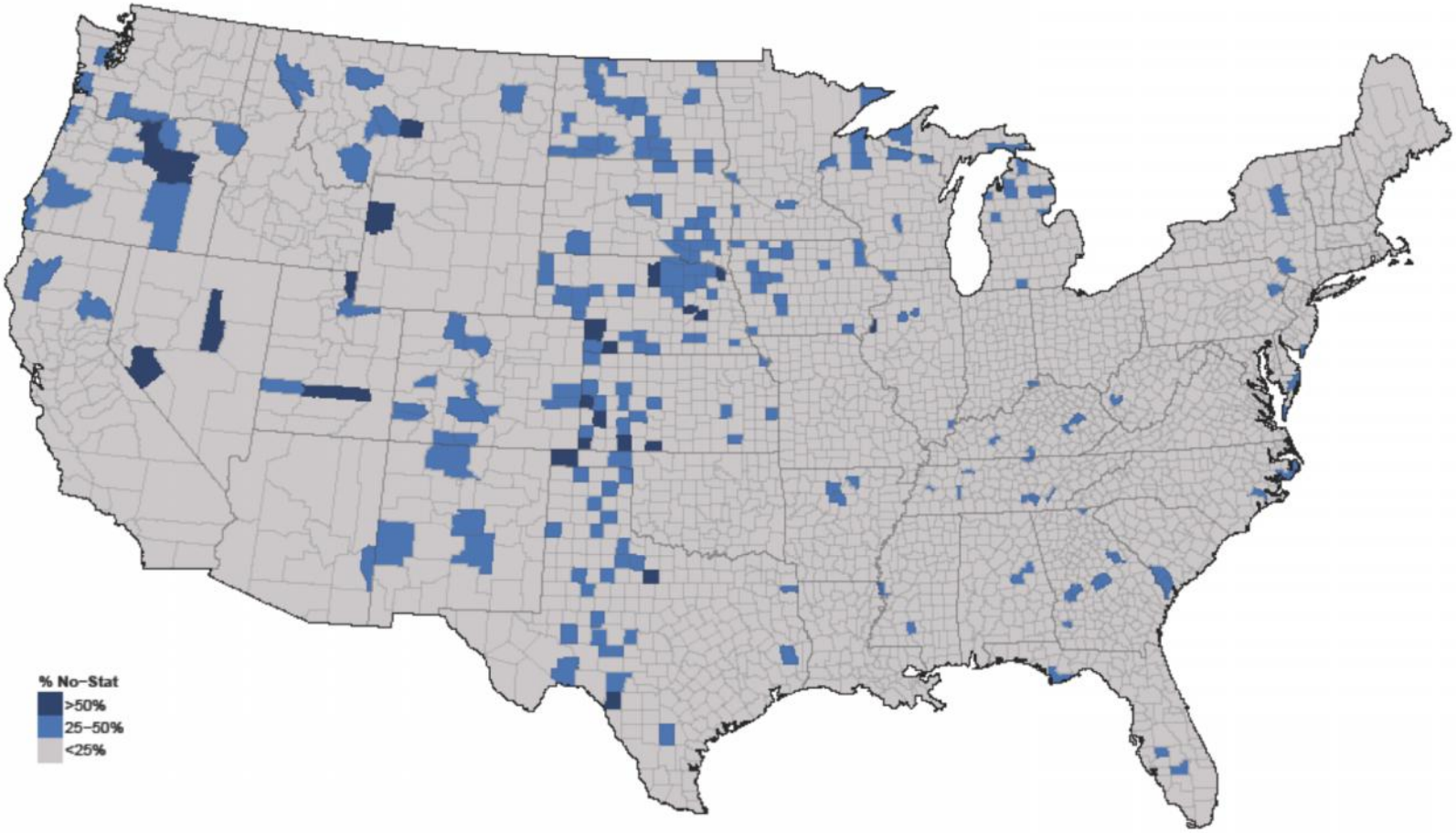
Pre-specified ABS coverage threshold

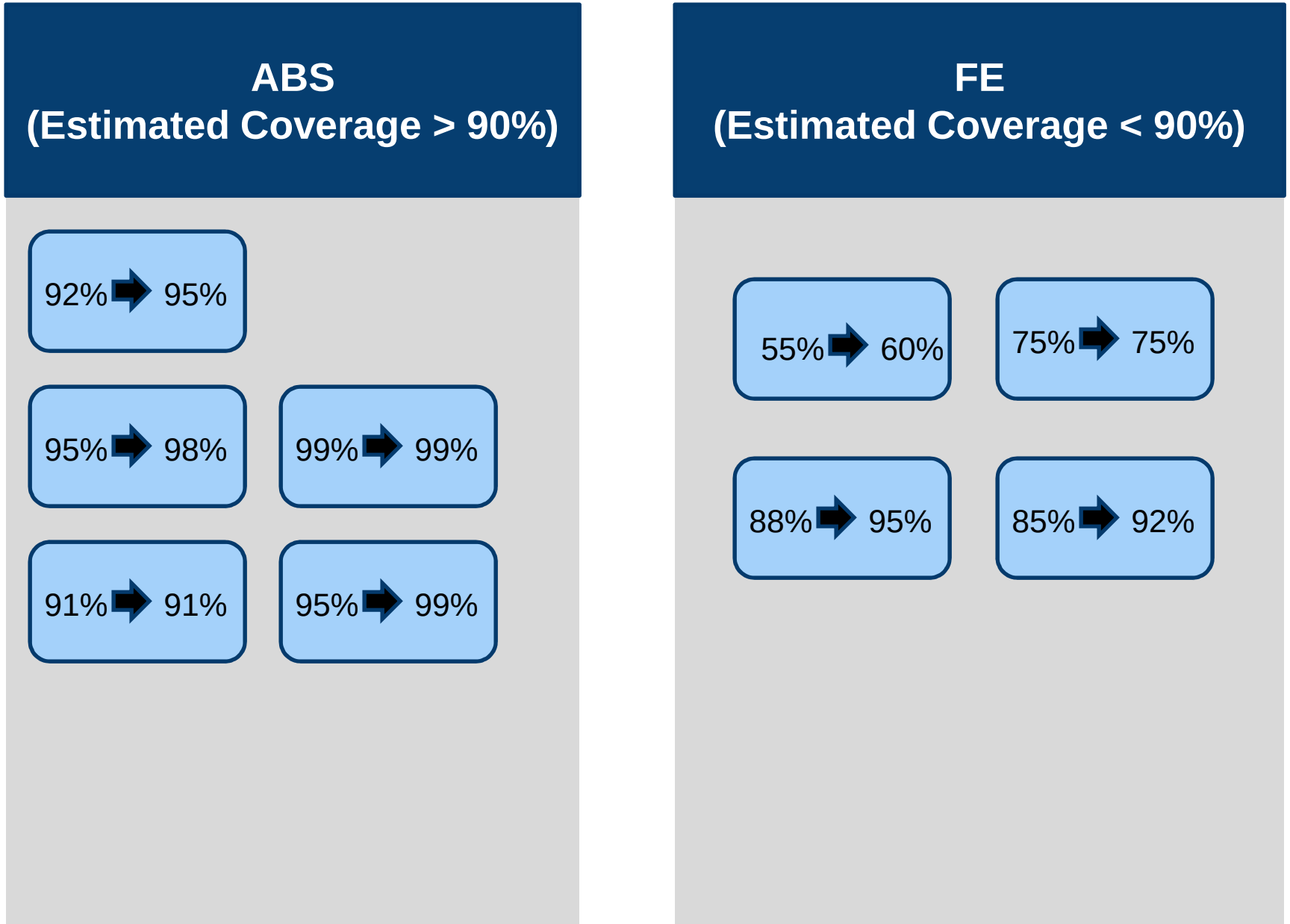
Hybrid ABS/Field Enumerated Frames (cont.)

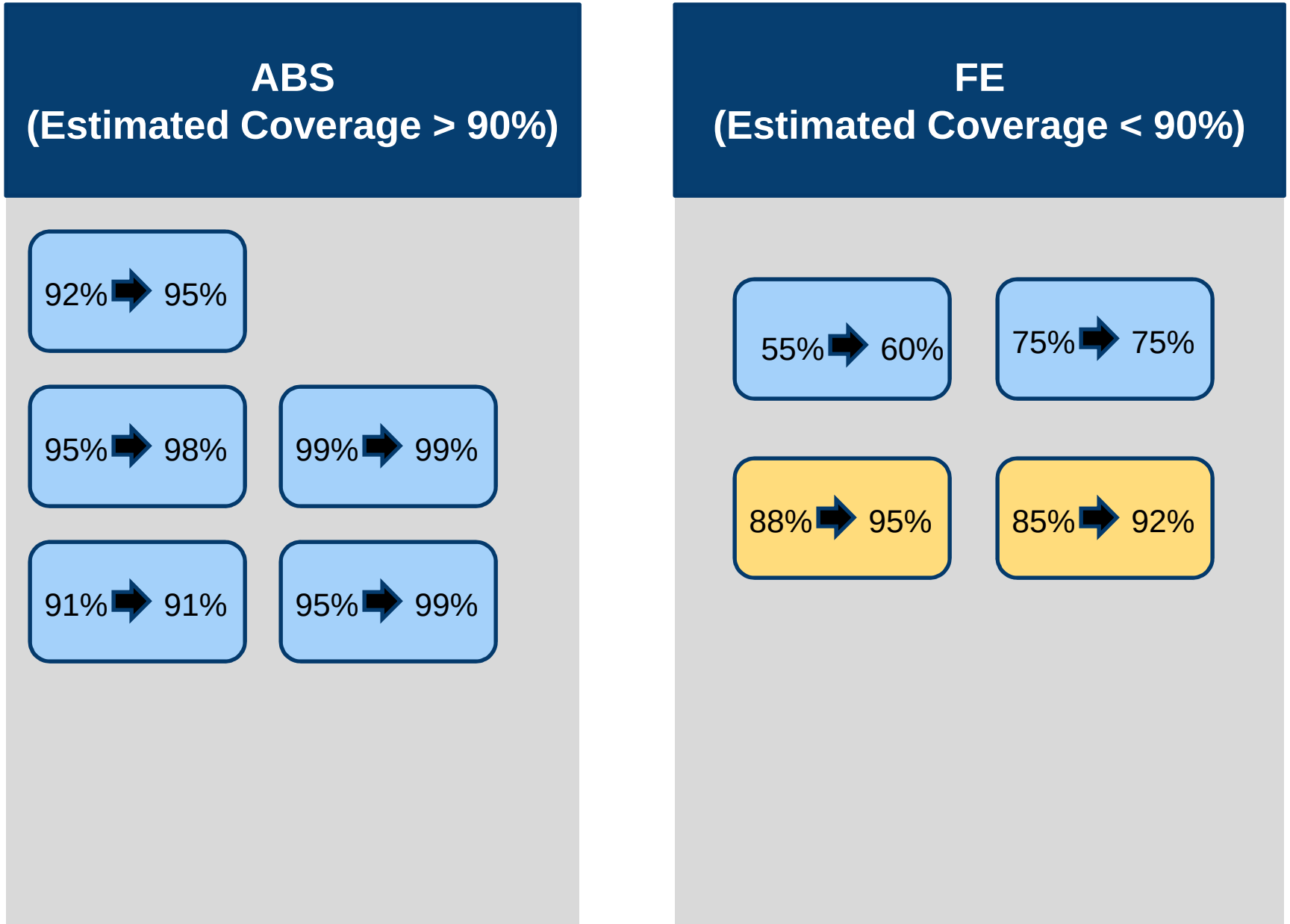
- Provides cost savings of ABS when possible
- Retains FE only where necessary to maintain coverage
- The ABS coverage threshold is not the expected coverage of the hybrid frame
- Field supplementation might still be needed
- The more segments allocated to ABS, the higher the cost savings
- The No-Stat file has the potential to move segments that otherwise would rely on FE to the ABS portion of the frame

RTI INTERNATIONAL

% Locatable Addresses from No-Stat File




ABS
(Estimated Coverage > 90%)
92% 95%
95% 98%
99% 99%
91% 91%
95% 99%
FE
(Estimated Coverage < 90%)
55% 60%
75% 75%
88% 95%
85% 92%


ABS
(Estimated Coverage > 90%)
92% 95%
95% 98%
99% 99%
91% 91%
95% 99%
FE
(Estimated Coverage < 90%)
55% 60%
75% 75%
88% 95%
85% 92%

Hybrid Frame Evaluation

- Compared two potential sampling frames
  - CDS Only (traditional ABS frame)
  - CDS + No-Stat
- Constructed a national frame of area segments
  - Geocoded March 2013 CDS and No-Stat addresses into Census Block Groups (CBGs)
  - Collapsed CBGs to meet minimum size criteria
- Evaluated a national PPS Design under the two frames
- Size measure = Population 18+

RTI INTERNATIONAL

Hybrid Frame Evaluation (Cont.)

- Specified coverage model:

$$\text{Estimated ABS Coverage} = \frac{\text{\# Locatable Addresses}}{\text{\# Number of Housing Units}}$$

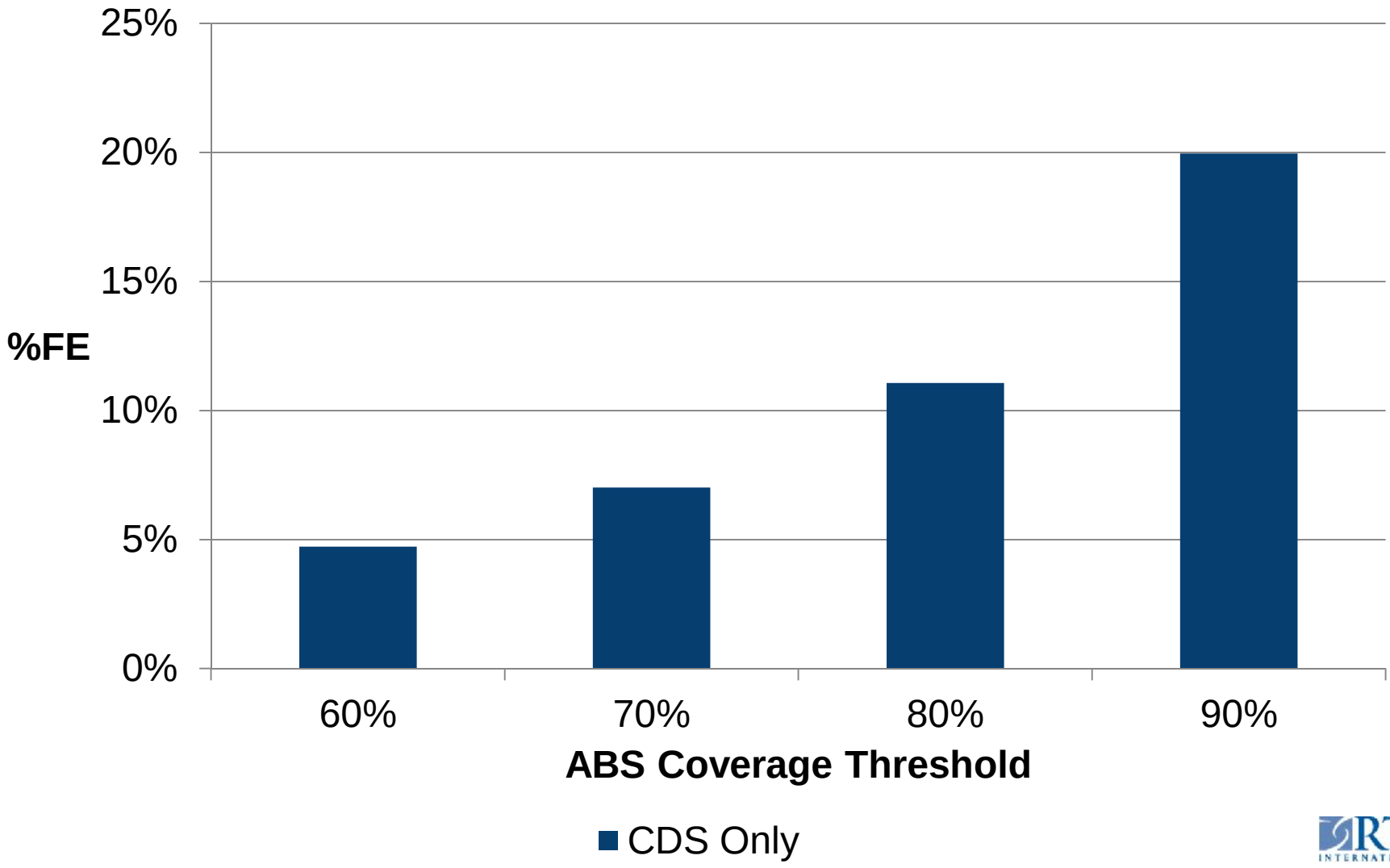
- Examined the expected distribution of segments across ABS and FE for several potential ABS coverage thresholds
- National and State-Level estimates of No-Stat savings

RTI INTERNATIONAL

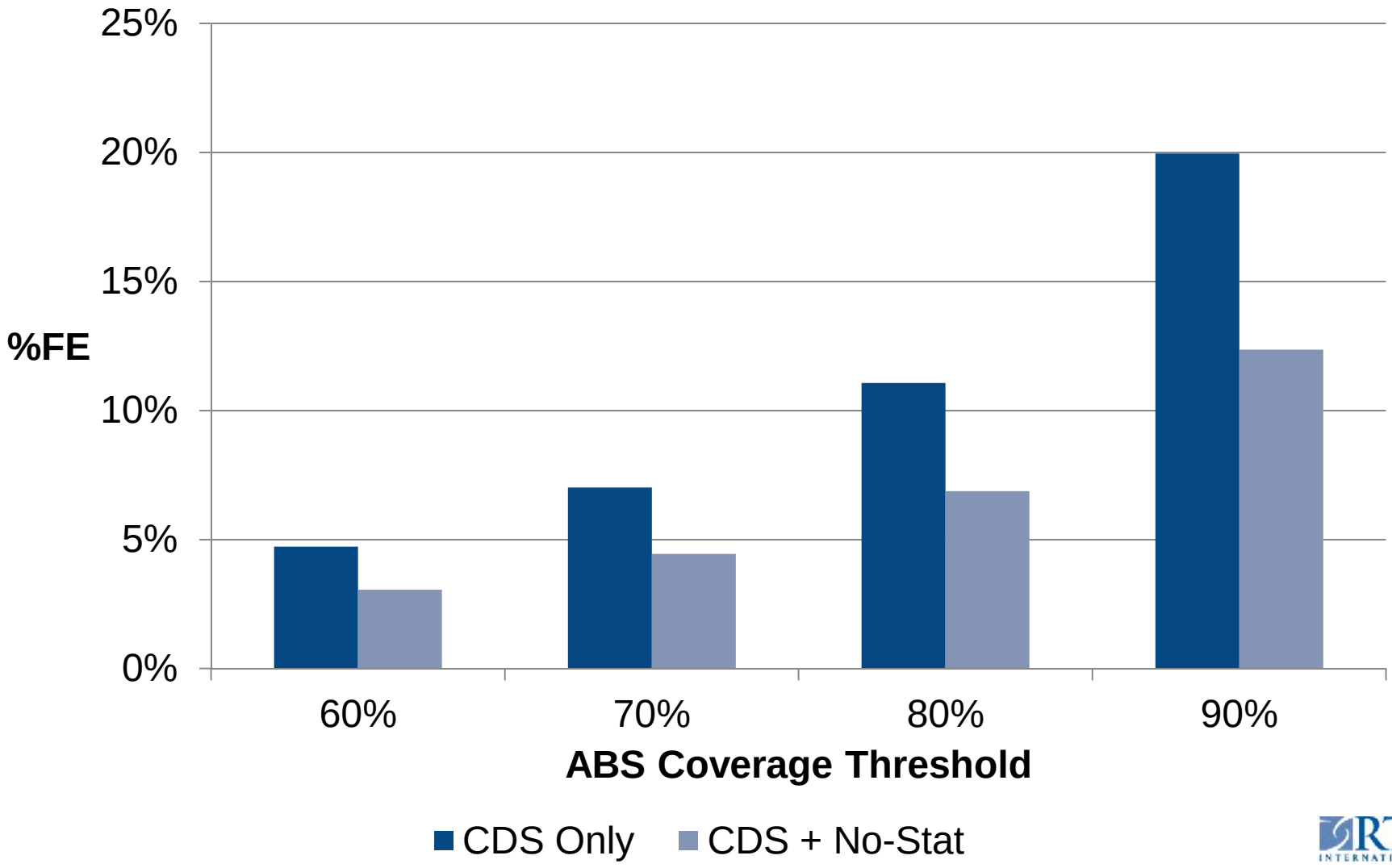
Hybrid Frame Properties

- 215k segments
- 63% contain one or more No-Stat addresses
  - Range from 0.01% No-Stat to 100% No-Stat
  - Median 4%

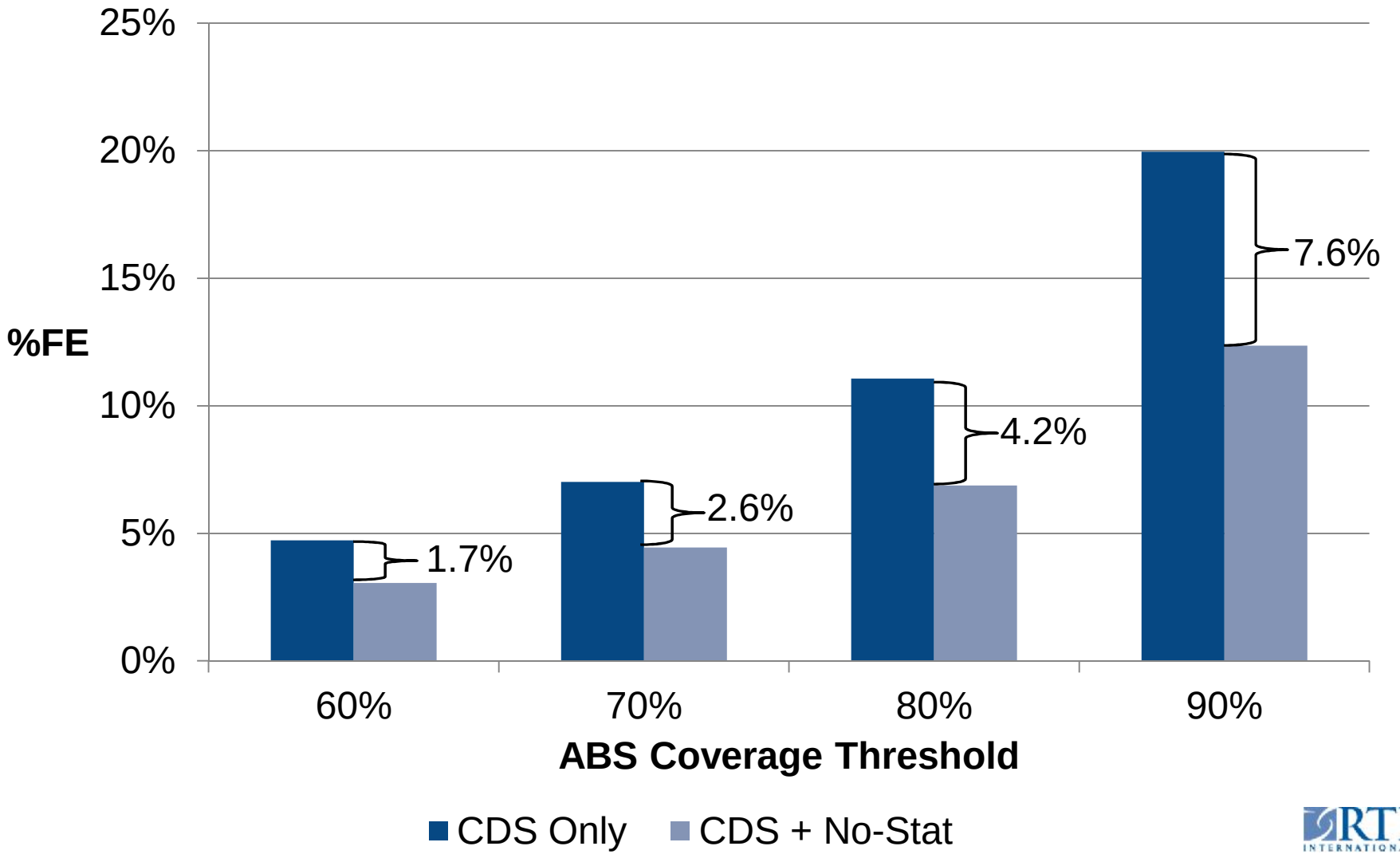
% Segments Allocated to FE by ABS Coverage Threshold



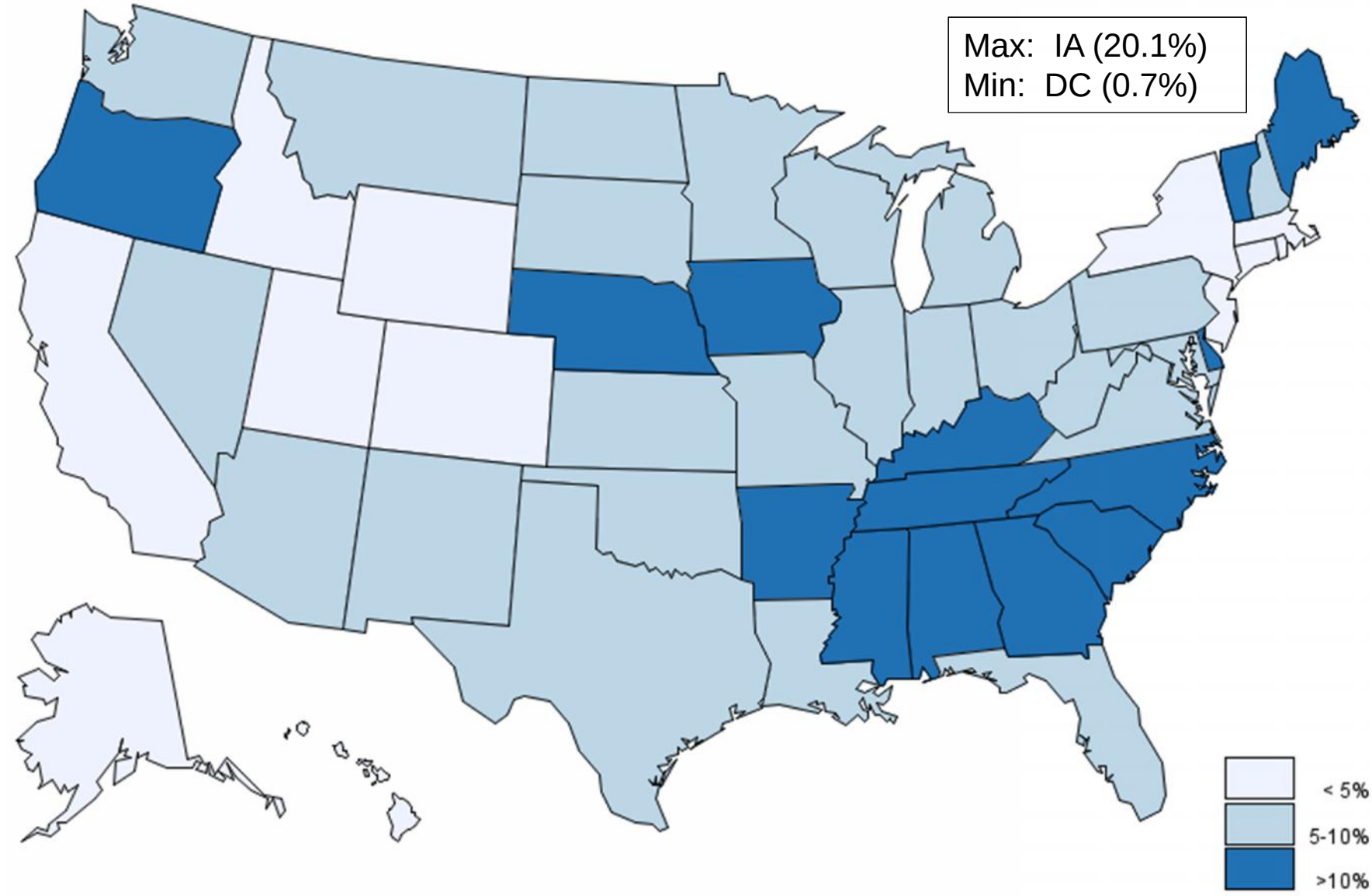
% Segments Allocated to FE by ABS Coverage Threshold



% Segments Allocated to FE by ABS Coverage Threshold



% Segments No-Stat Moved to ABS, by State (90% ABS Cov. Threshold)



Conclusions

- The No-Stat file improves the efficiency of ABS frames for in-person surveys by moving segments that would rely on FE to the ABS portion of the frame
- Efficiency gains are greater for higher coverage thresholds
- Efficiency gains vary widely across states

RTI INTERNATIONAL

Limitations

- Efficiency gains are dependent on the coverage predication model and the ABS coverage threshold specified
- Vary based on sample design
- Results are limited to in-person household surveys

Future Research

- No-Stat addresses for multi-mode designs
  - Mailing to No-Stat addresses
  - Phone append and accuracy rates for No-Stat addresses for mixed mode designs
- Accuracy of No-Stat indicators

Contact

Bonnie Shook-Sa

Research Statistician

RTI International

bshooksa@rti.org