Considerations when using Address-Based Sampling (ABS)

Joe McMichael

Survey Statistician

International Field Directors & Technologies Conference

May 22, 2012

Quick Poll

This is a very difficult audience to know…

Why are you here?

Who has taken an ABS course or webinar? Which one?

Who is using/used an ABS design?

ABS is Hot

Address-Based Sampling is changing the ways we as survey researchers think about and conduct surveys.

- Many well attended sessions (APPOR, JSM and IFD&TC)
- Publish papers (POQ, Survey Practice, Field Methods, conference proceedings)
- Courses
- Design changes for many large surveys:
  - National Survey of Family Growth
  - American National Election Study
  - General Social Survey
  - Nielsen TV Dairy
  - Knowledge Networks now recruits households for its KnowledgePanel
  - many more…

What is ABS?

- What is Address-Based Sampling (ABS)?
  - Frame or Methodology?
  - A sampling frame consisting of a list of addresses derived from the United States Postal Service (USPS)
  - Addresses serve as a surrogate to housing units
  - 123+ million addresses
- CDS File or DSF^2?
  - DSF^2 – Delivery Sequence File ☹
  - CDS File - Computerized Delivery Sequence File ☺
- Valassis & Compact Information Systems (CIS)

Basic CDS file data elements

- Street name and number/box number
- City, state, nine-digit ZIP Code
- Carrier route number
- Delivery-sequence number
- Vacant address indicator
- Seasonal delivery flag
- Address Type Indicators (curb, cluster box, door slot, etc.)
- Residential/Business indicator
- Drop Count

What is typically appended to the CDS?

- Geocoded Information
  - Latitude / Longitude
  - Census Geography (Block, Tract, etc…)
  - Geocode Accuracy Indicator
- Area Demographics
  - Usually Census-based

What else can be appended?

- Anything that can be appended to an address
  - Name(s)
  - Phone Number
  - Email

- Age, Race, Ethnicity, Income, Education, Marital Status, # Children, # Adults, Rent or Own, Presence of specific age groups, and more…

- Little is known about the accuracy of these data
- Sources need to be identified and studied

Why ABS?

Depends largely on mode and target population

- High Coverage
- Lower Cost
- Speed?
- Ability to target small geographies
- Multi-mode options
- Ancillary data leads better stratification

- Everyone else is doing it…

ABS Contact Method & Survey Mode

Contact Method

- Mail
- In-person
- Phone
- Email
- Carrier Pigeon

Survey Mode

- PAPI (self-administered)
- CATI (in-bound)
- CATI (out-bound)
- Web
- CAPI (interviewer administered)

Contact Method: Mail

Modes:

- PAPI (self administered)
- Web
- CATI in-bound
- Others (ask for phone, send cell phone)

Advantages:

- Very high coverage

Disadvantages:

- Self-administered (may not be a problem)

Special considerations:

- Drop units and throwbacks
- How to address?
- OWTGM

Contact Method: In-person

Modes:

- CAPI
- CATI in-bound

Advantages:

- Response Rate

Disadvantages:

- Cost
- Coverage

Special considerations:

- Unlocatable addresses
- Clustering (larger clusters and geocoding error)

Contact Method: Phone

Modes:

- CATI out-bound

Advantages:

- Systems already in place
- Many surveys already designed for this mode
- Lower Cost

Disadvantages:

- Coverage
- Bias likely

Special considerations:

- Matching (reduces coverage, complicates, more bias)

Mixed-Mode

- PAPI & CATI
  - Concerns about mode effects
    - Self-administered vs. interviewer-administered
    - Visual vs. auditory
  - Concerns about phone coverage bias
  - Design to measure bias
- PAPI & Web
  - Concerns about mode effects
  - Coverage bias from internet access

Hybrid Frames (ABS + ???)

Field Enumerated:

- Traditional C&L
- List Assisted, Dependent, or Enhanced Listing
- GeoFrame

ABS Frame Supplementation:

- Field Methods
  - CHUM
  - Waksberg Approach
- List Methods
  - NoStat File
  - Other Address Sources

Others:

- Aerial or Satellite Listing

Odds and Ends

- Census vs. Postal Geography
- Geocoding Error and ABS
- ABS Coverage Prediction

Census vs. Postal Geography

- Postal Geography
  - Not a geography at all
  - Created to make efficient mail delivery
  - Zip, carrier routes, zip+4
- Census Geography
  - Partitions the U.S. into well defined areas
  - Created to canvas the US and report summary statistics
  - Block, Block group, Tract, MSA, CDP, PUMA, many more…
  - ZCTAs (Zip Code Tabulation Areas)
- Crosswalks are inexact
- Which one to use depends on your needs

Geocoding Error and ABS

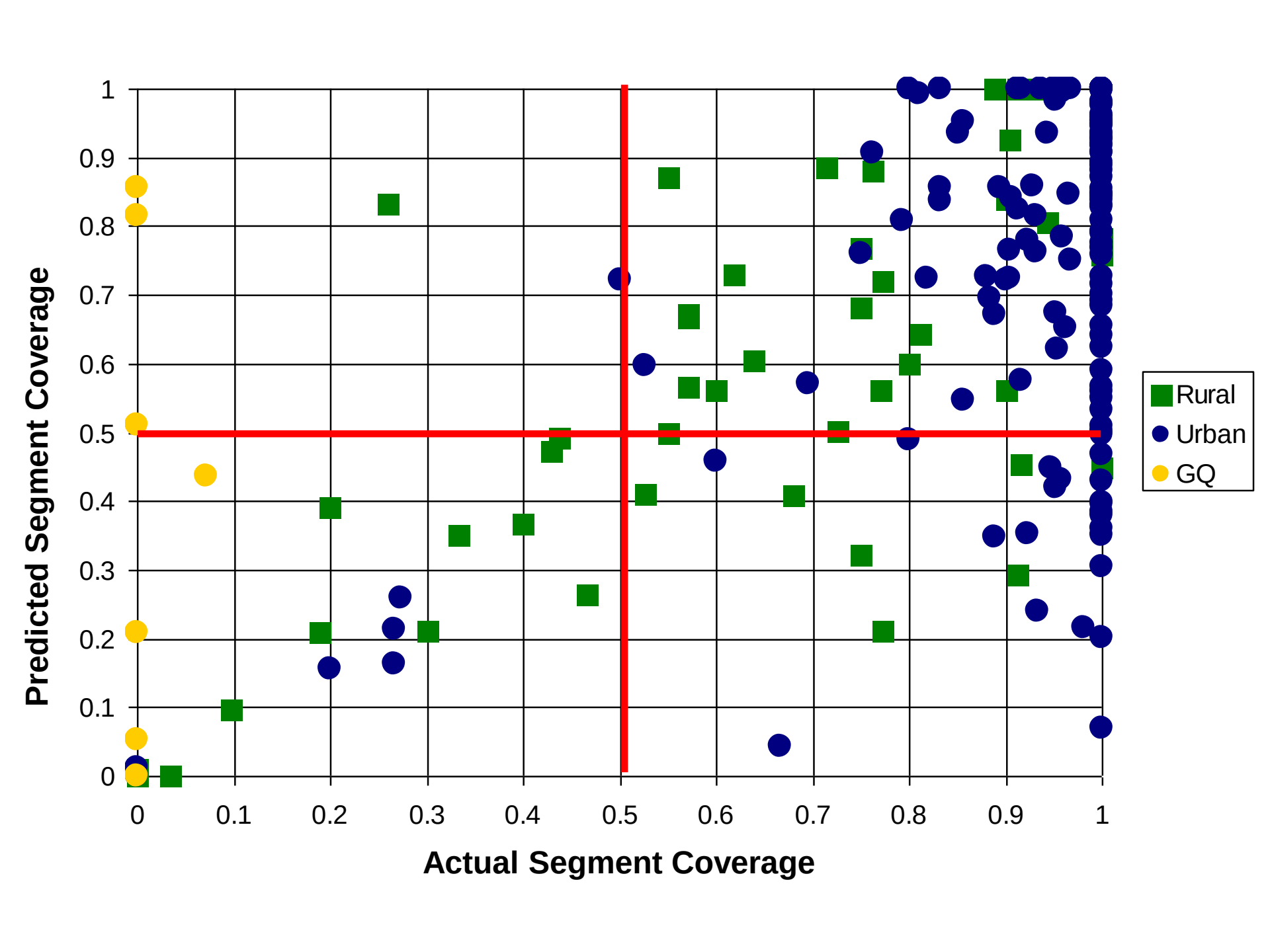
- Serious problem for small clusters
- Unlocatable addresses (PO Boxes, RR/HC)
- Rural Areas

Stratification

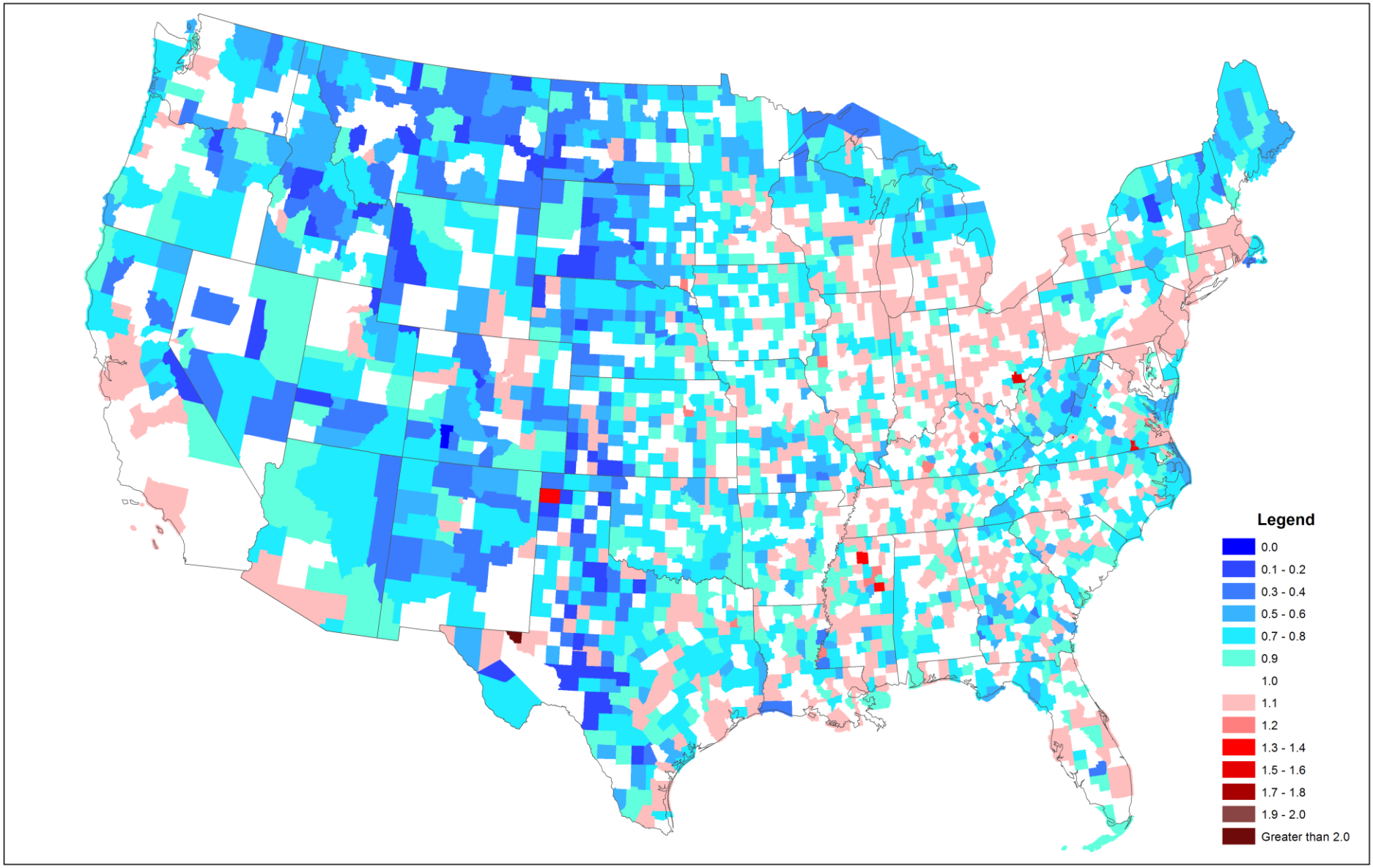
- Very promising
- Census demographics
- Surname for Hispanic and Asian populations

Coverage Prediction

- Model based
- Naive: Census to ABS ratio
- Over coverage vs. Under coverage

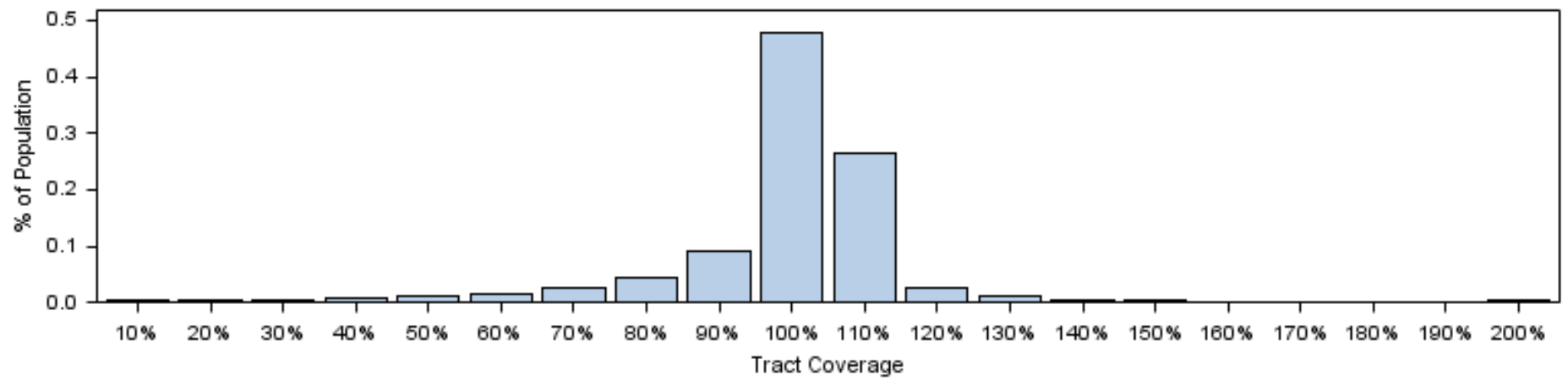

Predicted Segment Coverage
Actual Segment Coverage
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Rural
Urban
GQ

County Coverage Map

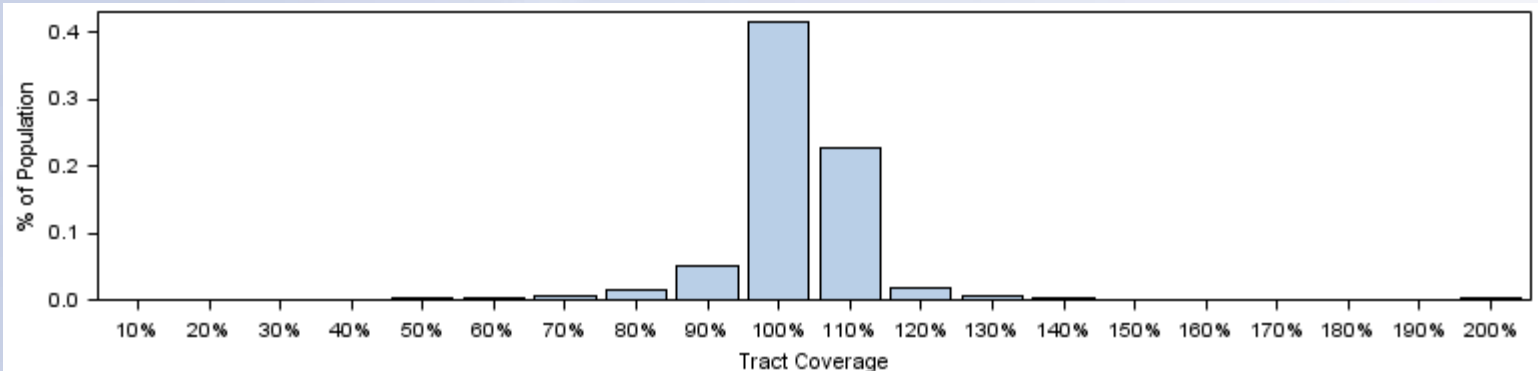


National Tract Coverage by % Population

- 90% of population lives in tracts with 90-110% coverage

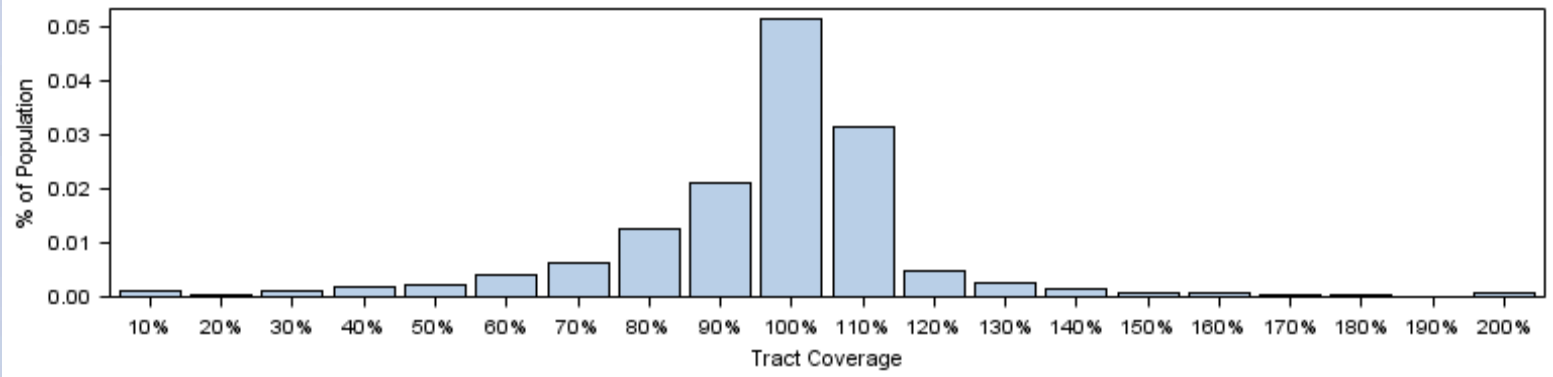


Urbanicity – Tract Coverage by % Population



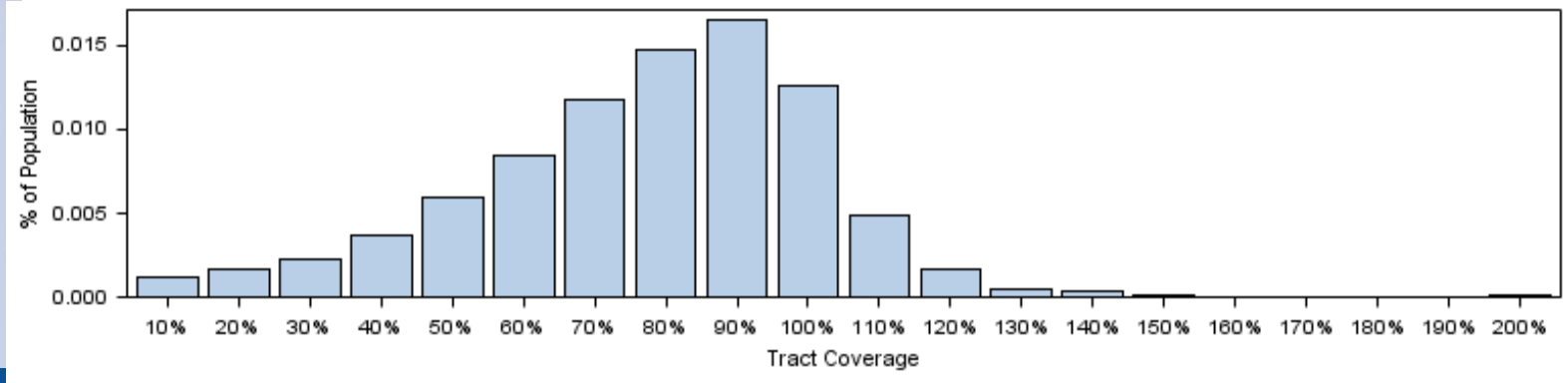
Urban

78% population



Some Urban

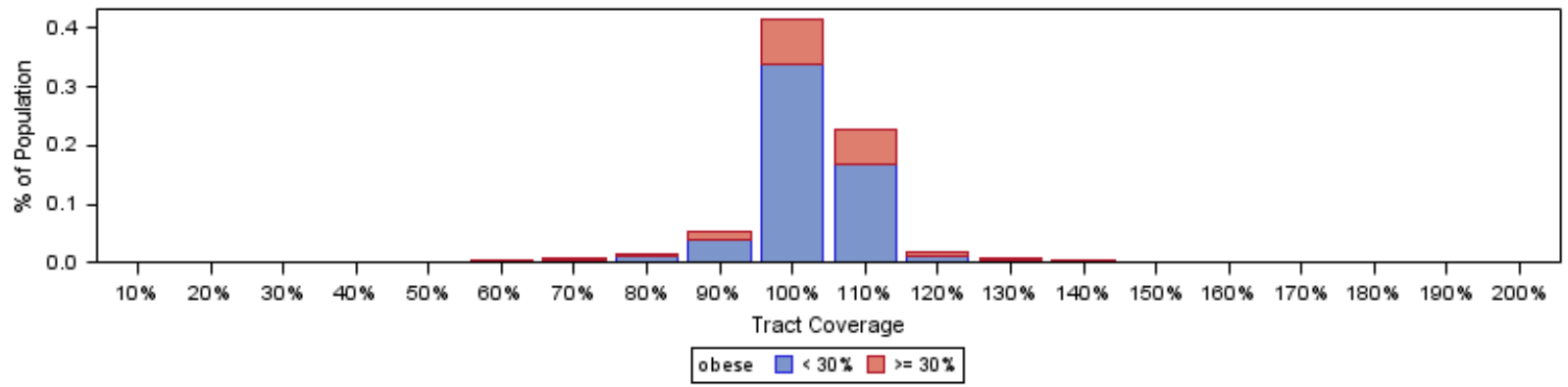
15% population



Rural

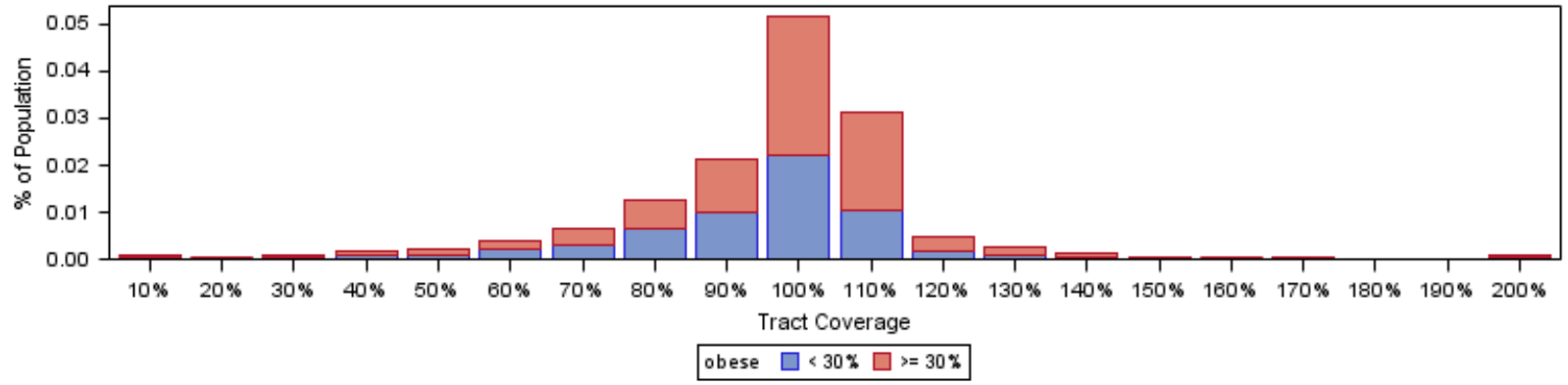
8.6% population

Obesity and Urbanicity



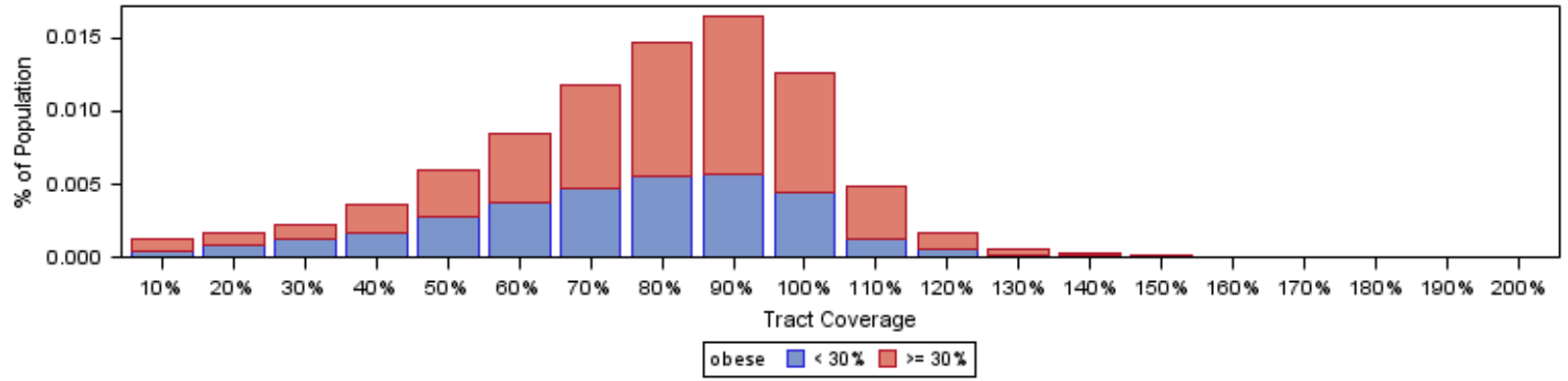
Urban

78% population



Some Urban

15% population



Rural

8.6% population

RTI INTERNATIONAL

Contact Information

Joe McMichael
Statistician
919.485.5519
mcmichael@rti.org