Mode Differences in a Mixed-Mode Address-Based Survey Design

René Bautista
Ting Yan
Rupa Datta
Peter Hepburn

AAPOR 2011 Conference
Phoenix, Arizona

Outline

Mode Differences in a Mixed-Mode Address-Based Survey Design

- Background
  - Address-Based Surveys (ABS)
  - Mixed Mode Designs (MM)
  - Address-Based Surveys (ABS)+Mixed Mode Designs (MM)
- Research Question
- Data and Methods
- Results
- Conclusions

Background

- Address-Based Sampling (ABS)
  - Aim: Decrease coverage error in surveys
  - Extensive use of the US Postal Service (USPS) Delivery Sequence File (DSF)
  - Telephone numbers can be matched to DSF-sampled addresses
  - USPS-DSF as an alternative to:
    - RDD sampling frame construction
    - Listing in multi-stage area surveys
  - ABS designs facilitate implementation of mixed mode data collection methods
Background (cont’d)

Mixed-Mode Designs (MM)

- Aim: Increase both likelihood of participation and coverage of the population of interest
- Use of multiple modes to approach and conduct interviews in a single study
  - For instance, a mail survey with telephone follow-ups
- Widely used in survey research
- Experimental research to study potential mode effects
  - Random allocation of respondents to different modes
  - In practice, extensive use of sequential mixed modes
  - Cost efficiency

Address-Based Sampling+ Mixed-Mode Design (ABS+MM)

- Little is known about survey designs using both address-based sampling and mixed mode data collection strategies
- Unclear to what extent an ABS+MM design helps recruit hard-to-reach populations
  - “Mode differences” — In the context of this study:
  - Data Collection Strategy Differences

Research Question

- Are there different recruitment patterns among Hard-to-Reach populations when using one mode or another in an ABS+MM design?
  - What types of respondents are recruited by which modes?
DATA AND METHODS

Survey Data

2010 Census Integrated Communications Program Evaluation (CICPE)  
Address-Based Sampling + Mixed Mode Design (ABS+MM)

<table>
<thead>
<tr>
<th>Wave 1</th>
<th>Wave 2</th>
<th>Wave 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sample Size</td>
<td>N=2,671</td>
<td>N=668</td>
</tr>
<tr>
<td>Hispanic</td>
<td>461</td>
<td>118</td>
</tr>
<tr>
<td>Non-Hispanic African American</td>
<td>377</td>
<td>111</td>
</tr>
<tr>
<td>Non-Hispanic Whites &amp; Other</td>
<td>404</td>
<td>99</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>457</td>
<td>107</td>
</tr>
<tr>
<td>Asian</td>
<td>542</td>
<td>114</td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>430</td>
<td>119</td>
</tr>
</tbody>
</table>

Fieldwork period
- Sept 9, 2009 - Jan 16, 2010
- Jan 19 - Mar 14, 2010
- April 19 - July 5, 2010

All cross sectional data n=4,879

Data Collection Strategy

2010 Census Integrated Communications Program Evaluation (CICPE)  
Address-Based Sampling + Mixed Mode Design (ABS+MM)

Selected Sample → Phone numbers matched to ABS addresses

Yes → Telephone (CATI) → Tel & Field

No → Field Only (Paper & Pencil)

Non Interviewed (Sub sampling) Tel & Field: tel #'s matched but not completed on the phone (Harder cases)
Field Only: tel #'s not matched, sent straight to the field
Regression Model

- Multinomial Logistic Model

**Dependent variable (Y):**
- Race/Ethnicity
  - Hispanic
  - African American
  - White (Ref cat)
  - American Indian/Alaska Native
  - Asian
  - Native Hawaiian & P Island
- Homeownership
  - Renter
  - Owner (Ref cat)
- Age
  - Younger (<45 yrs)
  - Older (45 yrs+) (Ref cat)
- Education
  - High School
  - Some college or more (Ref cat)
- Income
  - <50K or less
  - =50K+

**Independent variables (X’s):**
- Data Collection Strategy
  - Telephone only (Ref cat)
  - Telephone and Field
  - Field Only

Descriptive Statistics

- All cross-sectional data (n=4,879)

**Data Collection Strategy**
- Telephone Only: 22%
- Tel & Field: 42%
- Field: 36%

RESULTS
Multinomial Logistic Model

- All cross sectional data (n=4,879)

<table>
<thead>
<tr>
<th>Race / Ethnicity</th>
<th>Tel &amp; Field vs. Telephone</th>
<th>Field Only vs. Telephone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hispanic</td>
<td>0.80**</td>
<td>0.38</td>
</tr>
<tr>
<td>Non-Hispanic African American</td>
<td>0.48*</td>
<td>0.13</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>-0.41**</td>
<td>1.27***</td>
</tr>
<tr>
<td>Asian</td>
<td>0.42†</td>
<td>0.77***</td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>1.14</td>
<td>0.66***</td>
</tr>
<tr>
<td>Age (&gt;45 yrs = 1)</td>
<td>0.66***</td>
<td>0.66***</td>
</tr>
<tr>
<td>Education (High School=1)</td>
<td>0.36*</td>
<td>0.24*</td>
</tr>
<tr>
<td>Homeownership (Renter=1)</td>
<td>0.44***</td>
<td>0.27†</td>
</tr>
<tr>
<td>Income (&lt;59K = 1)</td>
<td>0.35**</td>
<td>0.19†</td>
</tr>
</tbody>
</table>

* p<.10, † p<.05, ** p<.01, *** p<.001

First Differences in Predicted Values

- "Discrete Changes" w/ 95% Confidence Intervals
- Difference Between Two Predicted Values
- Example:

Example:

Pr (Y=Tel Only | X)

Hispanics: 14% vs. 23%
African Americans: 18% vs. 23%
Non-Hispanic African Americans: 6% vs. 9%
American Indian/Alaska Natives: -7% vs. 16%
Asians: 4% vs. 17%
Pacific Islanders: -5% vs. 5%
Age (>45 yrs = 1): 4% vs. 6%
Education (High School=1): 2% vs. 3%
Homeownership (Renter=1): -5% vs. 5%
Income (<59K = 1): 5% vs. 15%

FrD (Y=Tel Only | X)

Hispanics: -10% vs. 0%
African Americans: -11% vs. 5%
Non-Hispanic African Americans: 2% vs. 8%
American Indian/Alaska Natives: -12% vs. 6%
Asians: -40% vs. -30%
Pacific Islanders: -5% vs. 0%
Age (>45 yrs = 1): -3% vs. 0%
Education (High School=1): -18% vs. -11%
Homeownership (Renter=1): -10% vs. 10%
Income (<59K = 1): -25% vs. -15%

Race / Ethnicity

- First Differences (All cross sectional data)
Demographic Characteristics

First Differences (All cross sectional data)

FrD (Y=Tel Only | X)  FrD (Y=Tel & Field | X)  FrD (Y=Field only | X)

Subgroups Most Likely to Be Recruited by Each Data Collection Strategy (Summary)

Address-based survey + Mixed-mode design (ABS-MM)

Data Collection Strategy       Subgroup

Telephone                      NonHispanics Whites & Other
  Tel #’s matched and completed on the phone

Telephone & Field             Hispanics
  (Tel #’s matched but not completed on the phone, subsample sent to the field (Harder cases)

Field Only                    American Indian
  Tel #’s not matched, sent straight to the field

  Younger
  Renter
  Low income

Conclusions & Discussion

• ABS+MM: cost efficient sampling strategy
• 2010 CICPE survey:
  – 1/4 of the complete cases → Telephone Only
  – 3/4 of the complete cases → Tel & Field and Field Only
• As expected
  • Hard-to-Reach populations: Harder cases (Tel & Field)
• Further research needed to validate results
Thank You

• Bautista-Rene@norc.org