The Undercounted: How NORC at the University of Chicago built the probability-based AmeriSpeak Panel to improve representativeness of America’s most difficult-to-reach populations

University of Michigan: February 12, 2020

Michael Dennis
David Dutwin
Dan Costanzo

The NORC Presentation Team

J. Michael Dennis, Ph.D.
- Executive Director, AmeriSpeak
- SVP, NORC
- Probability panel expert

David Dutwin, Ph.D.
- Chief Scientist, AmeriSpeak
- SVP, NORC
- Past President, AAPOR
- Expert in scientific use of survey samples

Dan Costanzo
- Senior Manager, AmeriSpeak
- Business Development
- Academic sector lead
- 10+ years telephone research center experience

Outline

- AmeriSpeak: America’s High Response Rate Probability Panel
- Mode Effects: A Unique Test Using the AmeriSpeak Panel
- The Impact of NonResponse Follow-Up on AmeriSpeak Data Quality

All About AmeriSpeak
AmeriSpeak By the Numbers

- Number of Participating Households ➔ 35K
  (50 States + DC)
- Client Surveys Completed ➔ 300+
  (Since June 2015)
- Panel Recruitment Response Rate ➔ 34%
  (AAPOR RR3)

AmeriSpeak.NORC.org/research

How Does AmeriSpeak Address the Challenge of Representing All of the Population?

2–STAGE RECRUITMENT

2ND stage selects segments within each NFA (1,514 selected)
- DSF used for addresses, supplemented with in-person enumeration in rural areas to increase coverage to 97%
- NFAs with at least 1.5 million selected with certainty
- At least one NFA per state
- Segments with higher % of young adults and minorities (Hispanic, NH-Black) oversampled
- High Hispanic tracts oversampled

AmeriSpeak Sampling Frame: NORC National Frame

- Uses the NORC National Frame
  - Area probability design based on 126 national frame areas (NFAs)
  - 2ND stage selects segments within each NFA (1,514 selected)
  - DSF used for addresses, supplemented with in-person enumeration in rural areas to increase coverage to 97%
  - NFAs with at least 1.5 million selected with certainty
  - At least one NFA per state
  - Segments with higher % of young adults and minorities (Hispanic, NH-Black) oversampled

AmeriSpeak Recruitment Methodology

- Initial Recruitment
  - Pre-notification postcard
  - +15 days, 9 x 12 recruitment packet w/ pre-incentive, study brochure, and privacy policy
  - +18 days, reminder postcard
  - Call-ins allowed throughout
  - +25 days, call-outs to matched telephones

- NRFU Recruitment
  - Federal Express study brochure and enhanced pre-incentive
  - In-person recruitment
How do we sample for NRFU?

- ~1 in 5 non-respondents are subsampled for NRFU
- We use consumer vendor data to target and oversample HHs that are likely to have
  - Young adults (18-34 years)
  - Hispanic, non-Hispanic Black
- Benefits:
  - Significantly improves our recruitment rate; from 6% to 33-35%
  - Brings in younger, Hispanic & NH-Black, High school or less panelists
- Cost:
  - Each NRFU recruit is ~1.5x more costly than an initial recruit
  - Slightly increases the design effect

Response Rates

- Panel Recruitment Rate: 34%
- Panel Retention Rate: 88-89%
- Survey Completion Rate: Varies depending on survey; 25-45%
- Cumulative Response Rate: Product of the above; approximately 9-15%

Panel Weighting Procedure

- HH base weight
  - Selection probability of addresses
  - Combine different sample sources
  - Combine different years
  - NRFU (non-response follow up)
  - Unknown eligibility
  - Ineligible: businesses, vacation homes, vacant properties, no one 18+ in HH, etc.
  - Eligible: HHs with 18+
  - Non-response among eligible HHs
Panel Weighting Procedure (cont.)

- Panelist final weight
  - Non-response among eligible 18+
    - About 2:1 eligibles per HH, we recruit only 1:1
  - Raking
    - Current raking dimensions:
      - Age group (18-24, 25-29, 30-39, 40-49, 50-59, 60-64, 65+ years);
      - Sex (male, female);
      - Census Division;
      - Education (less than high school, high school graduate, some college, college graduate);
      - Race/ethnicity (Hispanic, non-Hispanic Black, non-Hispanic White, All Other);
      - Housing tenure (owner, renter, other); and
      - Household phone status (cell-only, dual user, landline-only/phoneless).
    - Looking into raking by age * gender interactions
    - Raking to more variables or categories of variables increases DEFF

AmeriSpeak Panel Demos vs. Benchmark

AmeriSpeak Panel Demos vs. Benchmark (cont.)

AmeriSpeak Panel Demos vs. Benchmark (cont.)
AmeriSpeak Panel Demos vs. Benchmark (cont.)

Race/Ethnicity

- NH-White
- NH-Black
- Hispanic
- NH-Asian
- NH-All Other

Benchmark
AmeriSpeak Unweighted
AmeriSpeak Weighted

Gender

- Male
- Female

Benchmark
AmeriSpeak Unweighted
AmeriSpeak Weighted

AmeriSpeak Panel Demos vs. Benchmark (cont.)

Presence of Children

- With 1+ Children in HH
- Without children under 18

Benchmark
AmeriSpeak Unweighted
AmeriSpeak Weighted

AmeriSpeak Profile-based Estimates vs. Benchmarks

<table>
<thead>
<tr>
<th>Estimate</th>
<th>AmeriSpeak</th>
<th>Benchmark</th>
<th>Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ever Told You Have A Stroke</td>
<td>2.5%</td>
<td>3.2%</td>
<td>-0.7%</td>
</tr>
<tr>
<td>Ever Told You Have Cancer (Other Than Skin Cancer)</td>
<td>9.7%</td>
<td>9.8%</td>
<td>-0.1%</td>
</tr>
<tr>
<td>Ever Told You Have A Depressive Disorder</td>
<td>18.5%</td>
<td>19.0%</td>
<td>-0.5%</td>
</tr>
<tr>
<td>Ever Told You Have Diabetes</td>
<td>12.5%</td>
<td>11.8%</td>
<td>0.7%</td>
</tr>
<tr>
<td>Ever Told Blood Pressure High</td>
<td>32.0%</td>
<td>33.4%</td>
<td>-1.4%</td>
</tr>
<tr>
<td>Ever Told You Have Cancer (Other Than Skin Cancer)</td>
<td>5.7%</td>
<td>6.8%</td>
<td>-1.1%</td>
</tr>
<tr>
<td>New Stroke/Cerebrovascular Event</td>
<td>11.5%</td>
<td>10.5%</td>
<td>1.0%</td>
</tr>
<tr>
<td>New Stroke/Cerebrovascular Event</td>
<td>18.5%</td>
<td>18.0%</td>
<td>0.5%</td>
</tr>
<tr>
<td>New Stroke/Cerebrovascular Event</td>
<td>18.5%</td>
<td>19.0%</td>
<td>-0.5%</td>
</tr>
<tr>
<td>Mean Num of Cigarettes Now Smoke Each Day (Among Every Day Smokers)</td>
<td>15.8</td>
<td>15.6</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Note:
1. The 2013-2014 National Adult Tobacco Survey (NATS) is used as benchmark for the mean number of cigarettes now smoke each day.
2. The 2017 Behavioral Risk Factor Surveillance System (BRFSS) is used as benchmark for all other estimates.
3. Both AmeriSpeak profile-based estimate and benchmark are weighted.
AmeriSpeak Profile-based Estimates vs. Benchmarks: Chronic Health Conditions

Has a doctor ever told you that you had...

<table>
<thead>
<tr>
<th>Condition</th>
<th>AmeriSpeak</th>
<th>Benchmark</th>
<th>Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>A heart attack</td>
<td>2.0%</td>
<td>2.6%</td>
<td>-0.6%</td>
</tr>
<tr>
<td>Angina or coronary heart disease</td>
<td>2.4%</td>
<td>4.3%</td>
<td>-1.9%</td>
</tr>
<tr>
<td>Stroke</td>
<td>2.4%</td>
<td>2.0%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Any kind of diabetes or high blood sugar</td>
<td>11.5%</td>
<td>10.8%</td>
<td>0.7%</td>
</tr>
<tr>
<td>Cancer (other than skin cancer)</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Depression</td>
<td>10.6%</td>
<td>10.6%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Mode Effects: A Unique Test Using the AmeriSpeak Panel

Benefits of a mixed-mode design (selection effects)

- Researchers are increasingly using a mixed-mode survey design to obtain a representative sample.
- Allowing people to complete surveys on either the phone or the web can increase coverage and enhance representativeness.
  - A large portion of the public does not have access to the internet or prefers not to take surveys on the web (Bennett et al. 2017).
  - Many households are wireless only and are difficult to reach via phone (Blumberg & Luke 2017).
- Research shows significant attitudinal/behavior differences between those with and without internet access in United States (Dutwin and Buskirk 2017).

Challenges with a mixed-mode design (measurement effects)

- A mixed-mode approach can potentially lead to mode measurement effects.
- There could be differences between those who complete the survey on the web and those who complete it on the phone based on the interview mode.
  - Phone mode features interviewers while web mode is self-administered.
  - Phone mode presents questions verbally while web mode presents questions visually.
Research question

- When there are differences between those who complete the survey on the web and those who complete it on the phone, are those due to mode selection effects or mode measurement effects?
  - Researchers want to maximize representativeness and capture mode selection effects.
  - Researchers want to minimize mode measurement effects.
- The challenge is that respondents often select their mode so it is difficult to disentangle mode selection and measurement effects.
- A survey experiment is needed to explore whether differences are selection effects or measurement effects.

Experiment by Sterrett, et al. 2018

Panelist usual mode  |  Experimental study mode
---|---
Phone-Phone  |  Phone
Web-Phone  |  Web
Web-Web  |  Web

Sample details

- Drew a nationally representative sample.

<table>
<thead>
<tr>
<th>Group</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web-Web</td>
<td>1,001</td>
</tr>
<tr>
<td>Web-Phone</td>
<td>1,017</td>
</tr>
<tr>
<td>Phone-Phone</td>
<td>880</td>
</tr>
<tr>
<td>Total</td>
<td>3,098</td>
</tr>
</tbody>
</table>

Survey details

- A 15-minute survey with wide range of topics including:
  - Political attitudes
  - Views on social issues and the economy
  - Personal finances
  - Participation in social groups
  - News behavior
  - Personal health and medical care
Evidence of mode selection effects

Would you say that in general your health is...?
% Excellent or very good

Differences between Phone-Phone and Web-Phone group are significant in regression controlling for age, gender, race/ethnicity, income, education, and partisanship.

Thinking about your financial situation, do you have any of the following... a credit card?
% Yes

Differences between Phone-Phone and Web-Web group are significant in regression controlling for age, gender, race/ethnicity, income, education, and partisanship.

In general, do you think the benefits the United States gets from legal immigration outweigh the risks, or are the risks to the United States great enough so that legal immigration should be further limited?
% Outweigh

Differences between Phone-Phone and Web-Phone group are significant in regression controlling for age, gender, race/ethnicity, income, education, and partisanship.

Generally speaking, would you say that most people can be trusted or that you can’t be too careful in dealing with people?
% Most people can be trusted

Differences between Phone-Phone and Web-Web group are significant in regression controlling for age, gender, race/ethnicity, income, education, and partisanship.
Evidence of mode selection effects

Differences between Phone-Phone and Web-Web group are significant in regression controlling for age, gender, race/ethnicity, income, education, and partisanship.

Evidence of measurement mode effects

Questions such as this have potential mode effects related to social desirability bias and differences between interviewer/self-administered modes.

Preliminary takeaways

- There is strong evidence that offering a mixed-mode design improves sample coverage and the representativeness of the survey.
  - Those who choose to complete a survey in one mode (phone) are very different on some key attitudinal and behavioral issues than those who choose to complete the survey in another mode (web).
  - Many AmeriSpeak phone panelists are recruited by in-person non-response follow-up and these cases help improve sample representativeness (Bilgen 2017).
- Researchers should be aware of potential mode effects and design the questionnaire to mitigate such effects.
  - Limit questions prone to social desirability bias.

The Impact of Non-response Follow-up on AmeriSpeak Data Quality
Research Questions

What is the impact of nonresponse follow-up (NRFU) during panel recruitment on AmeriSpeak Data Quality and Study Estimates?

- What is the impact of F2F NRFU on panel response rates (AAPOR RR3)?
- Does F2F NRFU improve demographic representation of the panel sample?
- Does F2F NRFU have an impact on the study estimates?
  - In what ways NRFU panel recruits differ in their opinions, attitudes, and behaviors in comparison to the initial recruited panelists among different AmeriSpeak studies?

AAPOR RR3 (2014-2018 AmeriSpeak Panel Recruits)

<table>
<thead>
<tr>
<th>Description</th>
<th>Response Rate (AAPOR RR3, Weighted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household Response Rate due to Initial Recruitment</td>
<td>5.6%</td>
</tr>
<tr>
<td>Household Response Rate due to NRFU</td>
<td>28.6%</td>
</tr>
<tr>
<td>Household Response Rate</td>
<td>34.2%</td>
</tr>
</tbody>
</table>

NRFU boosts response rate by 6.1 times

Panel Composition (2014-2018 AmeriSpeak Panel Recruits)

- Unweighted Panel Distribution by Panelist Type and Age
- Unweighted Panel Distribution by Panelist Type and Race
Panel Composition (2014-2018 AmeriSpeak Panel Recruits)

Unweighted Panel Distribution by Panelist Type and Education

- NRFU recruits tend to be more educated compared to F2F recruits.
- Initial recruits overrepresent higher education.
- Mail/Phone recruits tend to be more interested in current events.
- Mail/Phone recruits are more pro-science, compared to F2F recruits.
- F2F recruits more strongly opposed to gun control policy.

How interested are you in current news events?

- Mail/Phone recruits tend to be more interested in current events.

Science and technology are making our lives healthier, easier, and more comfortable

- Mail/Phone recruits tend to be more pro-science, compared to F2F recruits.

Favor or oppose the proposal: “Prohibiting a person convicted of drunk and disorderly conduct to carry a loaded gun in public”

- Mail/Phone recruits more strongly support the policy.
- F2F recruits provide a middle-of-the-road view.
To what extent do you think people like you are treated fairly in the civil legal system? [Asked of Low-Income Households]

- Not at all & Rather
- Some of the time
- Most of the time & All of the time
- Don't know

Mail/Phone recruits more likely to perceive institutional unfairness

**p < 0.0001**

Source: Legal Services Corporation “Justice Gap” Study (Weighted)

Do you consider yourself to be . . . [Asked of Working adults age 18-64 in private sector]

- Democrat
- Independent
- Republican
- Something else

Mail/Phone recruits more likely identify as Democrats

**p < 0.0001**

Source: AARP Retirement Survey (Weighted)

When it comes to politics, do you usually think of yourself as...

- Externally Liberal
- Liberal
- Slightly Liberal
- Moderate
- Slightly Conservative
- Conservative
- Extremely Liberal

Mail/Phone recruits more likely report a political ideology whether Liberal or Conservative

**p < 0.0001**

Source: AmeriSpeak Panel Recruitment Survey (Weighted)

National Spending Priorities: % Reporting the Country Spends “Too Little”

- Improving and protecting the nation’s health
- Social security
- Highways and bridges
- Improving and protecting the environment
- Solving the problems of the big cities
- Improving the conditions of blacks
- Mass transportation
- Supporting scientific research

Mail/Phone % - F2F % > 4pp

F2F recruits trend towards conservative positions on government spending

**p < 0.0001**

Source: AmeriSpeak Panel 2016 General Social Survey (Weighted)
SUMMARY – IS FACE TO FACE RECRUITMENT MAKING A DIFFERENCE?

Mail/Phone panelists – recruited with less effort
more prone to being very interested in news, pro-science, and liberal policy positions

F2F panelists – recruited after greater effort
less prone to being very interested in the news and more likely to report conservative policy positions

QUESTIONS!

TESS Experiments

- Funded by National Science Foundation
- Must be experimental and original
- Free access to AmeriSpeak Panel
- www.tessexperiments.org

THANK YOU.

Michael Dennis
Dennis-Michael@norc.org
650-537-7950

David Dutwin
Dutwin-David@norc.org
312-759-4027
amerispeak.norc.org