AmeriSpeak

CELEBRATING 10 YEARS

Avoiding Fraud in Market Research

Leveraging Probability Panels and
Non-Response Follow-Up Strategies

April 2, 2025

David Dutwin

David Dutwin
Executive Director, AmeriSpeak

- Joined AmeriSpeak in 2019
- Research background on election methodology, surveying low-incidence populations, the use of big data in survey research
- Senior Fellow at the Program for Opinion Research and Election Studies at the University of Pennsylvania
- President of the American Association of Public Opinion Research (AAPOR) from 2018-2019
- Taught courses in survey research and design, political polling, research methods, rhetorical theory, media effects, and more at the University of Pennsylvania, the University of Arizona, and West Chester University

The Problem

Fraud in opt-in survey data

AmeriSpeak

Risks of Opt-in Survey Sampling

Sometimes called opt-in, convenience sampling, or non-probability surveys, whatever you call it, often leads to a data quality issues that can result in inaccurate results.

Pew Research Center

Online opt-in polls can produce misleading results, especially for young people and Hispanic adults

Greenbook

The Rising Issue of Bad Data in Online Surveys: Causes and Contributing Factors
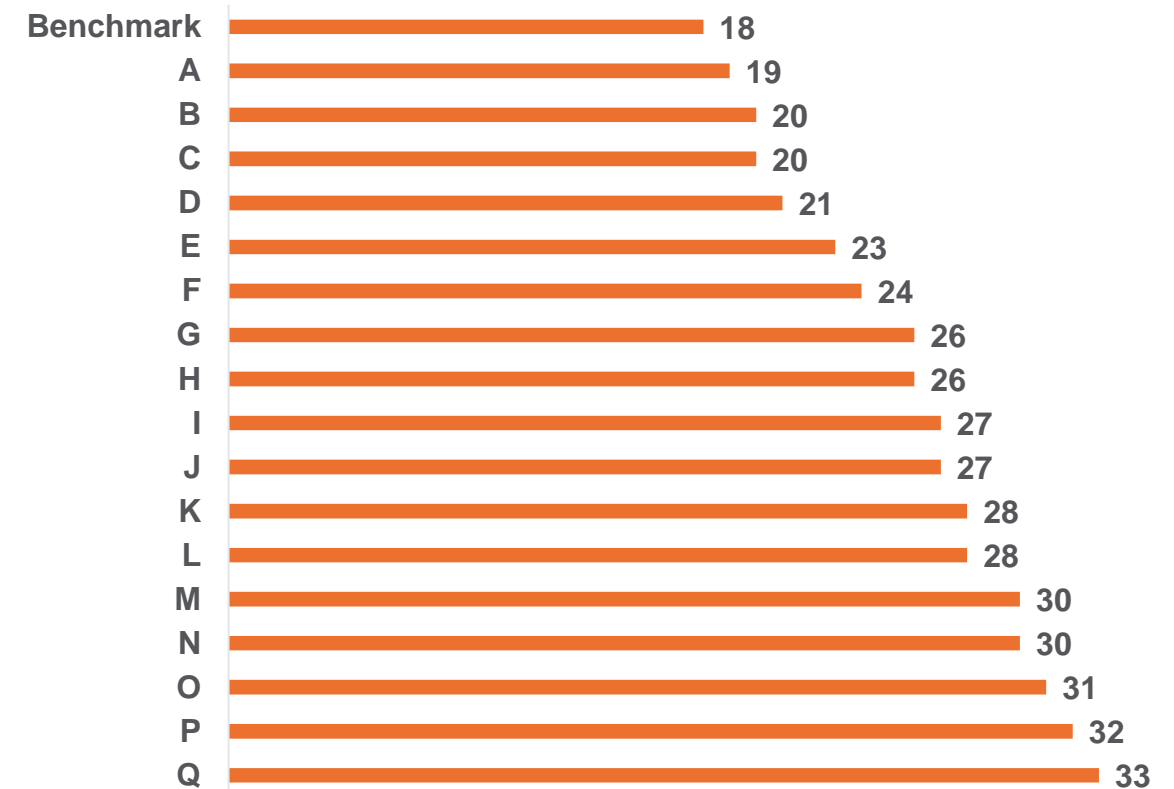
DATA QUALITY, PRIVACY, AND ETHICS   February 25, 2025

Truth To Be Told: Five Realities About Online Sample That Compromise Data Quality

Greenbook

- **Limited Participant Validation** Within Online Panels Presents a Significant Challenge to Ensuring Unique and Authentic Respondents

- Most Panels Rely on a **Self-Selection** Model With No Limits on Respondent Participation

- The Majority of Suppliers **Blend Sample From Multiple Sources**, Even Those Touting Their Proprietary Panels

- **Fraud** Easily Blends Into Your Data

- Marketing Materials Often Misrepresent the Supplier's Actual Capabilities, Including Quality, Panel Size, and Profiling Information
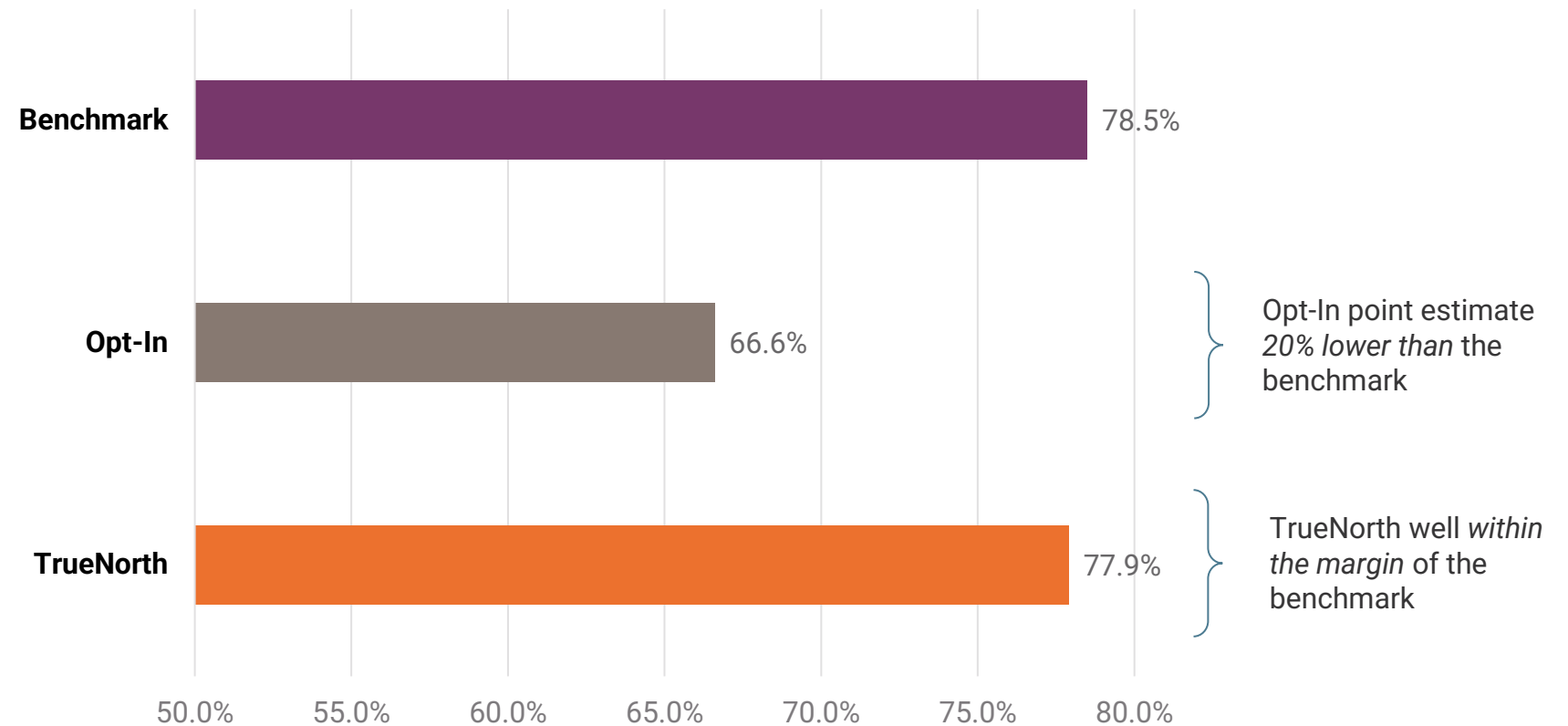
- The Advertising Research Foundation (ARF) Foundations for Quality study.

- 17 U.S. Nonprobability panel providers.

- Benchmark questions asked to assess Nonprobability accuracy.

- Results: Wide variation across panels in the survey estimates.

Comparison of Smoking Prevalence; Benchmark vs. 17 Opt-In Panels

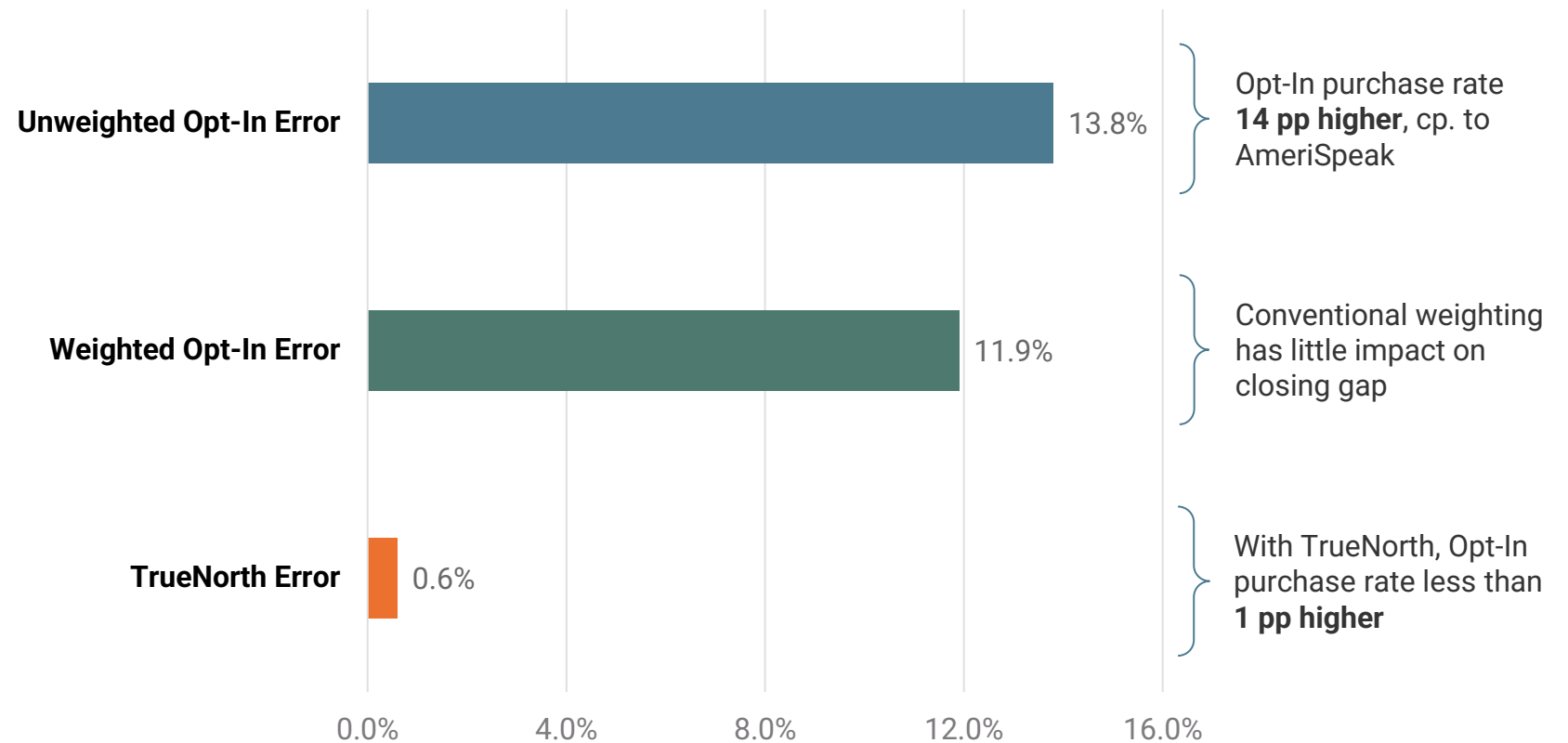| Panel	Value
Benchmark	18
A	19
B	20
C	20
D	21
E	23
F	24
G	26
H	26
I	27
J	27
K	28
L	28
M	30
N	30
O	31
P	32
Q	33

Buying a single company stock usually provides a safer return than a stock mutual fund? (Percent False)

Point Estimates

Benchmark: SCF 2016

Benchmark 78.5%

Opt-In 66.6%

Opt-In point estimate *20% lower than* the benchmark

TrueNorth 77.9%

TrueNorth well *within the margin* of the benchmark

Buying a single company stock usually provides a safer return than a stock mutual fund? (Percent False)

Average Absolute Error

Unweighted Opt-In Error	13.8%
Weighted Opt-In Error	11.9%
TrueNorth Error	0.6%

Opt-In purchase rate **14 pp higher**, cp. to AmeriSpeak

Conventional weighting has little impact on closing gap

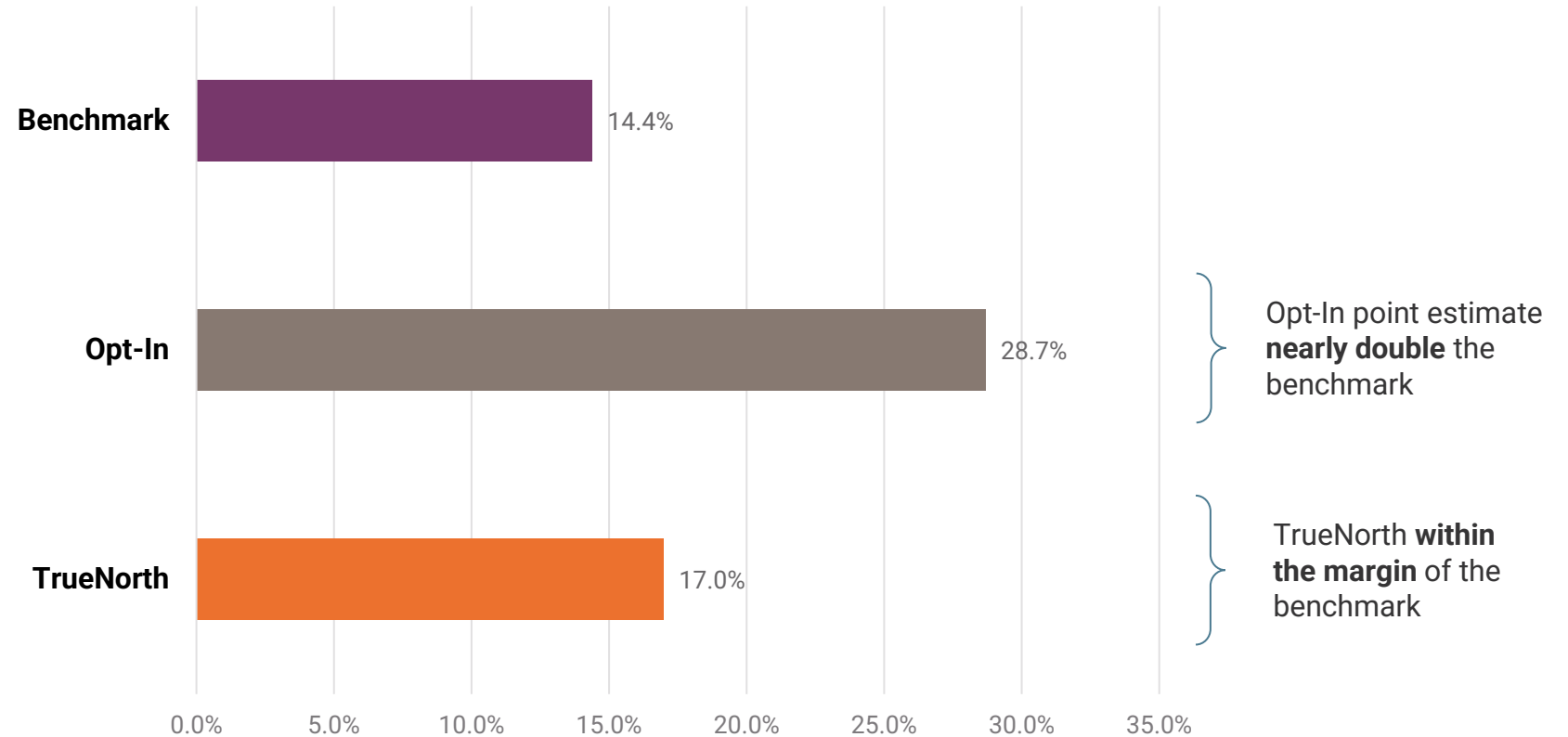With TrueNorth, Opt-In purchase rate less than **1 pp higher**

0.0%   4.0%   8.0%   12.0%   16.0%

Benchmark: SCF 2016

How likely are you to purchase a new car or light truck in the next 12 months?

Point Estimates on Car Purchase

Benchmark — 14.4%

Opt-In — 28.7%

TrueNorth — 17.0%

Opt-In point estimate **nearly double** the benchmark

TrueNorth **within the margin** of the benchmark

0.0% 5.0% 10.0% 15.0% 20.0% 25.0% 30.0% 35.0%

How likely are you to purchase a new car or light truck in the next 12 months?

Average Absolute Error Across All Items on
Car Purchase

Unweighted Opt-In Error — 14.3%

Weighted Opt-In Error — 13.3%

TrueNorth Error — 2.8%

Opt-In purchase rate **14 pp higher**, cp. to AmeriSpeak

Conventional weighting has little impact on closing gap

With TrueNorth, Opt-In purchase rate only **2-3 pp higher**
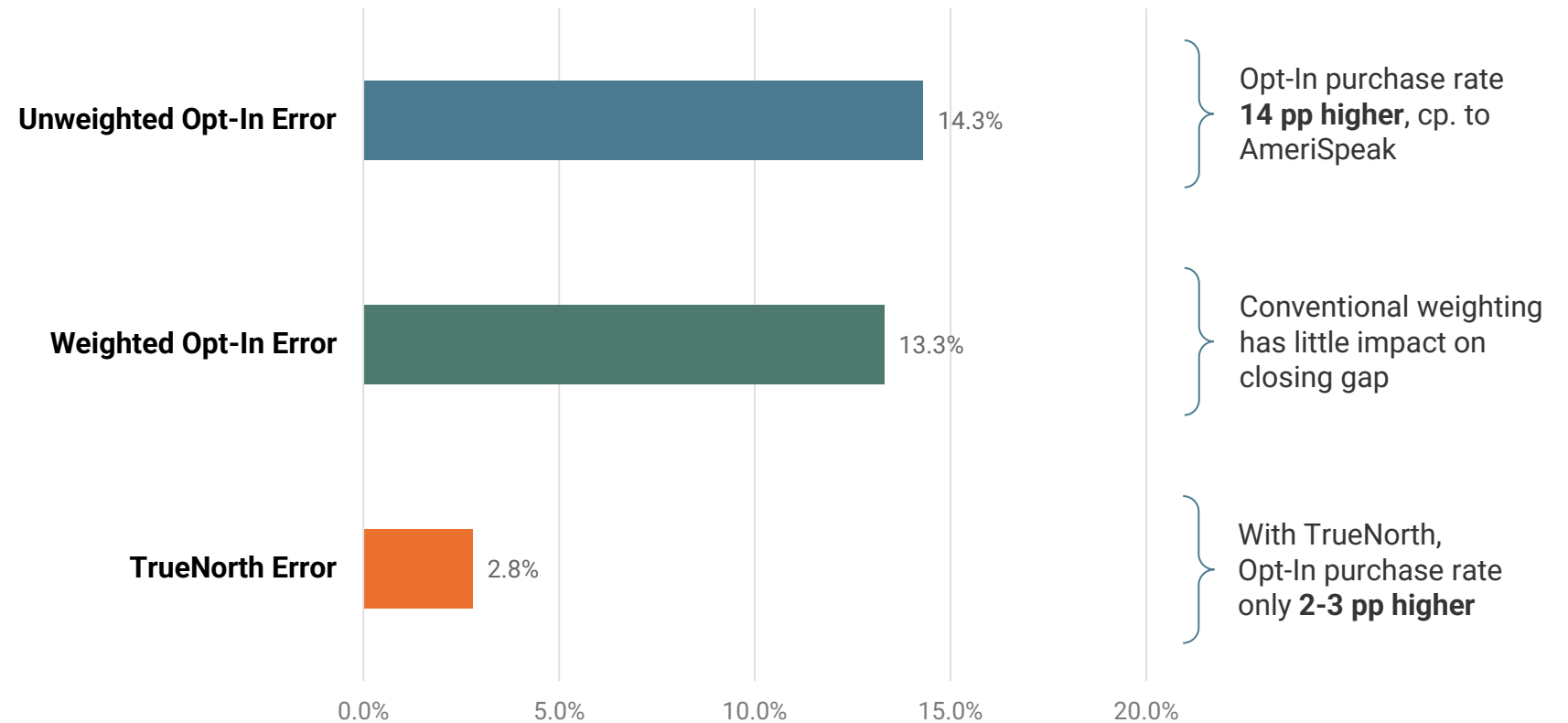
Rethinking Probability Panels

Part of the solution?

AmeriSpeak

Probability panels: A refresher

Key Features

- **Random Selection:** Participants are chosen randomly, minimizing bias and enhancing the accuracy of the data.

- **Representative Sample:** The panel reflects the diversity of the population, providing reliable insights into public opinion and behavior.

- **High Data Quality:** The rigorous selection process ensures high-quality, reliable data.

Pew Research Center

Comparing Two Types of Online Survey Samples

Opt-in samples are about half as accurate as probability-based panels

Larger errors on online opt-in samples, especially for 18- to 29-year-olds and Hispanic adults

Average absolute error on benchmark variables

	Probability-based panels	Opt-in samples
All adults	2.6	5.8
Ages 18-29	3.6	11.2
Hispanic	3.6	10.8

Public Opinion Quarterly

Apples to Oranges or Gala versus Golden Delicious? Comparing Data Quality of Nonprobability Internet Samples to Low Response Rate Probability Samples

FREE

David Dutwin, Trent D. Buskirk   Author Notes

How Probability Differs from Opt-in Surveys

Opt-in Sampling

Pros:

✓ Fast and Cheap

✓ Niche Targeting

Cons:

✗ Greater Variance of Estimates, thus

✗ Inconsistent Results

✗ Unknown Bias

Probability Sampling

Pros:

✓ Accuracy and Representativeness

✓ Minimizes Bias
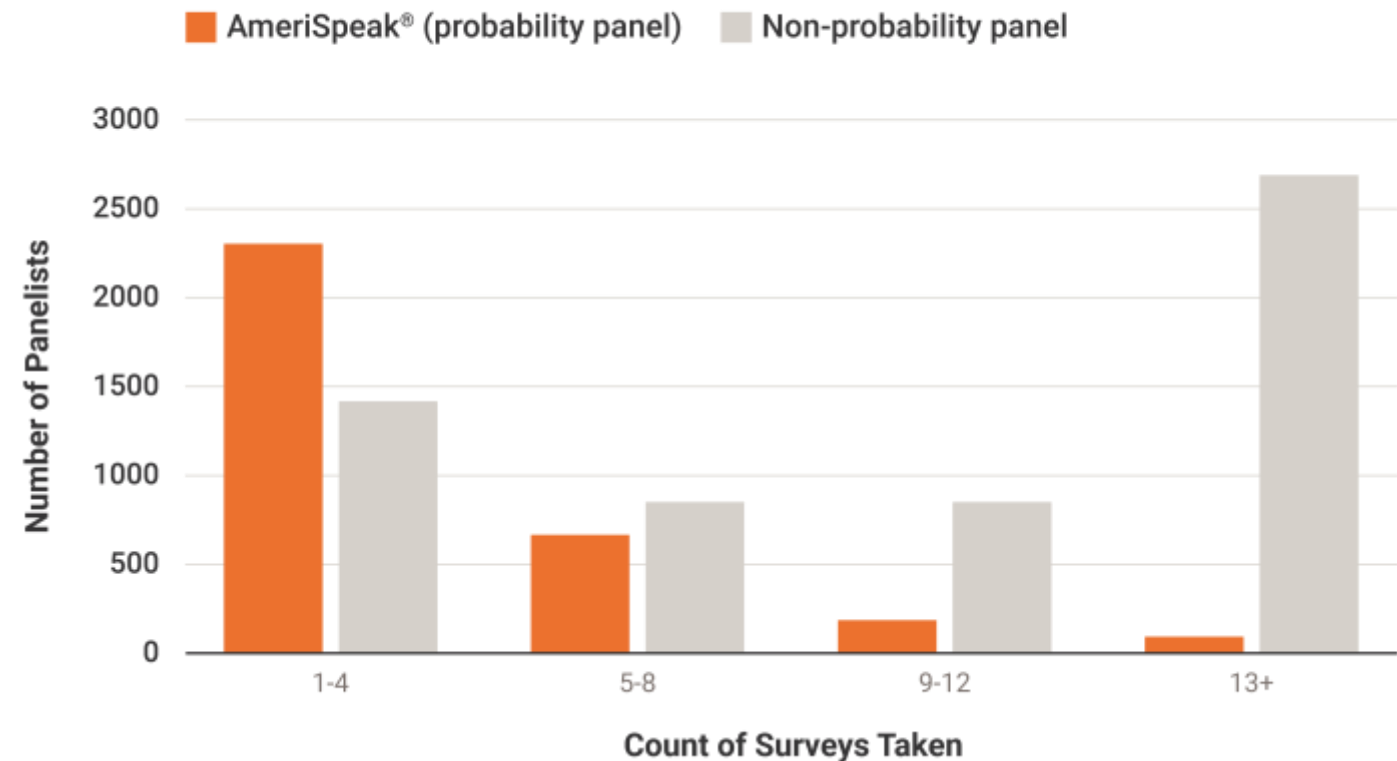
✓ Generalizability

✓ Statistical Rigor

Cons:

✗ **Cost and Time**

✗ Declining Response Rates

Survey Taking Frequency Differences Between Probability and Nonprobability Panelists

Case4Quality Study: 3% heaviest users do 20% of all surveys; **40% of all are completing 100 surveys daily**.

Surveys taken in the past 30 days by panel type

■ AmeriSpeak® (probability panel)   ■ Non-probability panel

Slide 10: Question: "How many surveys have you completed in the past 30 days?"
Source: The survey was administered in 2022 to 3,331 AmeriSpeak® probability panelists and 5,855 nonprobability panelists.

Multiple Panel Membership Differences
Between Probability and Nonprobability Panelists

Average number of panel memberships by panel type
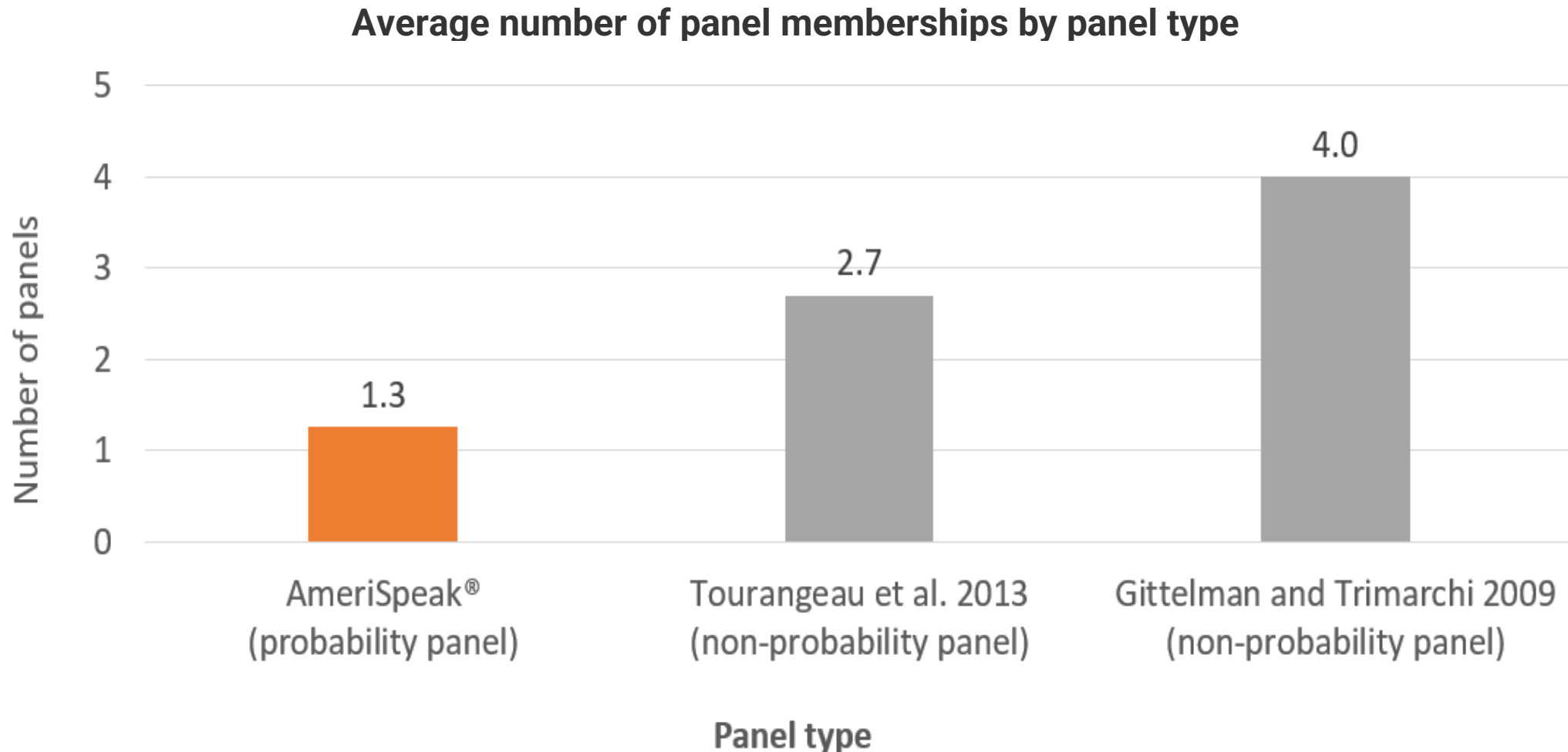
Non-response follow-up recruitment builds a more accurate and representative composition, capturing a wider range of participants who might otherwise be missed

NORC National Frame

Invited Sample

Initial Recruitment
Postcards, Recruitment Package, Telephone

Recruitment Profile Survey
Web, Phone

Non-Responders

Random Selection

Not Selected

Responders

Active for Client Surveys

Non-Response Follow-Up
- Face-to-Face Interviewing
- FedEx or Priority Mailing
- Enhanced Incentive
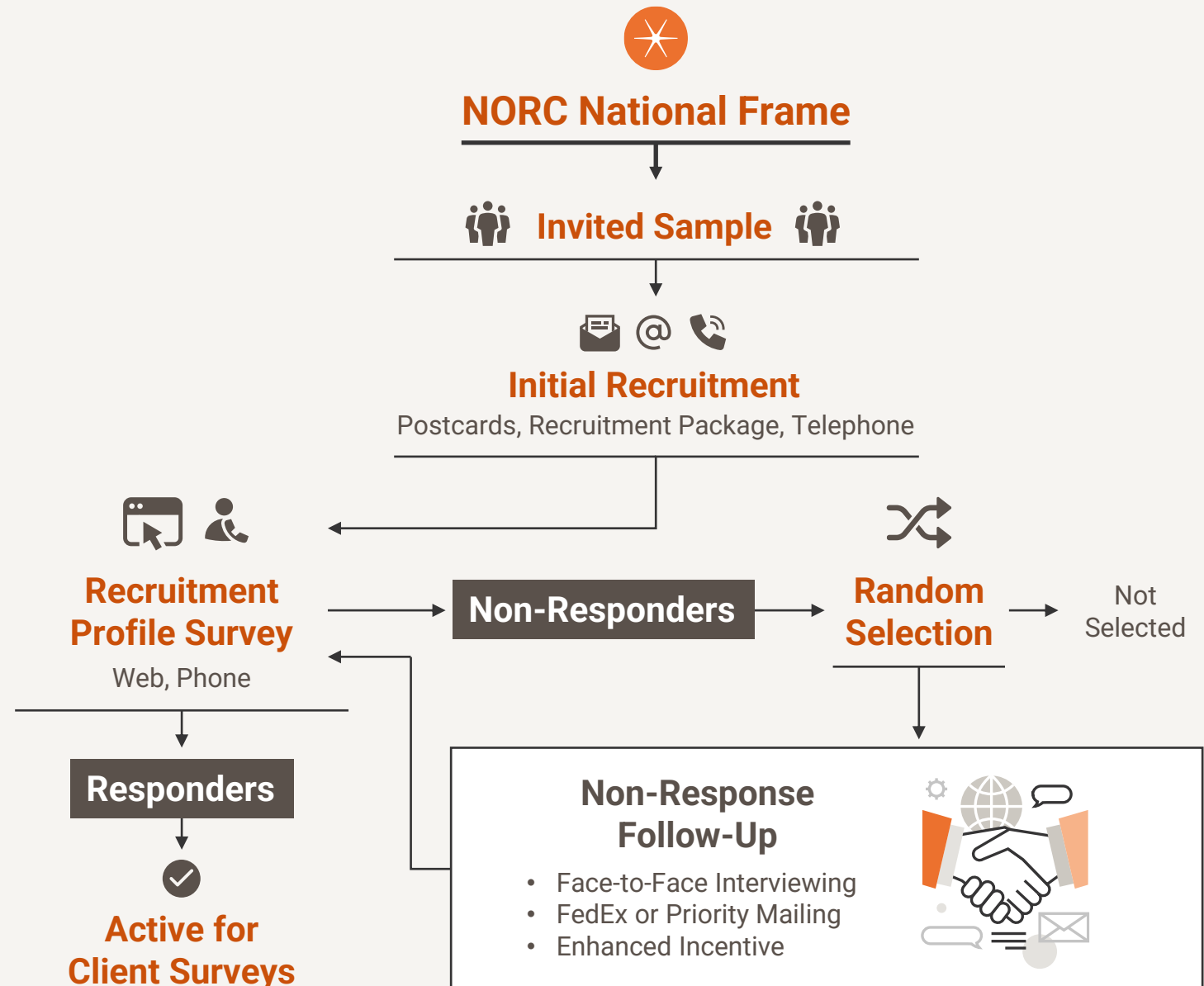
If Opt-In is not to be trusted, and Probability is too expensive and too slow, what is/are the solution(s)?

AmeriSpeak

Probably Leveraging:
Use-Cases where probably makes sense

AmeriSpeak

Fastest and least expensive general population probability research available

- Access AmeriSpeak panel quality

- Runs two-four times monthly

- Four-day survey fields, 1,000 interviews

- Roughly $1 per question per interview

- Ask three questions, get 30 back

TrueNorth

Leverage some probability data to model and reduce bias in the opt-in samples

- Conduct 25% in AmeriSpeak, 75% in opt-in

- Results in a reduction of 50-70% of bias

- Lowers variance of bias substantially…the more biased the answer the more TrueNorth reduces it.

- Yes, more expensive than 100% opt-in and a longer field…but a reasonable compromise

Benchmarking and Calibration

Use a single probability survey to benchmark all of your opt-in surveys

- Field a benchmark survey (omnibus or custom)

- Weighting protocols developed to apply to every opt-in survey fielded afterward

- Allow "business as usual" but with a substantial increase in quality

Qualitative

Recruiting from a Representative Sample

- Panelists are "light use" qualitative respondents

- Profiling at panel recruitment minimizes screener falsification

Benchmarking and Calibration for AI (in development)

Expand the training data for synthetic data by incorporating data generated from high quality NORC surveys, including 10 years of AmeriSpeak data, the General Social Survey and other datasets

Questions?

Thank you.

David Dutwin
Executive Director, AmeriSpeak

Get Your Research Right

AmeriSpeak