Pew Research Center

ACF/OPRE's 2023 Methods Meeting: Addressing Unit Missingness in Social Policy Survey Research

Non-probability sampling

Courtney Kennedy

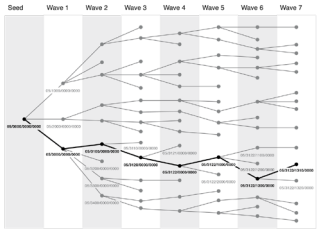
VP Methods and Innovation

Some types of non-probability samples

Volunteers in a clinical trial

Commercial list of adults in ethnic group

Respondent-driven sample



Online opt-in panel or marketplace

focus of this talk

When are online non-probability samples “fit for purpose”

Likely fit for purpose

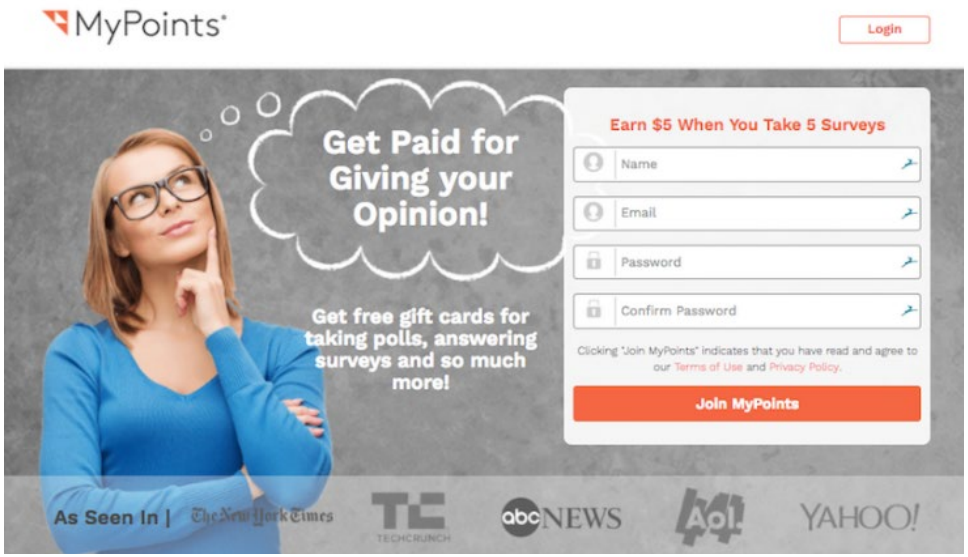
- Exploratory phase (e.g., before main study)
- A “rough estimate” is sufficient
- Some types of marketing research

Likely NOT fit for purpose

- Estimating a rare outcome
- Accuracy is important
- Estimates for racial and ethnic subgroups are important
- Estimates for young adults are important
- Testing for an effect that interacts with race, ethnicity, or age

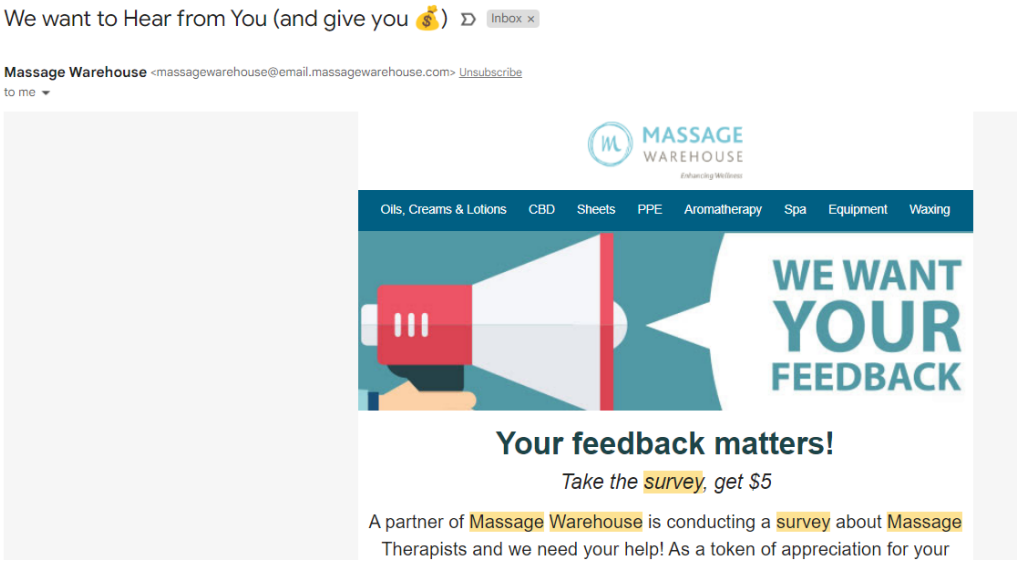
How online non-probability samples are created

websites where people sign up to make money



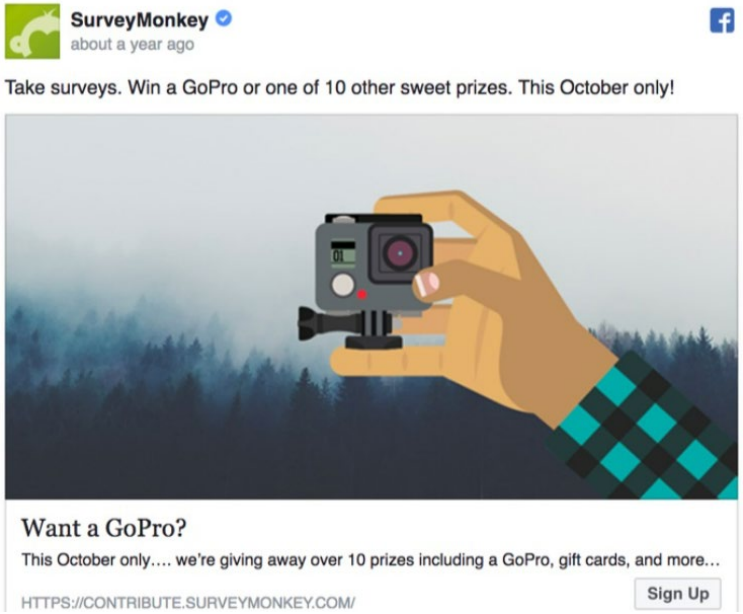
How online non-probability samples are created

emails to customer lists



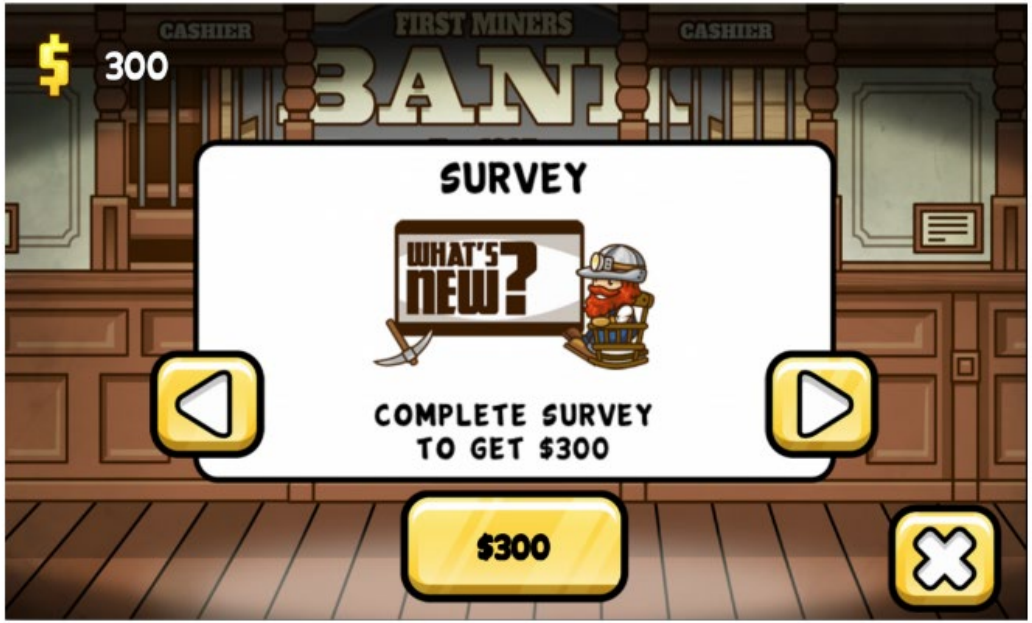
How online non-probability samples are created

ads on social media



How online non-probability samples are created

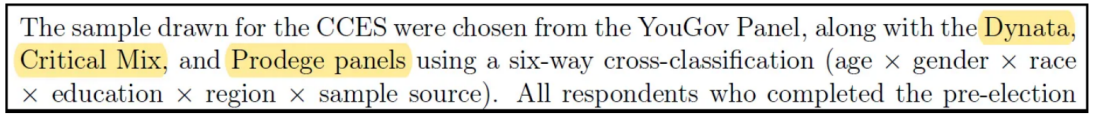
video games (in exchange for extra life or premium content)



How online non-probability samples are created

and, increasingly, by third party sourcing

Two of the most high profile surveys using nonprobability sample (CES, VoteCast) source from third-party panels.



The sample drawn for the CCES were chosen from the YouGov Panel, along with the Dynata, Critical Mix, and Prodege panels using a six-way cross-classification (age × gender × race × education × region × sample source). All respondents who completed the pre-election

Figure 1: Data sources for the Cooperative Election Survey (Ansolabehere, Schaffner, and Luks 2019, p.13)

Nonprobability sample

Nonprobability participants will include panelists from Cint, Prodege, or Dynata, including members of its third-party panels. Digital fingerprint software and panel-level ID validation is used to prevent respondents from completing the AP VoteCast survey multiple times.

AP VoteCast: 2022 Midterm General Election Methods Statement, NORC

Probability sampling

Population

Sampling frame

Selected sample

Responding sample

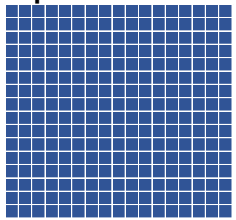
The *researcher* selects the survey sample

Non-probability sampling

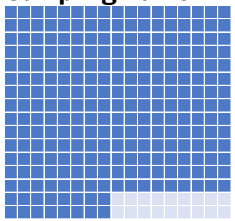
The *participants* select themselves into the survey

Probability sampling

Population



Sampling frame



Selected sample

Responding sample

The *researcher* selects the survey sample

Non-probability sampling

The *participants* select themselves into the survey

… which yields less representative samples and opens the door to fraud

How can we measure survey “accuracy”?

Benchmark Question:

Are you CURRENTLY covered by any of the following types of health insurance or health coverage plans? (American Community Survey)

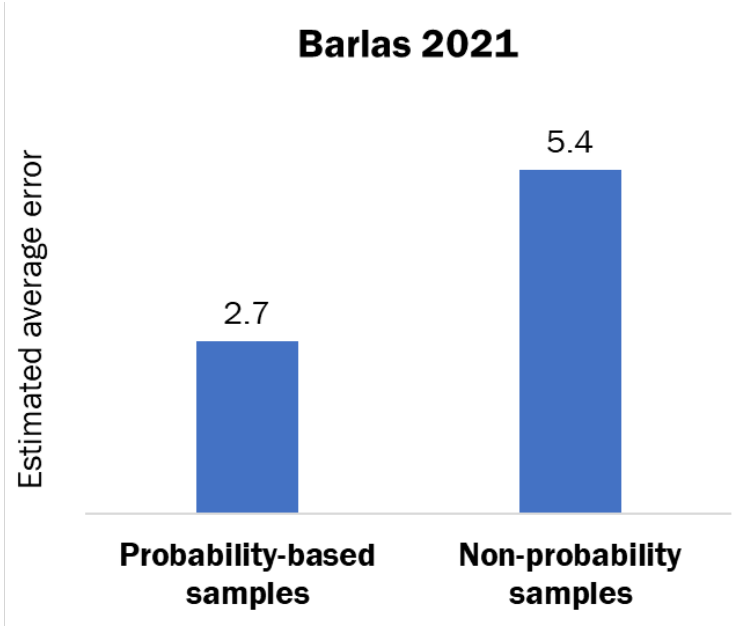
Benchmark estimate	Survey estimate	Difference	Absolute difference
89%	85%	- 4 pp	4 pp

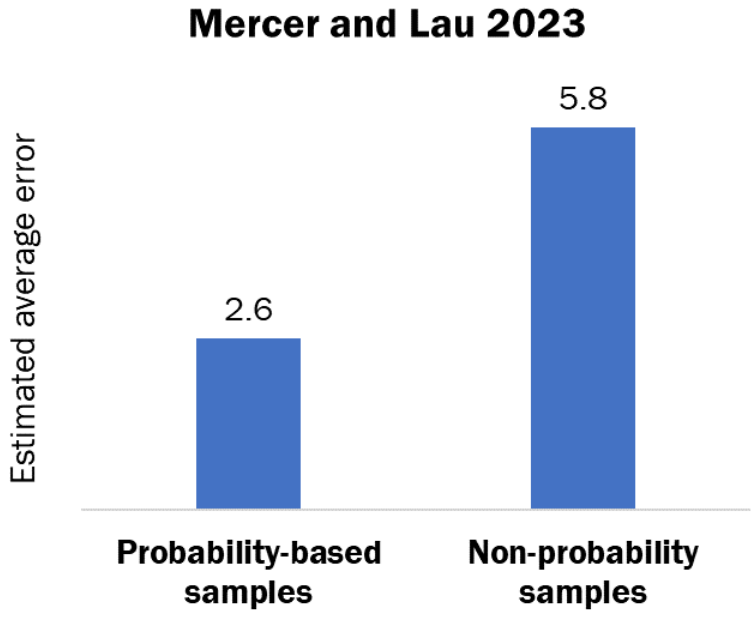
Studies often compute this for all the benchmark estimates and then examine the average of the absolute differences

Non-probability survey estimates tend to be less accurate

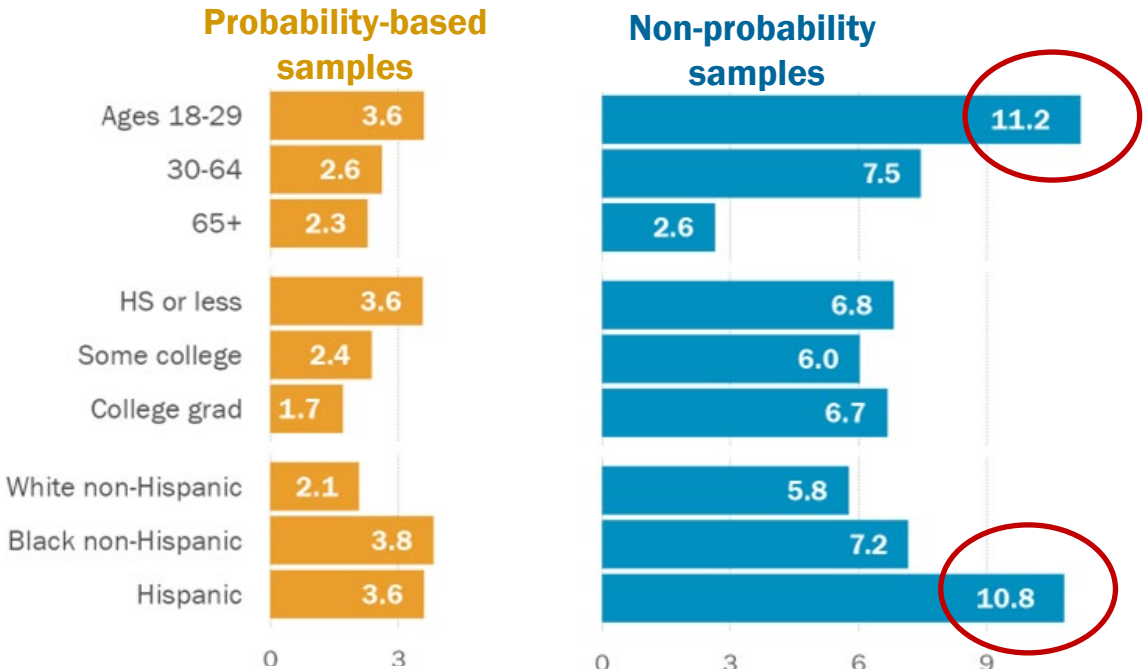
Found non-probability surveys were less accurate	Found non-probability surveys were just as accurate
Malhotra and Krosnick (2007)	Vavrek and Rivers (2008)
Chang and Krosnick (2008)	Ansolabehere and Schaffner (2014)
Yeager et al. (2011)	
Szolnoki and Hoffmann (2013)	
Erens et al. (2014)	
Sturgis et al. (2016)	
Dutwin and Buskirk (2017)	
MacInnis et al. (2018)	
Pennay et al. (2018)	
Silver (2018)	
Barlas (2021)	
Mercer and Lau (2023)	

Non-probability survey estimates tend to be half as accurate





Non-probability estimates for young adults and Latinos have especially large errors



Average absolute error on 25 benchmark variables

Source: Mercer and Lau 2023

“Some of the biggest threats to (non-probability) data quality are bots and cheaters. Often bots will complete surveys en masse, or a person will take surveys on behalf of someone else multiple times.”

Qualtrics website

https://www.qualtrics.com/support/survey-platform/survey-module/survey-checker/fraud-detection/

“Just 10 years ago, researchers would need to remove 5%-10% of all interviews from online (non-probability) samples because of poor quality. That proportion is now in the 35%-50% range.”

Geraci, John. 2022. *Pollarized*, p 153
(emphasis added)

Bogus cases tend to say “Yes” or “Agree” no matter what is asked

Are you licensed to operate a nuclear submarine?

- ○ Yes
- ○ No

Are you of Hispanic, Latino, or Spanish origin, such as Mexican, Puerto Rican or Cuban?

- ○ Yes
- ○ No

Source: Kennedy, Mercer, Lau forthcoming

Bogus cases lead to over-estimates of rare outcomes

Belief in conspiracy theories
(Lopez and Hillygus 2018)

Support for political violence
(Westwood et al. 2022)

Ingesting bleach to protect from COVID
(Litman et al. 2020)

Americans' favorability of Putin
(Kennedy et al. 2021)

Key takeaway #1

- Online non-probability vendors have measures in place to detect and remove bogus respondents. These include:
  - Screening out foreign IP addresses
  - Digital fingerprinting to prevent duplicate interviews
  - IP proxy testing
  - Validation of email addresses
  - Machine learning to flag suspicious patterns

- How well do these work?
  - Not very well. Geraci noted that the share of non-probability cases that need to be dropped because of poor quality is now “in the 35%-50% range” despite vendor checks.

Key takeaway #2

Statements like…

“Yes, this study was non-probability, but it’s OK because I used a ‘good’ panel”

… should not be taken at face value. Read Enns and Rothschild (2022) for a primer. In short, even nonprobability panels that were once considered ‘good’ now routinely source respondents from third-party suppliers, which are often a black box.

References

Ansolabehere, S., and Schaffner, B. (2014). Does survey mode still matter? findings from a 2010 multi-mode comparison. *Political Analysis*, 22(3), 285-303.

Barlas, F. (2021). "Representative Research: Assessing Diversity in Online Samples?" Presented at University of Michigan Webinar, November 10, 2021.

Chang, L., Krosnick, J.A. (2009). National surveys via RDD telephone interviewing versus the internet: comparing sample representativeness and response quality. *Public Opinion Quarterly*, Volume 73, 4, 641–678.

Dutwin, D., and Buskirk, T.D. (2017). Apples to oranges or gala versus golden delicious? comparing data quality of nonprobability internet samples to low response rate probability samples. *Public Opinion Quarterly*, 81, 213–239.

Enns, P.K. and Rothschild, J. 2022. "Do you know where your survey data come from? Outsourcing data collection poses huge risks for public opinion." Medium. Available at https://medium.com/3streams/surveys-3ec95995dde2.

Erens, B., Burkill, S., Couper, M.P., Conrad, F., Clifton, S., Tanton, C., Phelps, A., Datta, J., Mercer, C.H., Sonnenberg, P., Prah, P., Mitchell, K.R., Wellings, K., Johnson, A.M., and Copas, A.J. (2014). Nonprobability web surveys to measure sexual behaviors and attitudes in the general population: a comparison with a probability sample interview survey. *Journal of Medical Internet Research,* 16(12), e:276.

Geraci, J. (2022). *Poll-arized: Why Americans Don't Trust the Polls And How to Fix Them Before It's Too Late*. Houndstooth Press, p. 153.

Kennedy, C., Hatley, N., Lau, A., Mercer, A., Keeter, S., Ferno, J., and Asare-Marfo, D. (2021). Strategies for detecting insincere respondents in online polling. *Public Opinion Quarterly*, 85, 1050–1075.

Kennedy, C., Mercer, A., and Lau. "Exploring the Assumption That Online Nonprobability Survey Respondents Are Answering in Good Faith." *Survey Methodology*. Forthcoming.

Litman, L., Rosen, Z., Rosenzweig, C., Weinberger-Litman, S.L., Moss, A.J., and Robinson, J. (2021). Did people really drink bleach to prevent COVID-19? A tale of problematic respondents and a guide for measuring rare events in survey data. MedRxiv, https://doi.org/10.1101/2020.12.11.20246694

Lopez, J. and Hillygus, D.S. (2018). Why so serious? survey trolls and misinformation." Presented at the Annual Meeting of the Midwest Political Science Association, Chicago.

MacInnis, B., Krosnick, J.A., Ho, A.S., and Cho, M. (2018). The accuracy of measurements with probability and nonprobability survey samples: replication and extension. *Public Opinion Quarterly*, 82, 707–744.

Malhotra, N., & Krosnick, J. (2007). The effect of survey mode and sampling on inferences about political attitudes and behavior: comparing the 2000 and 2004 ANES to internet surveys with nonprobability samples. *Political Analysis*, 15(3), 286-323.

Mercer, A. and Lau, A. (2023). "Comparing two types of online survey samples." Pew Research Center. Available at https://www.pewresearch.org/our-methods/.

Pennay, D.W., Neiger, D., Lavrakas, P.J., and Borg, K. (2018). The online panels benchmarking study: a total survey error comparison of findings from probability-based surveys and nonprobability online panel surveys in Australia (2 ed.) The Australian National University. http://csrm.cass.anu.edu.au/ research/publications/methods-research-papers.

Sturgis, P., Baker, N., Callegaro, M., Fisher, S., Green, J., Jennings, W., Kuha, J., Lauderdale, B., and Smith, P. (2016). Report of the inquiry into the 2015 British general election opinion polls. London: Market Research Society and British Polling Council.

Szolnoki, G and Hoffmann, D. (2013). Online, face-to-face and telephone surveys—Comparing different sampling methods in wine consumer research. *Wine Economics and Policy*, 2, 57-66.

Vavreck, L. and Rivers, D. (2008). The 2006 Cooperative Congressional Election Study. *Journal of Elections, Public Opinion & Parties*, 18(4), 355-366.

Westwood, S.J., Grimmer, J., Tyler, M. and Nall, C. (2022). Current research overstates American support for political violence. *Proceedings of the National Academy of Sciences*, 119(12).

Yeager, D.S., Krosnick, J.A., Chang, L., Javitz, H.S., Levendusky, M.S., Simpser, A., and Wang, R. (2011). Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly*, 75, 709–747.

Thank you