Public Opinion Polling in the USA: Online Methodologies

A Technical Report submitted to the Department of Computer Science

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Ethan Chen

Spring, 2022

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Daniel G. Graham, Department of Computer Science

Rosanne Vrugtman, Department of Computer Science

Public Opinion Polling in the USA: Online Methodologies

CS4991 Capstone Report, 2022

Ethan Chen
Computer Science
The University of Virginia
School of Engineering and Applied Science
Charlottesville, Virginia USA
ec5rdn@virginia.edu

Abstract
Polling, the tabulation of opinions from a group of persons, is the primary method of collecting public opinion data in the United States. As traditional public opinion polling via phone calls has become more and more expensive, the vast majority of public opinion polling is now conducted via a variety of cheaper online methodologies. Recent studies find that, overall, online polling methodologies fare no worse than traditional phone-based methodologies, but still suffer from unique issues regarding their accuracy. Future research into online-only polls, hybrids of online and phone-based polls, and text messaging-based polls will likely yield insights into the merits of these innovations in polling methodology.

1. Introduction

Polling, the tabulation of opinions from a group of persons, is the primary method of collecting public opinion data in the United States. However, while the art of polling has existed since the 1930s, recent developments in American life, culture, and technology have greatly complicated the traditional process, requiring the field to evolve in response [1]. As polling via phone calls has become more expensive, the vast majority of public opinion polling is now conducted via a variety of cheaper online methodologies, bringing the field into the modern age [2].

To give some perspective on just how much online polling methodologies have displaced traditional phone-based surveys, according to a 2020 report from Pew Research Center, more than 80% of U.S. public opinion polls are conducted using online opt-in polling [2]. However, the popularity of these techniques has attracted criticism from political observers.

2. Related Works

Traditionally, public opinion polling in the U.S. consisted mainly of phone calls to a random sample of Americans [3]. However, decreasing response rates and the advent of cell phones has significantly increased the cost and difficulty of conducting polls in this fashion [4].

As shown in Figure 1 below, Pew Research Center found that response rates for telephone-based polls have fallen and stabilized at around 9%, down from a high of 36% in 1997. With declining response rates, more calls must be made to reach the same number of respondents, greatly increasing the expense [4].

As for the widespread transition from landline to cell phones among Americans, Pew Research Center has estimated that each cell phone interview costs about three times as much as a comparable landline-based interview [1]. At the same time, the widespread adoption of the Internet and the

polarization of American politics along demographic lines has made possible the development and widespread use of online methodologies for polling [3].

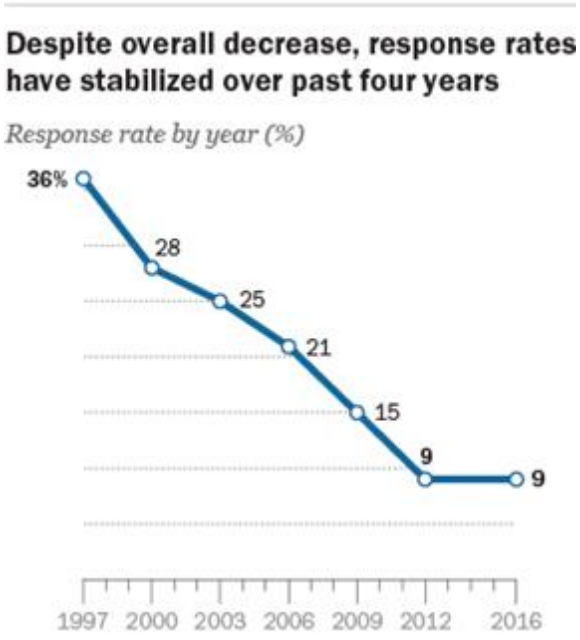

Figure 1. This graph presents landline telephone response rates over time [5].

Online polling methodologies differ greatly from their telephone-based counterparts. Traditional telephone-based polls generally work off probability samples, where every member of a target population has an equal chance of being asked to participate in a survey. If the population sample is large enough, its statistics will likely reflect the population at large. In contrast, online polling methodologies use nonprobability samples, where respondents do not have an equal chance of being asked, and thus the raw sample of responses will likely have very different statistics than the population at large [6].

To account for this, pollsters use a process called weighting to adjust the sample data to more accurately reflect the characteristics of the population at large. For instance, a pollster may ask respondents for their age, race, and educational attainment. Since they already know the actual demographics of the target population due to other sources like the U.S. Census Bureau, they would adjust the sample to match those demographics. In the United States, since demographics strongly correlate with political preferences, weighting allows for polls to reveal more accurate political statistics [6].

In general, pollsters prefer probability samples to nonprobability samples for a variety of reasons, and thus telephone-based polls to online methodologies. As such, the cost-motivated switch in methodology has attracted critics. In 2019, Nate Cohn of *The New York Times* claimed that online polls were unproven and a "plain step back from the historical level of accuracy of the so-called gold standard polls that have adhered to traditional best practices", showing that online election polls had an average error of 5.3%, while live-interviewer telephone polls had an average error of 3.5% in nonpresidential general election polls since 1998 [3]. Even online pollsters like Pew Research Center have found issues unique to online surveys, such as bogus respondents that finish surveys quickly in order to receive completion rewards, and thus exhibit biases such as a bias toward positive answers [2].

Still, for all the benefits phone-based polls supposedly receive from using a probability sample, differential nonresponse bias (where certain groups have higher response rates than others) has caused them to no longer yield representative samples. Thus, almost all phone-based polls adjust their data as well to yield more accurate results. In fact, a recent analysis of 2020 pre-election polling by the American Association for Public Opinion Research (AAPOR) found that "[no] mode of interviewing was

unambiguously more accurate", suggesting that online polling methodologies performed no worse than telephone-based polling methodologies in 2020 [7].

3. Meta-Data Review

For this report, a large user of panel-based online polling amount of the cited research was sourced from Pew Research Center, one of the most prominent pollsters in the United States, and both a user and developer of online polling methodologies.

Studies reviewed for this report include:

[2] Kennedy et al. 2020. Assessing the risks to online polls from bogus respondents.
This report describes the risk of bogus respondents to online polls, a phenomenon unique to them due to their use of incentives to get respondents to fill out their surveys. A significant number of respondents do not answer the survey seriously, and introduce bias into the survey results

[4] Keeter. 2019. Growing and improving Pew Research Center's American Trends Panel. This report describes in great detail Pew Research Center's American Trend Panel, their principal tool for conducting online polls. It includes their methodology for recruiting and interviewing respondents, as well as weighting results to produce public opinion data.

[6] Kennedy et al. 2016. Evaluating online nonprobability surveys.
This study evaluated various nonprobability samples for online surveys, and found that there exists a wide range of data quality, as well as patterns of bias in the data.

[8] Nate Silver. 2021. The death of polling is greatly exaggerated.
This article describes *FiveThirtyEight*'s methodology for rating pollsters. It includes their studies into the accuracy of various types of polling methodologies.

4. Conclusions

In 2021, *FiveThirtyEight*, a data journalism organization that also analyzes polls, made a big change to how they rated pollsters. In the past, live-caller phone-based pollsters were treated as more accurate than their online-based counterparts and were rated as such. However, in the wake of the 2020 U.S. elections, *FiveThirtyEight*'s research found this is no longer be true: "there's just not much evidence that live-caller polls are consistently outperforming other methods as far as poll accuracy goes" [8]. Accordingly, they removed this assumption from their ratings.

This reflects a trend in online polls where, through research and innovation, the art of polling has matured enough to become just as accurate as their traditional counterparts. While Nate Cohn's 2019 analysis may have found online polls to be unproven and slightly less accurate than live-called phone polls, both the AAPOR's analysis and *FiveThirtyEight*'s analysis of 2020 election polls found online pollsters at least no worse [3, 7, 8]. While there still exist problems specific to online methodologies, they have become a very viable solution to the rising cost of phone-based polls, and their widespread use should not be a cause for concern.

5. Future Work

FiveThirtyEight's analysis of poll accuracy found a significant divergence between the accuracy of online-only polls, and hybrid methodologies that incorporate both online and phone-based components. In their research, hybrid polls were found to be very accurate, and just as accurate as live-caller phone polls, while online-only polls were much less accurate [8]. Future studies into

the viability of online-only polls may be warranted, and may require separate treatment from the hybrid methodologies used by more prominent firms like Pew Research Center [4].

In addition, recent U.S. elections have seen the introduction of text messaging-based polls, where users are texted a request to complete a survey, usually online. *FiveThirtyEight*'s analysis of these polls found them to be highly accurate, but cautioned that there only exists a relatively small sample of polls [8]. Future research into this category of polls would likely be very enlightening.

References

[1] Pew Research Center. 2008. The Impact Of "Cell-Onlys" On Public Opinion Polling. Retrieved from http://pewrsr.ch/10OIRLC

[2] Courtney Kennedy, Nick Hatley, Arnold Lau, Andrew Mercer, Scott Keeter, Joshua Ferno, and Dorene Asare-Marfo. 2020. Assessing the Risks to Online Polls From Bogus Respondents. Retrieved from https://pewrsr.ch/2tKW2Et

[3] Nate Cohn. 2019. No One Picks Up the Phone, but Which Online Polls Are the Answer? Retrieved from https://t.co/2h6LU3Uyql

[4] Scott Keeter. 2019. Growing and Improving Pew Research Center's American Trends Panel. Retrieved from https://pewrsr.ch/2SoPHon

[5] Scott Keeter, Nick Hatley, Courtney Kennedy, and Arnold Lau. 2017. What Low Response Rates Mean for Telephone Surveys. Retrieved from https://www.pewresearch.org/methods/2017/05/15/what-low-response-rates-mean-for-telephone-surveys/

[6] Courtney Kennedy, Andrew Mercer, Scott Keeter, Nick Hatley, Kyley McGeeney, Alejandra Gimenez. 2016. Evaluating Online Nonprobability Surveys. Retrieved from https://www.pewresearch.org/methods/2016/05/02/evaluating-online-nonprobability-surveys

[7] Josh Clinton, Jennifer Agiesta, Megan Brenan, Camille Burge, Marjorie Connelly, Ariel Edward-Levy, Bernard Fraga, Emily Guskin, D. Sunshine Hillygus, Chris Jackson, Jeff Jones, Scott Keeter, Kabir Khanna, John Lapinski, Lydia Saad, Daron Shaw, Andrew Smith, David Wilson, Christopher Wlezien. 2021. Task Force on 2020 Pre-Election Polling: An Evaluation of the 2020 General Election Polls. Retrieved from https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/AAPOR-Task-Force-on-2020-Pre-Election-Polling_Report-FNL.pdf

[8] Nate Silver. 2021. The Death of Polling Is Greatly Exaggerated. Retrieved from https://53eig.ht/3tIkO1k