Public Polling of Perceptions of Higher Education: The Methodological Landscape

Tetsuya Yamada

Ph.D. Student

University of Pittsburgh

August, 2023

This report was generously supported by the EdFunders LEAD Impact Group through the Society for Research in Educational Effectiveness (SREE) summer fellowship. A special acknowledgment goes to Chelsea Koproske at the Bill & Melinda Gates Foundation, Kimberly Lent Morales at the Helios Foundation, Laura M. Stapleton at the University of Maryland, and Jenn Kush at SREE for their guidance and support. The information presented here does not represent the views of the EdFunders LEAD Impact Group, the Bill & Melinda Gates Foundation, the Helios Foundation, SREE, the University of Maryland, and the University of Pittsburgh. I am responsible for all errors. Address correspondence to the author; e-mail: tey15@pitt.edu.

Executive Summary

Despite public perceptions of higher education having been increasingly surveyed in public polls, public polls have come under critical scrutiny due, in part, to the misprediction of major polls in the 2016 presidential election. Polling Americans' opinions on higher education is still essential because higher education is considered a public good and largely depends on public funding. Given such social contexts, survey users must understand polling methodology and how best to utilize polling methods in their decision-making. The present study aims to provide consumers of survey findings with a comprehensive view of the current methodological landscape of public polls on perceptions of higher education.

American Public Opinions on Higher Education

Based on expert knowledge and online search, the present study identified public polls on higher education by eight organizations across time (1993-2023).

- Polls predominantly reported a low and decreasing confidence/sentiment in higher education. In a Gallup survey, 36% of U.S. adults expressed "a great deal/quite a lot" of confidence in higher education in 2023, dropping from 57% in 2015.
- Conservatives have less favorable views on higher education. New America recently reported that 78% of Democrats in 2023 believed that postsecondary institutions were leading the nation in a positive direction, but only 41% of Republicans felt the same.
- Although higher education has diverse benefits, polls predominantly surveyed the perceived private economic benefits and their cost-benefit implication. It was reported that many Americans perceived the economic benefits of higher education but questioned whether it was worth the financial outlay.

Three Common Polling Methods and Their Strengths

1. Random digit dialing (RDD): This is a probability sampling method in which telephone numbers are randomly picked. Interviewers can assist and encourage respondents during the telephone survey. RDD can help ascertain the most recent public opinion because of its ability to rapidly collate responses.
2. Probability-based panels: These are national survey panels formed through the random sample selection from a population database. A typical sampling method is address-based sampling (ABS) with a high household coverage rate (90–98%).
3. Online opt-in polls: These differ from probability sampling in that respondents are recruited based on accessibility (e.g., online ads). They use statistical adjustments, such as matching and weighting, to make inferences of the population estimates. Conducting online opt-in polls is considered inexpensive and time-efficient.

RDD has been the "gold standard" for assessing public opinions of higher education and is still a common method. Some polls have recently shifted from RDD to probability-based panels.

Three major methodological concerns

1. Construct validity*: The same construct may not be measured across sentiment question items.* There was a variation in public sentiment toward higher education across surveys (25–67%). *The same question item may measure different constructs depending on individuals.* Higher education is not a single entity (e.g., universities vs. community colleges) and has different layers (e.g., associate, bachelor, and graduate). The entities and levels a respondent has in their mind affect their general attitude toward higher education. *Some question items on the benefits of higher education need to be clarified.* Items need to be more specific to answer *what* the benefits of higher education are.

2. Low response rates for RDD and probability-based panels: The response rate for RDD is low and has been consistently dropping. Only 8.9% of landline and 3.2% of cell phone samples responded in a 2019 Pew poll. For probability-based panels, the final response rate tends to be relatively higher than for RDD, but the nonresponse pattern is complex. Nonresponse occurs at the panel recruitment, panel attrition, and survey participation stages.
3. Reliability of statistical adjustments for online opt-in polls: These polls claim the representativeness of estimates based on statistical modeling rather than the design of the sampling frame.

Best practices

1. Construct validity: Pollsters should specify constructs of public attitudes toward higher education.
   First, categorize public attitudes as follows:
   -public confidence/sentiment
   -benefits of higher education
   -specific issues regarding higher education

   The benefits of higher education are further specified as four types:
   -private economic benefits (e.g., individuals' income and employment)
   -private non-economic benefits (e.g., health and happiness)
   -public economic benefits (e.g., economic growth and tax revenues)
   -public non-economic benefits (e.g., democracy and reduction in health care costs)

   For each category of public attitudes, poll creators can then decide if they are interested in *general* higher education or *specific* entities (e.g., four-year universities or community colleges) and levels of higher education (e.g., bachelor's or graduate).

2. Low response rates: A low response does not necessarily mean nonresponse bias and increasing the response rate does not necessarily reduce nonresponse bias, which exists only when the variable of interest is correlated with the response propensity. Pollsters should assess the potential for nonresponse bias occurring as below:

   -comparison of the response rate across subgroups
   -comparison of respondents' characteristics with ones in a benchmark survey
   -comparison of panelists' characteristics before and after panel attrition and survey participation (for probability-based panels)
   -comparison of respondents' characteristics before and after the follow-up recruitment

3. Reliability of statistical adjustment: Sample matching may infer the population estimate from online opt-in panels. Successful sample matching, however, requires sufficient time, resources, and expert knowledge and needs to meet certain conditions:

   -the use of sample matching based on target probability sampling
   -the availability of variables in both the panel and the target sample
   -the use of a large and diverse sample to allow close matching
   -the use of post-survey weight to account for the remaining unrepresentativeness

Recommendations for Survey Users

Topline message: Public polls should be continuously used to gauge public perceptions of higher education despite their limitations. Despite potential nonresponse bias, the sample estimate from the probability-based sampling frame can infer the representative public opinions in theory. Further, higher education is considered a public good and largely supported by public tax, so survey users must understand the general public's opinions of higher education. Public polls are unique survey methods for gauging these opinions.

With this topline message, here are recommendations on using public polls for survey users.

Clarify the goals and purposes of your study. Identifying your research objectives is beneficial in determining what polling methods to use and in what way.

Check if question items correctly capture the construct you are interested in. Carefully consider how item wording and polling settings affect the interpretations of respondents in the answering process.

Analyze political contexts that may affect the answers when interpreting the result. Public confidence and sentiment toward higher education are susceptible to the short-term evaluation of political events and leaders; thus, understanding current political contexts regarding higher education is vital in interpreting the polling result.

Understand where survey errors are likely to occur in your polling method. Highlight potential survey errors in polling by using the total survey error framework as in this present paper. Investigate how errors were addressed in the pollsters' reports.

Large amounts of nonresponses do not necessarily lead to a nonresponse bias. A nonresponse bias is relatively absent if there is no association between the variable of interest and the response propensity. Likewise, attempts to increase response rates (e.g., through cash incentives) may only reduce the nonresponse bias if sample members who would otherwise refuse to respond are recruited.

Telephone surveys are not "dead" despite large numbers of nonresponses. Respondents' characteristics can be compared with those in a benchmark survey with a higher response rate. Weighting can also help adjust for sample demographics based on knowledge of who is likely to respond in telephone surveys. If this analysis partly justifies the representativeness of completed cases, telephone surveys can still be used.

Probability-based panels are promising but be wary of the nonresponse mechanism. While the higher coverage of the sampling frame in probability-based panels helps reduce coverage error, the complex nonresponse process during recruitment, panel attrition, and poll participation may introduce potential nonresponse bias.

Online opt-in polls are not advisable for inferring population estimates. Online opt-in polls appear inexpensive initially, but ensuring the representativeness of the sample estimate of high-quality polls, like those using sample matching, demands time, budget, and expertise.

Online opt-in polls can be used for exploratory studies and experiments. Quick and inexpensive online opt-in polls can be used for testing the construct validity of measures and the influence of information interventions on public perceptions.

Contents

Terminology

Address-based sampling (ABS): ABS is a sampling method that randomly recruits survey samples from residential address lists (Dillman et al., 2014). The list typically comes from a computerized delivery sequence file (CDSF) maintained by the U.S. Postal Service (USPS). CDSF can differentiate between residential and business addresses.

Probability sampling: Every individual in a sampling frame is assigned a known and nonzero likelihood of being selected for the sample, which enables the survey's findings to be representative of the entire target population (Dillman et al., 2014).

Probability-based panels: Probability-based panels are national survey panels formed by randomly selecting samples from a population database (Kennedy et al., 2023). A typical sampling method is ABS. Some panels recruit samples from random digit dialing. Once individuals join the panel, they participate in multiple surveys over time.

Online opt-in polls: Respondents are not randomly selected from the general population. Instead, they are gathered from diverse online sources, such as advertisements on social media or search engines, websites that offer incentives for participating in surveys, or voluntary enrollment in an opt-in panel (Kennedy et al., 2023). Certain opt-in samples are obtained from non-probability-based panels.

Random digit dialing (RDD): RDD is a probability sampling method that typically generates four-digit numbers randomly after selecting area codes and prefixes based on a geographical area (Dillman et al., 2014). As the number of cell phone users increased, the typical RDD had dual-frame sampling for landlines and cell phones.

Sampling frame: A sampling frame is a list of a population of interest from which sample members are drawn (Kennedy et al., 2023). Coverage error occurs if the sampling frame does not include all members of the population.

Sample matching: A benchmark random sample is chosen from a population, then respondents from the panel who are closely matched to the ones in the random sample are selected for the survey (Rivers, 2007).

Total survey error (TSE): This is the collection of all errors that may arise in a survey's design, implementation, processing, and analysis (Biemer, 2010). A survey error is defined as the deviation of a survey response from the underlying true value (Biemer, 2010). The deviation can occur due to survey bias (deviation of the expected value of an estimate from its true value) or survey variance (variations over replicated survey implementation in the departure of the actual estimate from its expected value) components (Groves, 2006).

-**Construct validity (specification error)**: The concept implied by the survey question contrasts with the concept that was intended to be measured (Biemer, 2010).

-**Measurement error:** The difference between the estimate produced and the true value because respondents answer inaccurately to survey questions. (Dillman et al., 2014). This contains errors from interviewees, interviewers, survey questions, and several interview factors (Biemer, 2010).

-**Processing error:** This covers errors in editing, data input, coding, assignment of survey weights, and summary of the survey data (Biemer, 2010).

-**Coverage error:** This occurs when the list from which sample members are drawn does not accurately represent the population on the characteristics. Coverage is influenced by both who has access to the mode and what lists of frames are available to sample members of the target population (Dillman et al., 2014).

-**Sampling error:** This is the difference between the estimate produced when only a sample of units on the frame is surveyed and the estimate made when every unit on the list is surveyed (Dillman et al., 2014).

-**Nonresponse error:** This is the contrast between the estimate obtained when only a portion of the sampled units respond and the estimate when all of them respond (Dillman et al., 2014).

Weighting: Weighting is used to adjust the sample distribution of variables to match the one for the population distribution.

-**Poststratification:** Poststratification adjusts weights by calculating the proportion of the sample size to the estimated population total within some categories (Baker et al., 2013).

-**Propensity score adjustment:** Propensity score adjustment estimates the probability that cases come from online opt-in polls by combining online opt-in poll datasets with benchmark population ones and fitting a statistical model (Mercer et al., 2018).

-**Raking:** With raking, pollsters select a group of variables for which the population distribution is already established. Then, through an iterative process, the weights assigned to each case are modified until the distribution of the sample matches that of the population for those variables (Mercer et al., 2018).

1. Background

"American Confidence in Higher Ed Hits Historic Low" (Blake, 2023).

"Americans Are Losing Faith in College Education" (Belkin, 2023).

Such sensational headlines catch our attention in the current news media. Public polls and media gauging the public perception of U.S. higher education is a relatively recent trend. For instance, Gallup continuously measured public confidence in institutions, such as Congress, the church, and media, back in the 1970s; however, it was in 2015 that Gallup started to assess levels of confidence in higher education in America (Gallup, n.d.). The public perceptions of higher education have been increasingly measured in polls at a time when issues in higher education, including affirmative action in college admissions and student loan cancellation, have garnered national attention.

Public skepticism and negative images of higher education are harmful to its organizations' legitimacy. Institutions can achieve legitimacy when their existence is taken for granted (Loveless, 1997). As almost all colleges rely on public funds, public support influences the allocation of these resources (e.g., federal support for financial aid and research funds) (Association of Governing Boards of Universities and Colleges, 2018). Universities may also lose their revenue base if those who doubt its value refrain from sending their family members to higher education.

Despite increasing attention on gauging public perceptions of higher education in public polls, such polls face increased public scrutiny. Most falsely predicted that Hillary Clinton would defeat Donald Trump in the 2016 U.S. presidential election, and many Americans were left wondering whether polling was flawed and what actions pollsters could take (Kennedy et al., 2023). Since 2016, the majority of national public polls have changed

their methods and they have become more diversified (Kennedy et al., 2023). Under such social contexts, survey users must understand these diverse methods of polling perceptions of higher education and their methodological strengths and weaknesses.

Although the public may doubt their efficacy, public polls are unique survey methods for gauging the representative opinions of Americans on higher education. There are other survey methods, such as expert opinion, focus groups, and proxy information (e.g., rates of enrollment in higher education) to gauge public perceptions of higher education; however, these methods cannot capture the general public's attitude. As higher education may contribute to public benefits (e.g., national development and democracy) and largely depends on public funding, it is essential to gauge Americans' overall opinion on this matter.

This study aims to present consumers of survey findings with an overview of the current methodological landscape of public polling on perceptions of higher education. The intention is to provide insight for survey consumers who use data from public polls in their decision-making regarding higher education policies and practices. The structure of the paper is as follows. First, it summarizes Americans' opinions on higher education and introduces public polling methods on perceptions within this field. Second, the methodological strengths and weaknesses of those methods are discussed based on the total survey error framework and three critical methodological concerns are highlighted. Third, best practices for tackling these concerns are discussed. The paper concludes with recommendations for interpreting and using public polls. The present article is not a comprehensive report on survey methodology in public polls, and only topics that are unique to public polls on perceptions of higher education are covered. Thus, other issues such as question order, methodological details of weighting methods, open-ended questions, item scales, methods to detect poor responses, and questionnaire design are not discussed. Readers should consult survey methodology literature (e.g., Dillman et al., 2014; Groves et al., 2004) for a deeper understanding of these topics.

2. General Findings

2.1. Summary of American Opinion on Higher Education

Various public attitudes towards higher education have been surveyed in public polls. Some have investigated confidence in higher education using holistic institutional confidence polls, while others have studied it with polls regarding higher education. The present paper summarized Americans' opinions on higher education in three categories: 1. general confidence/sentiment in higher education, 2. benefits of higher education, and 3. opinions on specific issues in higher education. Based on expert knowledge and online search, the present study identified public polls on higher education by eight organizations across time (1993-2023). This summary is not an exhaustive list of all public polls on perceptions of higher education. Instead, it provides an overall picture of the major public polls. As this paper aims to gauge the American public's general views on higher education, studies focusing on American adults were included in the analysis.

Table 1 summarizes the general confidence/sentiment in higher education. In general, polls reported a low and decreasing confidence/sentiment in higher education. Gallup started to measure confidence in higher education in 2015 by directly asking, "How much confidence [do] you, yourself, have in higher education." Institutional confidence items on other political and social institutions have been used in Gallup polls since the 1970s. It reported that 36% of U.S. adults expressed "a great deal/quite a lot" of confidence in higher education in 2023, decreasing significantly from 48% in 2018 and 57% in 2015 (Brenan, 2023).

Similarly, the public's directional feelings on higher education have been investigated in several surveys. For instance, the Pew Research Center (2019a) reported in 2019 that 50% of U.S. adults said college/universities have a positive effect on the way things are going in the

country these days, which had dropped from 55% in 2017 and 61% in 2010. However, there is a considerable variation in general sentiment (25–67%) across surveys. Question items differ across studies, and this issue is discussed further below.

Furthermore, it is common for pollsters to report the difference in attitudes towards higher education by subgroups, among which political polarization is prominent. New America recently reported that 78% of Democrats in 2023 believed that postsecondary institutions were leading the nation in a positive direction, but only 41% of Republicans felt the same (Nguyen et al., 2023). Differences in opinion by group, such as generation, gender, race, and income, were reported in some polls.

Table 1: Summary of general confidence/sentiment in higher education (HE)

	Main findings	Period	Question items	Sample
1. Gallup (2023)	A total of 36% of U.S. adults expressed "a great deal/quite a lot" of confidence in HE in 2023. It had dropped from 48% in 2018, 57% in 2015, and 44% in 2017. It showed that 59% of Democrats had confidence in 2023, with 19% for Republicans and 32% for independents.	2015, 2017, 2018, 2023	Please tell me how much confidence you, yourself, have in higher education. (a great deal/quite a lot/some/very little)	Probability sample: RDD for both landline and cell phones. 1,013 in 2023 Min. Quota: landline: 20% and cell phone: 80% Weighting: probabilities of selection, nonresponse, sample demographics, and double coverage of landline and cell phones
2. Pew (2019b)	A total of 50% of U.S. adults in 2019 said colleges/universities have a positive effect on the way things are going in the country these days. This had dropped from 55% in 2017, 63% in 2015 and 61% in 2010. While 67% of Democrats /Democrat-leaners held a favorable view in 2019, this was 33% for Republicans/Republican-leaners.	2010, 2012, 2015–2017, 2019	Are colleges/universities having a positive or negative effect on the way things are going in the country these days? (positive/negative/neither/don't know)	Probability sample: RDD for both landline and cell phones 1,502 in 2019 (landline: 302, cell phone: 1200) Weighting: probabilities of selection by household size and telephone usage, double coverage of landline and cell phones, and raking to match sample demographics
3. New America	About 25% of Americans believed that HE was fine the way it was in 2017/2018 (strongly/somewhat agree) (Fishman et al., 2018; Fishman et al., 2017). In addition, 24% of liberals and conservatives believed that HE was fine how it was in 2018. Follow-up questions asked reasons why colleges are not fine how it was. The top three reasons for liberals were "too expensive," "should be free for all," and "room for improvement." For conservatives, they were "too expensive," "room for improvement", and "degree is needed for a better life."	2017, 2018	Higher education in America is fine how it is. (Strongly/somewhat agree, neither agree nor disagree, strongly/somewhat disagree)	Probability sample: RDD for both landline and cell phones 1,000+600 oversample in 2018 (half of them cell phones) Oversamples of Black, Latino, Asian, and residents of North Carolina Weighting: information not available
4. New America /NORC	59% of U.S. adults believed that postsecondary institutions were leading the nation in a positive direction in 2023. It had risen from 55% in 2022 but dropped from 69% in 2020 (Nguyen et al., 2023). As well as this, 78% of Democrats in 2023 believed there was a positive effect, while this was 41% for Republicans. A total of 41% of Americans believed that HE is fine the way it is in 2023 (Democrats: 42% and Republicans: 40%).	2019–2023	Overall, do you think colleges and universities are having a positive or negative effect on the way things are going in this country? (Positive/Negative)	Probability sample: multistage probability sample of U.S. households in 2010 and additional address-based sample from AL, IO, ND, and WY. A total of 48 sampling strata by demographics were used to collect a sub-sample of panels. 1,497 in 2023 (web: 1,422 and phone: 75) (NORC, 2022) Oversamples of Black, Latino, and Asian Weighting: probabilities of selection, nonresponse, and sampling demographics
5. GBH (2018a,2018b)	67% of U.S. adults had a strongly/somewhat favorable view of American colleges/universities.	2018	Generally speaking, do you have a favorable/unfavorable impression of colleges and universities? (Strongly/Somewhat favorable) (Strongly/Somewhat unfavorable)	Probability sample: RDD for both landline and cell phone 1,002 (350 landlines and 652 cell phones) Weighting: probabilities of selection by telephone usage and double coverage of landline and cell phones. Raking to match sample demographics

Note: National Opinion Research Center at the University of Chicago (NORC). Great Blue Hill (GBH). Random digit dialing (RDD).

Table 2 summarizes findings on public perceptions of the benefits of higher education and specific issues in higher education. First, many of the polls asked about the perceived benefits and cost-benefits of higher education. In America, higher education has been considered a public good (e.g., contributing toward economic growth, democracy, and informed citizenry) in addition to its role as a private good (e.g., cultivating human capital and meeting workforce needs) (Leveille, 2006). It is therefore critical to categorize the benefits of higher education into "public" and "private." Each benefit can be further divided into economic and non-economic (social) subcategories. As outlined in Figure 1, there are four categories: private economic benefits, private non-economic, public economic, and public non-economic. Although higher education has diverse advantages, the polls this paper analyzed predominantly surveyed the perceived private economic benefits and their implication in terms of cost-benefit. The general discourse is that the public recognizes the private economic benefits of higher education, but many wonder if these outweigh the costs. For instance, The Wall Street Journal and National Opinion Research Center at the University of Chicago (NORC) (2023) showed that only 42% of U.S. adults believed a four-year college degree is worth the financial cost with regard to job prospects and life-long income. Similarly, the 2022 survey by New America revealed that only half of Americans (52%) believed that students could access an affordable, high-quality higher education after high school, although 76% of Americans thought that education beyond high school provides a good return on investment (Fishman et al., 2022).

In 2023, New America added new questions to their survey to gauge perceived private (e.g., good health) and public non-economic benefits of higher education (e.g., civic engagement and support for local business) (Nguyen et al., 2023).

Their findings indicated that people recognized the private economic benefits of higher education more than private non-economic ones. About 70% of Americans believed that undergraduate credentials were helpful for achieving a greater income from employment and better access to jobs with a decent wage (Nguyen et al., 2023). In contrast, only about half felt that those with credentials were satisfied with the communities they lived in and had better health compared to those without credentials (Nguyen et al., 2023). There was no notable difference in the perceived public economic (e.g., skilled workforce) and non-economic (e.g., civic participation) benefits of higher education.

In addition to the perceived benefits of higher education, many polls, such as New America and Public Agenda, investigated various issues related to higher education. These included the government's responsibility in funding higher education and equal access to higher education. New America and Great Blue Hill specifically gauged opinions on admission policies of higher education, such as the use of affirmative action in admissions and test-optional admission.

Figure 1: Four benefits of higher education

1. **Private economic benefits**: benefits of higher education on individuals' earnings, employment, levels of savings, working conditions, and mobility
2. **Private non-economic benefits**: individual benefits of higher education on health, longevity, happiness, leisure time, access to information about consumer choices, and personal status
3. **Public economic benefits**: benefits of higher education on economic growth, tax revenues, availability of labor, and independence from government support at the society level
4. **Public non-economic benefits**: benefits of higher education at the society level on civic engagement, availability of leaders from diverse communities, charitable giving, and community service, voting, social cohesion, and reduction in public health care costs

Sources: Pusser & Doane (2001) and McMahon (2009).

Table 2: Summary of benefits of higher education (HE) and opinions on specific issues

	Main findings	Period	Aspects of higher education	Sample
New America New America/ NORC	76% of U.S. adults in 2022 believed that education beyond higher education offers a good return on investment for students (Nguyen et al., 2023). However, only about half (52% in 2022 and 53% in 2023) believed students could access an affordable, high-quality education after high school.	2017-2023	Four benefits of HE Cost-effectiveness of HE Affordability Public money spent on HE Accountability of tax money spent on HE Equal access and diversity in campuses Perceived quality/costs of online HE Test-optional policy	Same as New America for 2017–2018 and New America/NORC for 2019–2023 in Table 1
WSJ/NORC (2023)	Only 42% of U.S. adults answered that a four-year college degree is worth the cost in terms of job prospects and lifetime income.	2023	Cost-effectiveness of HE Affirmative action in college admission	Same as the New America/NORC survey in Table 1 1,019 (web: 980 and phone: 39)
Public Agenda Public Agenda /USA Today	Only 26% of U.S. adults in 2022, and 42% in 2016, responded that a college education is necessary to succeed in today's work world (Schleifer et al., 2022; Schleifer et al., 2016). This was a decrease from 55% in 2008–2009 but an increase from 31% in 2000. Moreover, 51% doubt the private economic benefit of HE investment in 2022.	1993, 1998, 2000, 2003 2007–2009, 2016, 2018, 2019, 2021, 2022	Importance of HE in success Private economic benefits of HE Cost-effectiveness of HE Public benefits of HE Affordability Equal access HE's financial management Public money spent on HE Accountability of HE outcomes Political climate in HE College career support	**1993–2016:** probability sample: RDD 1,006 (landline and cell phone) Weighting: probabilities of selection and demographics **2018–2019:** probability sample: information not available 1,000 (phone: 750, web: 250) Weighting: demographics **2021–2022:** non-probability sample: a combination of Ipsos's online panel, the other online panel, and river sampling (Ipsos, 2019) 1,662 adults in 2022 Weighting: raking on demographics
Teachers College	69% of American adults said public spending on HE had been an excellent or good investment in 2023. This had dropped from 76% in 2017 (Drezner & Pizmony-Levy, 2023; Drezner et al., 2018). A total of 46% of Democrats believed it was an excellent investment in 2023, while 24% of Conservatives believed so.	2017 2023	Public money spent on HE Four benefits of HE	Non-probability sample: Qualtrics online panel The sample is compiled using overall demographic quotas based on census percentages for representation 2019 in 2023 Oversample of Black, Asian, and Latinx. Weighting: household and group quarters
GBH (2018a, 2018b)	68% of adults agreed that college was worth attending, considering costs. In addition, 72% disagreed with the past Supreme Court decision that colleges can use race as one factor in deciding which applicants to admit.	2018	Cost-effectiveness of HE Affirmative action/admission criteria Diversity in student bodies Public money spent on HE Political climate/free speech Student support	Same as GBH in Table 1
Bipartisan Policy Center/AACU	A total of 60% of U.S. adults believed college degrees were worth the time and money involved (Finley et al., 2021).	2021	Private benefits of HE Cost-effectiveness of HE Skills HE should cultivate	Non-probability sample: online 2,200 Americans Weighting: demographics

Note: Wall Street Journal / National Opinion Research Center at the University of Chicago (WSJ/NORC). Great Blue Hill (GBH). Association of American Colleges and Universities (AACU). Random digit dialing (RDD).

2.2. Three Common Polling Methods

Three major polling methods are used to gauge public perceptions of higher education: random digit dialing (RDD), probability-based panels, and online opt-in polls. These three methods, however, are not mutually exclusive. Some poll makers (e.g., Public Agenda/USA Today) blended these methods and used multiple sample sources or interview modes. Table 3 shows the characteristics of these three polling methods. RDD is a telephone survey and has been the "gold standard" for assessing public opinions and is still a common polling method. Telephone surveys became common once most U.S. adults were reachable by telephone (Tourangeau, 2004). As of 2021, 97% of U.S. adults owned cell phones (Pew, 2021). As in Table 1, the latest polls by Gallup and the Pew Research Center used RDD.

Some polling organizations, such as New America and Public Agenda, have started to use probability-based panels to gauge public opinions on higher education. New America shifted from RDD to probability-based panels in 2019 when it began collaborating with NORC. Such a shift corresponds to the recent increase in the use of probability-based panels in polling industries in the U.S. (Kennedy et al., 2023). Residential addresses obtained from the U.S. postal service (USPS)'s computerized delivery sequence file (CDSF) are common sampling frames for probability-based panels. NORC's AmeriSpeak panel added additional addresses by in-person field listing for areas where CDSF coverage was inadequate (NORC, 2022). Phone calls and in-person visits also enhanced NORC's recruitment process. Standard interview modes were in the form of online surveys for probability-based panels, but poll organizers provided alternative

options for non-internet users. NORC offered both online and telephone interviews, as 8% of active panelists preferred to participate in telephone surveys while 92% preferred online (NORC, 2022). KnowledgePanel for Public Agenda's poll provided a web-enabled device and free internet service for households without internet access (Ipsos, n.d.).

The final and least popular method is the online opt-in poll. As Table 3 clearly shows, this is not a form of probability sampling; thus, selection bias is the biggest concern. A sampling frame does not exist because the sample is selected based on availability and accessibility. To adjust for the representation of the estimate, those who use online opt-in polls attempt to match the sample to population demographics benchmarks by using adjustment methods such as weighting.

Table 3: Characteristics of three major polling methods of gauging public perceptions of higher education

	RDD	Probability-based panels	Online opt-in polls
Probability sampling	Yes	Yes	No
Sampling frame	Phone numbers	Residential addresses	Not available
Sampling unit	Household/individuals	Household	Individual
Recruitment mode	Telephone	Mail	Online
Interview mode	Telephone	Online/Telephone	Online
Interviewers	Yes	Yes (for a telephone survey)	No

Note: Random digit dialing (RDD)

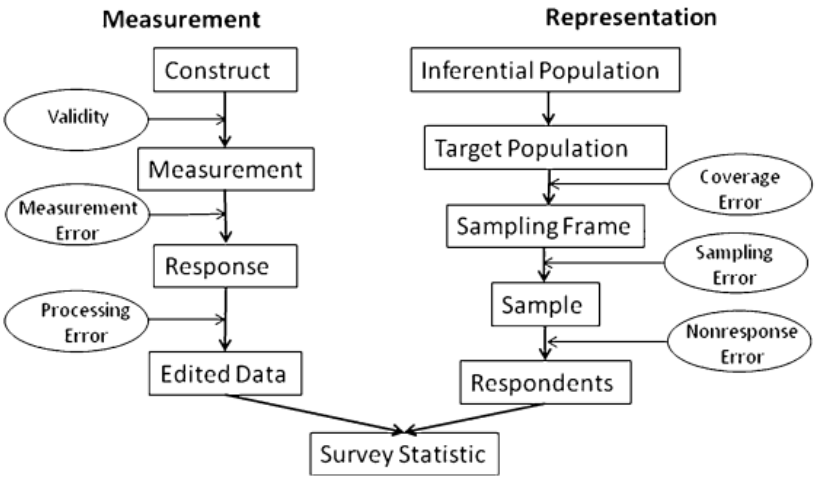
3. Methodological Strengths and Weaknesses

The present paper discusses the methodological strengths and weaknesses of three primary methods of polling perceptions of higher education, using the total survey error (TSE) framework as an analytical framework.

3.1. Total Survey Error Framework (TSE)

TSE refers to the collection of all errors that may arise in the survey design, implementation, processing, and analysis (Biemer, 2010). A survey error is defined as the deviation of a survey response from the underlying true value (Biemer, 2010). The deviation can occur due to survey bias (deviation of the expected value of an estimate from its true value) or survey variance (variations over replicated survey implementation in the departure of the actual estimate from its expected value) components (Groves, 2006). The TSE framework aims to improve our understanding of potential errors in each survey step and minimize them within the limited time and resources available to the pollsters. There is a trade-off between the reduction of errors and considerations of cost and time. Although surveyors need to understand TSE, the TSE framework does not specify the ideal threshold of error levels (Groves & Lyberg, 2010).

Figure 2: Total survey error components in measurement and representation inference processes



Source: The original figure appears in Groves et al. (2004, p.42), modified by Groves and Lyberg (2010, p. 856).

Figure 2 shows six survey errors in the measurement and representation inference processes. The measurement dimension delineates what data should be gathered about the observation units in the sample (Groves et al., 2004). The representational dimension focuses on the populations the survey depicts (Groves et al., 2004). Three errors can occur in the measurement process. Ideally, pollsters should start from the construct implied by the survey questions before writing the question items. A *construct validity* issue (specification error) arises if items do not measure the construct. Next, polltakers obtain responses from question items. A *measurement error* occurs if a difference between the measured and true response appears due to survey settings, respondents' motivation, and/or question items. A *processing error* occurs in the editing of responses because of erroneous data entry or survey weighting techniques. Three other types of error can occur in inferring representation. A *coverage error* occurs when some population members are not listed in the sampling frame. For instance, those who do not have phone numbers are

not in the sampling frame for RDD. Also, a family that recently moved may not be listed in the recent CDSF. Such cases cause coverage errors. A *sampling error* occurs when using a sample estimate instead of a population estimate. Sampling errors can occur in any survey, but a relatively large sample size can reduce their occurrence. Even after careful sampling, pollsters cannot guarantee that all of the members in a sample will respond to the poll. A *nonresponse error* occurs when respondents are different from nonrespondents in a way that affects the estimate (Dillman et al., 2014).

3.2. Mapping Methodological Strengths and Weaknesses

Based on the presented TSE framework, this study identified the methodological strengths and weaknesses of three major methods of polling public attitudes toward higher education (See Table 4 for a summary).

Construct Validity. Validity issues arise across all three methods. First, the same construct may not be measured *across* sentiment question *items.* The sentiment for the "fine the way it was" item was lower than others. Table 1 shows that about half of Americans had a positive feeling and confidence in higher education around 2018. However, only 25% of Americans believed that higher education was fine the way it was in New America's previous surveys in 2017 and 2018. There is a partisan gap in public sentiment in higher education for other items, but there is no partisan gap in the "fine how it is" item (Fishman et al., 2018; Fishman et al., 2017). Thus, this question item appears to measure a different construct than other sentiment and confidence questions.

Furthermore, the same item may measure different constructs *across people*. Higher education is not a single entity (e.g., universities vs. community colleges and Public vs. Private) and has different layers (e.g., associate, bachelor, and graduate). Entities and levels respondents have in their minds significantly affect their general attitudes toward higher education. Indeed, individuals would recall their own constructs when answering the general sentiment and confidence question. Although liberals and conservatives sometimes cited the same reason (cost) why higher education was not satisfactory, they also drew different reasons (e.g., equal access issues for liberals and degrees for a better life for conservatives). This suggests that people have different issues in their minds when they are asked to answer a question pertaining to sentiment.

Additionally, some question items on the benefits of higher education were unclear. Below are some sentences related to the benefits and cost benefits of higher education. Respondents answered their agreeableness to the above sentences in polls.

"A college education is still the best investment for people who want to get ahead and succeed" (Schleifer et al., 2022, p. 10).

"Education beyond high school offers the student a good return on investment for the student" (Fishman et al., 2022).

"Public four-year colleges or universities are worth the cost" (Nguyen et al., 2023).

Pollsters may seek to gauge the private economic benefits of higher education with the first sentence; however, whether this question concerns success in the workforce or in more general terms is unclear. Similarly, the second and third sentences may concern the cost-benefits of higher education in terms of private economic benefits.

However, some respondents may still consider higher education to be a good investment or worth the cost because of academic knowledge and critical thinking skills they gain.

Measurement Error. Measurement errors may occur across the three study types due to several factors, such as unclear question items, question structure (e.g., all-that-apply questions result in fewer selections), the wording and the order of items, the response option, and the visual layout.

The existence of interviewers is an issue for measurement error. Interviewers tend to be present in RDD. They can assist respondents in clarifying the meaning of questions if needed. Respondents may become engaged in the survey and answer accurately if interviewers explain the study's rationale and encourage them to participate. However, they may hide their true opinions under the existence of interviewers if question items are socially desirable and sensitive. Estimates become inconsistent across surveys when the interviewing process is not standardized.

Such interviewer errors do not occur for online probability-based panels and opt-in surveys. Online surveys are, in contrast, susceptible to respondents' skipping answers because self-administered surveys require more effort to answer. Online or mailing surveys may send cash incentives to respondents. Such incentives may increase errors in responses, because some participants may rush to receive the financial reward.

Conditional effects may occur for a panel survey (response changes due to having completed multiple surveys). This may be beneficial because participants can navigate questionnaires more efficiently in later studies. It may also have a negative effect because people produce less optimal answers when they reduce their amount of effort, leading to biased and inconsistent responses (Dillman et al., 2009).

Processing Error. Almost all polls reported the use of a weighting adjustment. Diverse methods were used to adjust for probabilities of selection, nonresponse, and representativeness of the survey sample characteristics (e.g., gender, age, and race). Some reported weighting methods in detail, but others did not. The weighing quality depends on the availability of correct and observable covariates (Dillman et al., 2009). It should also be noted that weighting may decrease biases but at the cost of increased variability in the estimates (Dillman et al., 2014). For instance, incorporating a large number of variables during weighting adjustment leads to a generally larger variability (Mercer et al., 2018). Also, the number of landline users has decreased (Van Dam, 2023), but the landline weight for population characteristics cannot be applied to the cell phone sample because cell phone-only users differ fundamentally from the landline sample (Berinsky, 2017).

Representation Inference. As in Table 4, the three errors in the representation inference process are not applicable to online opt-in polls because no sampling frame is constructed for non-probability sampling. The absence of a sampling frame is the biggest concern for inferring estimates that appear representative of the entire adult population in the U. S. These online opt-in polls tend to claim the representativeness of estimates by applying statistical adjustment. This claim is not based on the design of the sampling frame of probability sampling but on statistical modeling (e.g., raking adjustment, matching procedure, or propensity model). Those who advocate online opt-in polls insist that the problem of nonprobability sampling is equally significant as the challenge encountered by RDD due to the skepticism towards its non-representativeness stemming from the sizeable nonresponse rates in telephone surveys (Weisberg, 2018).

Coverage Error. RDD typically has dual-frame sampling for landline and cell phone numbers. If households with landlines and cell phones overlap in both frames, they are more likely to be listed as a sample. The area and exchange codes are listed based on geographical codes in RDD; however, they may not align with actual residency for cell phones because people often do not update their cell phone numbers after they move to a new location (Dillman et al., 2014). This can cause the undercoverage of people in some geographical locations, leading to selection bias.

The coverage rate for households is high for probability-based panels based on ABS. USPS's CDSF covers 90–98% of U.S. households (Harter et al., 2016) and is updated frequently. It can also differentiate between business and residential addresses (Dillman et al., 2014). Thus, CDSF can provide a high-coverage sampling frame. A problem occurs with people with multiple addresses, as they are more likely to be selected from the list. If the panel offers only online survey options, people who do not have internet access/skills are not covered.

The process of selecting a respondent within a household may lead to the under- or over-representation of some members, leading to selection bias. Multiple persons often live in one household. If a sampling unit represents a household, every eligible member of each sampled household should have an equal opportunity to be selected as the respondent (Dillman et al., 2014). In reality, males are more likely to answer mailing household surveys because their names tend to be on the address list, while women are more likely to answer telephone surveys (Dillman et al., 2014).

Sampling Error. Sampling errors occur in any sampling survey. The issue lies in how much variability between the sample estimate and the true population value is

tolerated. Predicting the acceptable sampling error can assist in estimating the necessary sample size. If pollsters require a precise estimate for subgroups, they need to have a bigger sample size. Most surveys report sampling errors for the total population, but some do not report sampling errors for subgroups. Reporting sampling errors for subgroups is critical, because group differences in the public perception of higher education are central arguments made in many reports (e.g., conservative vs liberal). If researchers wish to draw conclusions based on minority groups (e.g., Asian, African American, and Hispanic), they must oversample them to reduce sampling errors for those groups.

Nonresponse Error. Nonresponse occurs at the unit and item level. Unit nonresponse refers to a situation where no information is acquired from a number of elements within the sample. In contrast, item nonresponse refers to a situation where only some questions are unanswered (Bethlehem, 2016). This paper discusses the unit-level nonresponse issues.

The response rate for RDD samples is significantly low. The response rate for landline was 8.9% and cell phones 3.2% in a poll by the Pew Research Center in 2019 (Pew, 2019b). In general, the response rate to telephone public polls has been falling dramatically in the U.S. For instance, the telephone response rates for surveys conducted by the Pew Research Center consistently dropped from 36% in 1997 to 6% in 2018 (Kennedy & Hartig, 2019). Such a decrease in response rate occurred partly because of blocking and flagging systems to circumvent the recent increase in automated telemarketing calls and other unwanted calls (Kennedy & Hartig, 2019; Tourangeau, 2004). People may refuse to participate in polls because of busyness, lack

of civic engagement, privacy concerns, and a lack of trust toward telemarketers (Tourangeau, 2004). Substantial levels of nonresponse threaten the representativeness of a sample. A nonresponse bias occurs if the response propensity is correlated with the variable of interest. The initial sample size must be large to obtain sufficient completed responses.

The response rate for probability-based panels is also low, and the nonresponse mechanism is complex. The specific response rate on polls on perceptions of higher education was not available for NORC and Ipsos. NORC reported that the final response rate was between 10–20% depending on the specific study parameter (e.g., target population, survey length, time in the field, and importance of the topic) (NORC, 2022). Here the word "final" was used because it considered the whole nonresponse process of panel recruitment, panel attrition, and survey participation. Once the survey frame was constructed, the pollsters tried to recruit the sampled members onto the panel. The initial panel recruitment rate was 6% for NORC's AmeriSpeak panel (NORC, 2022), which was quite low. NORC attempted follow-up recruitment with enhanced incentives and in-person visits, then recruited about 28% for 2021 sample cases. Thus, only about a third of sampled members were recruited to the panel in the panel recruitment years. Next, a constructed panel was used for various polls over the years. Panel attrition occurs as the panel survives. The annual panel retention rate for the AmeriSpeak panel was about 85% (NORC, 2022). As time passes after the panel was constructed, attrition reduces the number of available cases. Lastly, Some panelists refuse to join a specific survey, which causes further nonresponse. Because of this, the cumulative nonresponse rate accounting for nonresponse at each stage may become low

even if the probability-based panel starts from a relatively representative sampling frame (Tourangeau, 2004). This may introduce bias into the estimates.

Cost and Time. These errors should be considered under budget and time constraints. Implementing an online opt-in poll tends to be relatively inexpensive and quick. Calling by telephone is laborious as only about 25–30% of numbers are actually working residential numbers (Dillman et al., 2014). The necessity for follow-up calls increases the survey cost. However, it takes a relatively short amount of time to finish RDD surveys. Constructing quality probability-based panels requires extensive resources, time, and expert knowledge, but once the panel is built, it does not require extra costs and time to implement each specific survey. It is worth noting that costs and time depend on factors such as the survey mode, follow-up frequency, and incentive amount.

Table 4: Methodological strengths and weaknesses of three polls under the total survey error framework

	RDD	Probability-based panels	Online opt-in polls
1. Construct validity	−**Different constructs across question items (sentiment in higher education varies from 25% to 67%)** −**Different constructs for one item across individuals** −**Higher education is not a single entity (e.g., 4-year universities vs. community college, public vs. private)** −**Unclear types of benefits of higher education (economic vs. non-economic)**		
2. Measurement error	+Interviewer assistance −Interviewer errors	+Interviewer assistance (for telephone interviews) +Conditional effects −Careless answers −Interviewer errors (for telephone interviews) −Incentive errors	+No interviewer errors +Experience (for online panel) −Careless answers −Incentive errors
3. Processing error	−Weighting error (dual frame)	-Weighting error	−Weighting error
4. Coverage error	−Overlap of landline and cell phone frames −Moving without changing cell phone number −Selection within a household (for landline)	+High coverage −Internet accessibility and skills −Multiple addresses −Selection within a household (multiperson dwellings)	**Not applicable**
5. Sampling error	−Imprecise estimates for subgroups	−Imprecise estimate for subgroups	**Not applicable**
6. nonresponse error	−**Extremely low response** (Pew: landline 8.9%. Cell phone: 3.2%).	−**Low response** (NORC: 10–20%) −**Low recruitment** (NORC: initial 6%) −**Panel attrition** (NORC: annual retention rate 85%)	**Not applicable**
Cost	High for calling people	Extremely high for making a panel, but moderate for implementing each survey.	Low
Time	Quick	Slow to make a panel by ABS Quick to collect responses	Quick

Note: + and − show the potential strengths to reduce or weaknesses to increase survey errors, respectively. Random digit dialing (RDD). National Opinion Research Center at the University of Chicago (NORC). Address-based sampling (ABS).

4. Best Practices in Poll Design and Implementation for Tackling Three Concerns

The previous section highlighted methodological strengths and weaknesses based on the TSE framework. Each weakness in measurement and representation inference may increase survey errors. The literature discusses various techniques for checking and minimizing each error. This paper highlights three methodological concerns that are critical or unique to polls on perceptions of higher education. As highlighted in red in Table 4, three major concerns were identified: 1. construct validity of measured items, 2. nonresponse in RDD and probability-based panels, and 3. reliability of statistical adjustment used in online opt-in polls. The following section discusses best practices for these three methodological concerns to check and minimize errors.

4.1. Construct Validity

Pollsters should specify constructs of attitudes toward higher education before developing question items based on three categories of perceptions (confidence/sentiment of higher education, benefits of higher education, and specific issues regarding higher education), as this paper summarized. As higher education is not a single entity/level, poll creators should consider whether they are interested in public perceptions of higher education in general or specific entities or levels of higher education. The present paper

proposes that the benefits of higher education should be summarized in four categories outlined in Figure 1 (1. private economic, 2. private non-economic, 3. public economic, and 4. public non-economic benefits). Major polls reported that Americans questioned the benefit of higher education; thus, it is important for surveyors to clarify the benefits of higher education in their studies. As public perceptions of the non-economic benefits of higher education have rarely been investigated in the past, they should be examined in polls. New America added questions to measure the benefits of higher education holistically in 2023. This is good practice and other polls should also measure the comprehensive benefits of higher education.

Some may question the necessity of asking general confidence and sentiment questions if they potentially measure different constructs depending on the individual respondent. Public confidence/sentiment questions have been frequently used in polls. The intention of these questions is to evaluate overall sentiment toward an institution (Berinsky, 2017). With little knowledge of specific higher education issues, the public can answer general confidence and sentiment questions. It may be true that some people answer these questions with a specific frame of reference in mind, and that frame may differ across individual people and contexts (Tourangeau et al., 2000). If this is the case, the question item cannot measure the same construct across people and time. However, pollsters do not necessarily specify question items all the time. The public may not have opinions on specific higher education issues at all. It should be noted that the poll in question is not for those who are specifically concerned with education, such as teachers and parents, but for the general adult population. Some may not have an opinion on specific higher education issues, but their opinions might be forcefully induced once they participate in the poll (Converse, 1987). Such a situation should be avoided because the

poll can result in crafting a public opinion that does not exist in reality.

The recommended practice is to combine general confidence/sentiment questions and specific questions in a theoretically meaningful way (Berinsky, 2017). With general confidence/sentiment questions, pollsters can glean the public's overall feeling toward the higher education sector. Additional specific questions can provide the meaning behind the response for general confidence/sentiment. It is also critical for pollsters to explain the political context regarding higher education at the time of polls (e.g., the Supreme Court's decision on affirmative action) to understand how such contexts may affect responses to general confidence/sentiment questions. The response to institutional confidence questions is susceptible to short-term evaluations of political events and leaders (Cook & Gronke, 2005). Consequentially, providing general contexts around higher education at the time of polls is good practice to interpret general confidence/sentiment toward the higher education sector.

4.2. Nonresponse in RDD and Probability-based Panels

It is ideal to achieve a 100% response rate for polls. Some survey researchers have faith in large response rates and argue that high response rates are always better than low ones, while survey results are unreliable under a high rate of nonresponse (Tourangeau, 2004, p. 786). To increase the response rate, polltakers follow up on nonresponse cases at another time or date. Leaving a voice message before a follow-up call is also effective (Dillman et al., 2014).

However, a low response rate does not necessarily signal the existence of a nonresponse bias. A nonresponse bias exists when the survey variable is correlated with

the probability of being a respondent (Groves, 2006). Thus, nonresponse bias is relatively absent if there is no association between the variable of interest (e.g., confidence in higher education) and the response propensity (Tourangeau, 2004). The low response rate itself may not be a major concern if a nonresponse bias is unlikely to occur. Hereafter, the paper presents methods to assess potential nonresponse bias. The effectiveness of the assessment largely depends on the availability of quality measures. It should also be noted that each variable of interest (e.g., multiple measures on public perceptions of higher education) may be susceptible to different nonresponse biases within the same poll (Groves, 2006).

RDD. The nonresponse bias for RDD is partly assessed by two methods. First, the response rate is compared across subgroups. Polltakers can imply there is no evidence of nonresponse bias if demographic groups (e.g., age, race, gender, urbanicity) are not common causes of the variable of interest and the response probability (Groves, 2006). People may claim no evidence of nonresponse bias when the response rates are similar across subgroups (Groves, 2006). Poll administrators can adjust for the biasing effects of nonresponse propensities through weighting if some groups have low response rates (Groves, 2006). Keeter et al. (2017) found that the politically engaged are likely to respond in telephone polls. Therefore, a poll may collect a measure of political engagement and reduce their weight based on the benchmark target. This method cannot assess the nonresponse bias correctly if other unmeasured factors are common causes of the variable of interest and the response propensity. The second and most common method is comparing respondents' characteristics in a benchmark survey (e.g., the General Social Survey (GSS)). Although the variable of interest is unlikely to be measured in the benchmark survey, obtaining similar estimates between respondents and

the benchmark survey provides some level of trust in the poll (Groves, 2006). A study carried out by the Pew Research Center (Keeter et al., 2017) showed that an RDD telephone survey with large amounts of nonresponses provided a measure of demographics very similar to that of a government survey that had a high response rate. The drawback of this method is that the variable of interest (e.g., public perceptions of higher education) does not normally exist in the benchmark survey and the measurement form may differ even if the common variable is collected in the benchmark survey (Groves, 2006).

Probability-based Panels. The uniqueness of nonresponses of probability-based panels is that they are the sum of nonresponses in panel recruitment, panel attrition, and survey participation. Methods that assess nonresponse bias for RDD can be applied to probability-based panels. Once respondents agree to join the panel, demographic information tends to be available for attrition and non-participation cases. Characteristics of panelists before and after attrition are compared to assess the influence of attrition on nonresponse bias. Similarly, panelists' characteristics are compared before and after non-participation. Again, the effectiveness of weighting is contingent on the available measured characteristics.

As highlighted above, a significant depreciation is likely to occur in the recruitment process. NORC conducted follow-up recruitment to increase their response rate. Here, the nonresponse bias can be indirectly assessed by comparing the characteristics of respondents in the initial recruitment and those in the follow-up. This method was used by NORC's AmeriSpeak panel. NORC (2022) found that the characteristics of panelists in the follow-up recruitment differ from those in the first recruitment, such as political orientation and access to media. NORC asserted that the

combined panel achieved a greater balance, making estimates more accurate. This comparison does not offer direct information about nonrespondents to a poll; however, it assumes that nonrespondents are similar to those respondents recruited in the extensive follow-up group (Groves, 2006).

Follow-up recruitment tends to see an increase in incentives offered in the recruitment of people to a survey, and NORC increased the incentives offered in their follow-up recruitment. Previous literature has discussed strategies for increasing recruitment. Sending advance letters with prepaid incentives to postal addresses is more effective than postpaid incentives in increasing recruitment for mail surveys (Dillman et al., 2014). However, increasing the response rate with cash incentives does not necessarily reduce nonresponse bias. By reviewing the past literature, Singer and Ye (2013) concluded that incentives may reduce nonresponse bias if those who would otherwise fail to respond are recruited. However, they said it will increase the nonresponse bias if incentives attract sample members who are already overrepresented. Thus, the above-mentioned comparison of characteristics of panel members in initial and follow-up recruitment is a sound approach to assessing how additional incentives can attract the sample groups who would otherwise end up refusing to participate in the poll. Although this approach can assess the impact of additional incentives on the nonresponse bias, it does not assess the impact of the incentive itself because the original recruitment tends to contain a small incentive by default.

4.3. Reliability of Statistical Adjustment in Online Opt-in Polls

The online opt-in poll is an emerging method, as conducting probability sampling

with high response rates is becoming increasingly challenging. Although research is inconclusive (Bethlehem et al., 2016), the existing evidence does not fully support the use of online non-probability-based polls with a simple weighting adjustment. Past studies show that online opt-in polls did not ensure the representative estimates of various measures (e.g., political participation, health, and alcohol consumption) even after post-survey weighting for demographics (e.g., poststratification), and estimates were less accurate than the RDD and probability-based internet panel methods (Chang & Krosnick, 2009; Yeager et al., 2011). There is a major concern with regard to estimates for minority groups. A study on nine online nonprobability samples by the Pew Research Center showed that the average estimates of several measures for Hispanics, blacks, and young adults by online opt-in polls by raking are more biased when compared with gold-standard government surveys (Kennedy et al., 2016). As many polls are concerned with the group difference in public attitudes toward higher education, it is not advisable for polltakers to simply adjust the estimate of the online opt-in polls using weighting.

Sample matching may be a potential method for inferring the population estimate from an online opt-in panel. Sample matching entails drawing a subsample from a benchmark probability-based survey, such as the Current Population Survey, and then selecting individuals in the online opt-in panel who closely resemble each one in the subsample of a probability-based survey (Kennedy et al., 2016). Sample matching uses covariates in sample selection, while weighting uses them after the survey is completed (Baker et al., 2013). In theory, selection bias is reduced when the characteristics used for matching are key variables linked to the outcomes, thus achieving balance in the covariates (p. 35). The essential point is to balance covariates for sample matching, which is automatically achieved by random sampling in theory. Similar to the issue for

weighting adjustment discussed previously, the critical covariates may not be identified or collected for the non-probability sampling (Baker et al., 2013). In such a case, the inference by sample matching is not valid.

Recent findings revealed that relatively accurate estimates were achieved by sample matching. Bethlehem (2016) found that sample matching provided a similar estimate for the percentage of voters in the population. Ansolabehere and Schaffner (2014) revealed that online opt-in polls with sample matching produced point estimates for political measures on validated benchmarks that were as accurate as an RDD survey. Ansolabehere and Schaffner (2014) also used propensity weighting to account for the remaining unrepresentativeness of the matched sample. A recent study by the Pew Research Center also revealed that combining sample matching and weighting methods like raking reduced the bias on estimates from 24 benchmark questions obtained from high-quality federal surveys (Mercer et al., 2018). However, it is debatable whether the reduction in bias was substantial. Mercer et al. (2018) reported that even the most successful adjustment strategy, a combination of sample matching and propensity weighting for 8,000 samples, did not reduce the average estimated bias below six percentage points from 8.4 percentage points unweighted. In addition, past literature has largely evaluated the effectiveness of sample matching in reducing the bias of the estimate in the political context; however, its effectiveness within a non-political context (e.g., opinions for higher education) is still unknown.

There are many factors to consider when attempting to achieve high-quality sample matching. Ansolabehere and Schaffner (2014) pointed out four essential points for high-quality opt-in polls: 1. the use of sample matching based on a target probability-based sample, 2. the availability of variables in both online opt-in panels and the target

benchmark survey, 3. a large and diverse sample to allow close matching, and 4. the use of propensity score weights to account for the remaining unrepresentativeness of the matched sample. Selecting the right variables for statistical adjustment is especially critical for achieving accurate estimates. Mercer et al. (2018) proposed that these correct variables include more than core demographic variables (e.g., political variables). The values of these variables should be measured in the same way, and the mode of data collection should be similar to eliminate mode effects (Bethlehem, 2016). Based on theories and previous research, pollsters should choose the right variables that are linked to the variable of interest. It should be noted that online opt-in panels also suffer from refusal or non-participation in the response state. Matched samples in an online panel may not necessarily accept the invitation to participate in the poll. Pollsters also have to consider the trade-off between bias reduction and variability increase (e.g., discarding observations in performing matching) (Mercer et al., 2018). While online opt-in polls are often considered cost-effective and quick, high-quality ones require sufficient time, resources, and expert knowledge in planning, implementation, and post-survey processing. Attaining these presents a significant challenge in practice.

Lastly, low-quality online opt-in polls are useful for conducting exploratory studies and experiments. The validity of question items for some constructs can be tested in online opt-in polls. While understanding public perceptions of higher education is vital, it is worth noting that these perceptions may not always be accurate. The study may provide information about higher education and investigate how their perceptions change. For instance, recent polls reported that the public doubts college credentials' value for money. Information intervention may test if their perceptions are corrected once they understand the actual rates of return to a college education.

5. Recommendations for Survey Users When Interpreting/Using Public Polls

Topline Message: Public polls should be continuously used to gauge the public perceptions of higher education despite limitations. Despite the methodological concerns discussed above, public polls are unique survey methods for collecting public opinions that are representative of U.S. adults. Surveys of college students and parents, and proxy information for public perceptions of colleges, such as college enrollment rate, cannot capture these representative opinions. Despite potential nonresponse bias, sample estimates from the probability-based sampling frame can infer the representative public opinions in theory. Further, higher education is considered a public good and largely supported by public tax, so it is critical for survey users to understand the general public's opinions on higher education. Therefore, they should still rely on public polls to measure public perceptions of higher education.

With this topline message, here are recommendations for survey users to interpret and use public polls.

Clarify the goals and purposes of your study. Clearly identify the goals and purposes of your study before starting to use polls. If the goal is not to make inferences of estimates from the sample to a target population, users may not need to concern themselves with errors in the representation process and potential risks concerning online

opt-in panels. Specifying constructs of public attitudes helps determine which survey questionnaires to use. Polling methods can be selected based on how prompt the result should be released (e.g., RDD). Clarifying your research objectives is beneficial in determining what polling methods should be employed.

Check if question items correctly capture the construct you are interested in. Once you identify constructs you would like to infer, you must analyze whether question items can correctly measure the intended constructs. Consider how item wording and polling settings affect the different interpretations of respondents during the answering process. Use the evidence cautiously if you suspect items do not measure the construct. No psychometric research has been conducted on public perceptions of higher education. Survey users or researchers should conduct such studies to understand the construct validity of public attitudes toward higher education.

Analyze political contexts that may affect the answer in interpreting the result. Respondents are not free from political contexts. Public confidence and sentiment toward higher education are susceptible to the short-term evaluation of political events and leaders; thus, understanding current political contexts regarding higher education is key to interpreting the polling result. Such analysis is helpful for comparing results across time and by a group.

Understand where survey errors are likely to occur for your polling method. None of the polling methods are infallible. Each method has its strengths and weaknesses. Holistically map out potential survey errors in the polling by using the TSE framework

as this paper did in Table 4. Investigate how errors were addressed in the pollsters' reports.

A high level of nonresponse does not necessarily lead to a nonresponse bias. A large nonresponse rate does not necessarily signify a large nonresponse bias. A nonresponse bias exists when the survey variable is correlated with the probability of being a respondent. Thus, nonresponse bias is relatively absent if there is no association between the variable of interest and the response propensity. A low response rate may not be of major concern if you can show that the key variable is unlikely to be related to the response propensity. Likewise, efforts to increase response rates (e.g., cash incentives) may not reduce the nonresponse bias unless sample members who would otherwise refuse to respond are recruited. They may increase the existing nonresponse bias if incentives recruit sample members who are already overrepresented.

Telephone surveys are not "dead" despite high levels of nonresponse. Some pollsters refrain from using telephone surveys due to a high nonresponse rate, but telephone surveys are not completely redundant. Although a large number of nonresponses signal the risk of nonresponse bias, nonresponse itself does not shape the nonresponse bias. The nonresponse bias occurs if the variable of interest is correlated with the response probability. Making a direct comparison between respondents and nonrespondents is infeasible, but the representativeness of respondents can be checked by comparing their characteristics with ones of a benchmark survey with a high response rate (e.g., GSS). Weighting can also be used to adjust for sample demographics based on knowledge of who is likely to respond in telephone surveys. If such analysis can indirectly justify the representativeness of completed cases, telephone surveys will still be used.

Probability-based panels are promising but be wary of the nonresponse mechanism. Some pollsters shifted from RDD telephone surveys to probability-based panels. The flexibility of survey modes (e.g., online and telephone) and intense follow-ups may increase response rates and achieve a more balanced panel. The higher coverage of the sampling frame by ABS in probability-based panels is beneficial for reducing coverage error. However, you cannot simply argue that probability-based panels are better than RDD. It should be noted that nonresponse processes are complicated for probability-based panels. Depreciation from a sample occurs in each state of recruitment, panel attrition, and participation in a specific poll. The nonresponse bias is still a significant concern for probability-based panels if the responded cases do not represent the population.

Online opt-in polls are not advisable for inferring population estimates. Although online opt-in polls are promising polling methods, their use is not recommended when the purpose of polls is to infer the population estimate of perceptions of higher education. Implementing online opt-in polls appears inexpensive; however, it requires time, budget, and expert knowledge to infer the representativeness of a sample estimate for high-quality online opt-in polls, such as with sample matching. If there is no alternative to non-probability polls, cautiously use high-quality online opt-in polls with an explicit discussion of the inferential limitation.

Online opt-in polls can be used for exploratory studies and experiments. Quick and inexpensive online opt-in polls can be used for testing the construct validity of measures and information interventions to influence public perceptions. As long as the

purpose of using online opt-in polls is not the inference of population estimates, they can be used for a variety of purposes.

References

Ansolabehere, S., & Schaffner, B. F. (2014). Does Survey Mode Still Matter? Findings from a 2010 Multi-Mode Comparison. *Political Analysis*, *22*(3), 285–303. https://doi.org/10.1093/pan/mpt025

Association of Governing Boards of Universities and Colleges. (2018). Public Confidence in Higher Education. https://agb.org/wp-content/uploads/2019/11/report_2018_guardians_public_confidence_REVISED.pdf

Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., & Tourangeau, R. (2013). Summary Report of the AAPOR Task Force on Non-probability Sampling. *Journal of Survey Statistics and Methodology*, *1*(2), 90–143. https://doi.org/10.1093/jssam/smt008

Belkin, D. (2023, March 31). Americans Are Losing Faith in College Education, WSJ-NORC Poll Finds -Confidence in value of a degree plummeted among women and senior citizens during pandemic-. *Wall Street Journal*. https://www.wsj.com/articles/americans-are-losing-faith-in-college-education-wsj-norc-poll-finds-3a836ce1

Berinsky, A. J. (2017). *Measuring Public Opinion with Surveys* (SSRN Scholarly Paper 2968092). https://doi.org/10.1146/annurev-polisci-101513-113724

Bethlehem, J. (2016). Solving the Nonresponse Problem With Sample Matching? *Social Science Computer Review*, *34*(1), 59–77. https://doi.org/10.1177/0894439315573926

Biemer, P. P. (2010). Total Survey Error: Design, Implementation, and Evaluation. *The Public Opinion Quarterly*, *74*(5), 817–848.

Blake, J. (2023, July 11). American confidence in higher ed hists historic low. *Inside Higher Ed*. https://www.insidehighered.com/news/business/financial-health/2023/07/11/american- confidence-higher-ed-hits-historic-low

Brenan, M. (2023, July 11). Americans' Confidence in Higher Education Down Sharply. *Gallup*. https://news.gallup.com/poll/508352/americans-confidence-higher-education-down-sharply.aspx

Chang, L., & Krosnick, J. A. (2009). National Surveys Via Rdd Telephone Interviewing Versus the Internet: Comparing Sample Representativeness and Response Quality. *Public Opinion Quarterly*, *73*(4), 641–678. https://doi.org/10.1093/poq/nfp075

Converse, P. E. (1987). Changing Conceptions of Public Opinion in the Political Process. *Public Opinion Quarterly*, *51*(4 PART 2), S12–S24. https://doi.org/10.1093/poq/51.4_PART_2.S12

Cook, T. E., & Gronke, P. (2005). The Skeptical American: Revisiting the Meanings of Trust in Government and Confidence in Institutions. *The Journal of Politics*, *67*(3), 784–803. https://doi.org/10.1111/j.1468-2508.2005.00339.x

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method, 3rd ed* (pp. xii, 499). John Wiley & Sons Inc.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed mode surveys: The tailored design method, 4th ed* (pp. xvii, 509). John Wiley & Sons Inc.

Drezner, N. D. and Pizmony-Levy, O. (2023). *American higher education widely viewed as a worthwhile investment benefiting individuals and society.* Teachers College, Columbia University. 1-18. https://academiccommons.columbia.edu/doi/10.7916/ttph-b559

Drezner, N. D., Pizmony-Levy, O., & Pallas, A. (2018). Americans views of higher education as a public and private good. Teachers College, Columbia University. 1-16. https://www.tc.columbia.edu/thepublicmatters/reports/Research-Brief-2-v10102018.pdf

Finley, A., Aborn, M., Ruddy, S., & Miller, K. (2021). *Is College Worth the Time and Money?* Bipartisan Policy Center. Association of American Colleges and Universities (AACU). 1-10. https://bipartisanpolicy.org/download/?file=/wp-content/uploads/2021/09/BPC_Fed-_State_Brief_R05.pdf

Fishman, R., Nguyen, S., & Woodhouse, L. (2022). *Varying Degrees 2022 -New America's Sixth Annual Survey on Higher Education-*. New America. https://www.newamerica.org/education-policy/reports/varying-degrees-2022/executive-summary/

Fishman, R., Nguyen, S., & Ezeugo, E. (2018). *Varying Degrees 2018 -New America's Annual Survey on Higher Education-*. New America. https://d1y8sb8igg2f8e.cloudfront.net/documents/Varying_Degrees_2018_Final_PDF. pdf

Fishman, R., Ekowo, M., & Ezeugo, E. (2017). *Varying Degrees -New America's Annual Survey on Higher Education-*. New America. https://na-production.s3.amazonaws.com/documecants/Varying-Degrees.pdf

Gallup (2023). Gallup News Service June Wave 1 -Final Topline-. Gallup. https://news.gallup.com/file/poll/508358/230710ConfidenceHigherEd.pdf

Gallup (n.d.). *Confidence in Institutions.* Gallup. https://news.gallup.com/poll/1597/confidence-institutions.aspx

Great Blue Hill (GBH). (2018a). *WGBH News Higher Education Poll: Top Line Data*. GBH. September 2018. https://www.wgbh.org/news/education-news/2018-09-16/wgbh-news-higher-education-poll-top-line-data

Great Blue Hill (GBH). (2018a). *Methodology: National Higher Education Poll*. GBH. September 2018. https://www.wgbh.org/news/education-news/2018-09-16/methodology-national-higher-education-poll

Groves, R. M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *The Public Opinion Quarterly*, *70*(5), 646–675.

Groves, R. M., & Lyberg, L. (2010). Total Survey Error: Past, Present, and Future. *Public Opinion Quarterly*, *74*(5), 849–879. https://doi.org/10.1093/poq/nfq065

Groves, R. M., Fowler Jr., F. J., Couper, M. P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2004) *Survey Methodology*. Wiley Series in Survey Methodology. 1st edition Paperback. John Wiley & Sons Inc.

Harter, R., Battaglia, M. P., Buskirk, T. D., Dillman, D. A., English, N., Fahimi, M., Frankel, M. R., Kennel, T., McMichael., J. P., McPhee, C. B., Montaquila, J., Yancey, T. & Zukerberg., A. L. (2016). *Address-based sampling.* American Association for Public Opinion Research. https://www.aapor.org/wp-content/uploads/2022/11/AAPOR_Report_1_7_16_CLEAN-COPY-FINAL-2.pdf

Ipsos (2019). Public Agenda/USA Today/Ipsos Hidden Common Ground Survey. Ipsos. December 2019. https://www.ipsos.com/en-us/news-polls/public-agenda-hidden-common-ground

Ipsos (n.d.). *KNOWLEDGEPANEL® A Methodological Overview*. Ipsos. https://www.ipsos.com/sites/default/files/ipsosknowledgepanelmethodology.pdf

Keeter, S., Hatley, N., Kennedy, C., & Lau, A. (2017). *What Low Response Rates Mean for Telephone Surveys*. Pew Research Center, 1-39. https://assets.pewresearch.org/wp-content/uploads/sites/12/2017/05/12154630/RDD-Non-response-Full-Report.pdf

Kennedy, C., Popky, D., & Keeter, S. (2023). *How Public Polling Has Changed in the 21st Century.* April 2023. https://www.pewresearch.org/methods/2023/04/19/how-public-polling-has-changed-in-the-21st-century/

Kennedy, C. & Hartig, H. (2019). *Phone survey response rates have resumed their decline.* Pew Research Center. February 2019. https://www.pewresearch.org/short-reads/2019/02/27/response-rates-in-telephone-surveys-have-resumed-their-decline/

Kennedy, C., Mercer, A., Keeter, S., Hatley, N., McGeeney, K., & Gimenez, A. (2016). *Evaluating online nonprobability surveys.* Pew Research Center, 1-60. https://assets.pewresearch.org/wp-content/uploads/sites/12/2016/04/Nonprobability-report-May-2016-FINAL.pdf

Leveille, D. E. (2006). *Accountability in Higher Education:A Public Agenda for Trust and Cultural Change.* Center for Studies in Higher Education. https://cshe.berkeley.edu/publications/accountability-higher-educationa-public-agenda-trust-and-cultural-change

Loveless, T. (1997). The Structure of Public Confidence in Education. *American Journal of Education*, *105*(2), 127–159.

McMahon, W. W. (2009). *Higher Learning, Greater Good: The Private and Social Benefits of Higher Education*. Johns Hopkins University Press.

Mercer, A., Lau, A., & Kennedy, C. (2018). *For Weighting Online Opt-In Samples, What Matters Most?* Pew Research Center, 1-69. https://www.pewresearch.org/methods/2018/01/26/for-weighting-online-opt-in-samples-what-matters-most/

National Opinion Research Center at the University of Chicago (NORC) (2022). *Technical Overview of the AMERISPEAK® Panel NORC's Probability-Based Household Panel*. NORC. https://amerispeak.norc.org/content/dam/amerispeak/research/pdf/AmeriSpeak%20Technical%20Overview%202019%2002%2018.pdf

Nguyen, S., Fishman, R., & Cheche, O. (2023). *Varying Degrees 2023 -New America's Seventh Survey on Higher Education-*. New America. https://www.newamerica.org/education-policy/reports/varying-degrees-2023/

Pew Research Center (2021). *Mobile Fact Sheet.* Per Research Center. April 2021. https://www.pewresearch.org/internet/fact-sheet/mobile/

Pew Research Center (2019a). The Growing Partisan Divide in Views of Higher Education. Pew Research Center. August 2019. https://www.pewresearch.org/social-trends/2019/08/19/the-growing-partisan-divide-in-views-of-higher-education-2/

Pew Research Center (2019b). *July 2019 Political Survey.* Pew Research Center. July 2019. https://www.pewresearch.org/politics/dataset/july-2019-political-survey/

Pusser, B., & Doane, D. J. (2001). Public Purpose and Private Enterprise The Contemporary Organization of Postsecondary Education. *Change: The Magazine of Higher Learning*, *33*(5), 18–22. https://doi.org/10.1080/00091380109601815

Schleifer, D., Friedman, W., & McNally, E. (2022). *America's Hidden Common Ground on Public Higher Education: What's Wrong and How to Fix It.* Public Agenda. 1-32. https://publicagenda.org/wp-content/uploads/Public-Agenda-HCG-Higher-Ed-Report-FINAL.pdf

Schleifer, D. & Silliman, R. (2016). *What's the Payoff? Americans Consider Problems and Promises of Higher Education.* Public Agenda. 1-11. https://files.eric.ed.gov/fulltext/ED591331.pdf

Singer, E., & Ye, C. (2013). The Use and Effects of Incentives in Surveys. *The Annals of the American Academy of Political and Social Science*, *645*, 112–141.

Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology*, *55*, 775–801. https://doi.org/10.1146/annurev.psych.55.090902.142040

Tourangeau, R., Rips, L. J., & Rasinski, K. (Eds.). (2000). *The psychology of survey response* (pp. xiii, 401). Cambridge University Press. https://doi.org/10.1017/CBO9780511819322

Van Dam, A. (2023, June 23). Barely a quarter of Americans still have landlines. Who are they? *The Washington Post.* https://www.washingtonpost.com/business/2023/06/23/landline-telephone-holdouts/

Wall Street Journal / National Opinion Research Center at the University of Chicago (WSJ/NORC) (2023). *WSJ/NORC Poll Match 2023.* WSJ. https://s.wsj.net/public/resources/documents/WSJ_NORC_ToplineMarc_2023.pdf

Weisberg, H. (2018). Total Survey Error. In L.R. Atkeson & R. M. Alvarez, (Eds.), *The Oxford Handbook of Polling and Survey Methods* (pp. 13-27). Oxford University Press.

Yeager, D. S., Krosnick, J. A., Chang, L., Javitz, H. S., Levendusky, M. S., Simpser, A., & Wang, R. (2011). Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples. *The Public Opinion Quarterly*, *75*(4), 709–747.