LION POLL

CENTER FOR SURVEY RESEARCH AT PENN STATE HARRISBURG

Lion Poll Survey:

Benefits, Challenges, and Study Methods

May 2019

Web panel methodologies are extremely cost-effective and, for smaller project budgets, allow for robust data collection that is much more affordable than phone and mail surveys. They can also produce results more quickly—several weeks compared to a few months for mail or phone methodologies. Nevertheless, there are concerns that are often raised when discussing this method: 1) Are panelists representative of the reference population? 2) Do respondents need to have Internet access at home to participate? 3) Will a web panel survey result in low-quality responses? 4) Will the survey yield a lower response rate? The following discusses these concerns in detail and provides more information about how the web panel methodology compares to other survey methodologies.

To be clear, web panels are a non-probability based data collection method. Respondents will not be selected from the general population at random; rather, only respondents who opted to participate in a paid web survey panel will be included in the sampling frame. The results are representative only of those who choose to participate. As a result, certain biases might exist that would prevent a direct comparison to Pennsylvania's general population. Namely, results may be biased against those who are less likely to participate in web survey panels (sample frame bias) or those panelists who choose not to participate in the study (non-response bias). In order to ensure that the results of the survey are not biased towards any particular location, age, or sex, the Center for Survey Research programs quotas into web panel surveys to guarantee that the final dataset will be representative of Pennsylvania's known population by region, and separately, by age/sex combined categories.

It is worth noting that the issue of underrepresentation is not unique to non-probability methods. Unlike a phone or mail survey methodology where weighting is often utilized, web panels allow for more direct targeting of the population of interest, which eliminates the need for weighting on specific variables of interest. Producing a representative dataset without weighting reduces overall survey error.

Next, although respondents who do not have access to the Internet at home are less likely to sign up for web survey panels, they are not excluded from participating, and may do so through the use of public Internet access, mobile phones, or Internet at their place of employment. It should also be noted that access to broadband

Internet is not necessary to participate as a web survey panelist. For the Lion Poll, respondents will not be viewing video, pictures, or soundbites as part of the web survey, which means that respondents could access and participate in the web survey via a dial-up Internet connection.

Furthermore, the research team will utilize a variety of strategies to ensure that online web panel datasets contain high-quality responses. First, the research team will use screening methods that obscure the types of respondents that we are not looking to survey. Since panelists are paid for their participation, this decreases the chance of the respondent fabricating answers to gain access to the study. For example, instead of confirming that they are a Pennsylvania resident, respondents would select their state from a list of states; only those from Pennsylvanian would continue. Next, checks will be built in to identify bots (short for "robot") and people who do not read the questions. Specifically, the research team will embed attention checks and straight-lining checks to confirm that the respondent was carefully reading and responding to questions. Straight-lining is when a respondent answers all questions in a series in the same way. To check for this, one item will be included in a question series in two different ways: once with positive wording and once with negative wording. Respondents who carefully read the questions will answer in contrasting ways. If respondents do not meet the eligibility requirements or appropriately answer quality check questions, they will be removed from the survey and will not be counted towards the required completed interviews. Additionally, a manual review will take place where outliers (i.e., surveys completed very quickly) and not-applicable open-ended responses (possibly from automated/bot responses) will be removed from the final dataset.

Finally, online panel surveys typically see a response rate of 1-3%, which is consistent with response rates obtained via phone survey and mail/web mixed mode surveys in the last few years. This is due, in large part, to the fact that survey respondents lead busy lives. In addition, the quota system results in participants being excluded from participating after their cohort has filled; this decreases the overall response rate because more panelists are invited than are ultimately able to participate. Finally, panelists may no longer be interested in completing surveys, despite having never officially ended their formal panel involvement. Because each survey

is different, the number of individuals who will be invited is not known in advance; this is similar to how sampling is done in phone survey methodologies.

Despite these disadvantages, the research team has found the results to be more representative of Pennsylvania's population than random selection methods. The Center for Survey Research (CSR) compared the demographic characteristics of its omnibus online web panel survey (Lion Poll) and its omnibus random phone survey (Penn State Poll) to known Census proportions for Pennsylvania. The web panel utilized quotas by age/sex combination and, separately, by region while the random phone survey was weighted by age/sex categories. The results suggest that the web panel sample more closely matches the population on demographic variables of interest than the so-called random phone survey. A direct comparison to income was not possible due to categories that did not align with Census data. The following table displays these findings:

Question	Response Choice	% of Respondents		% Difference from Census		Census %*
		Random phone	Web panel	Random phone	Web panel	
Gender	Male	48.4%	48.5%	-0.1%	0.0%	48.5%
	Female	51.6%	51.5%	0.1%	0.0%	51.5%
Age	18-34 years	28.3%	28.3%	0.0%	0.0%	28.3%
	35-64 years	49.7%	49.7%	0.0%	0.0%	49.7%
	65 and older	22.0%	22.0%	0.0%	0.0%	22.0%
Region	Northern	9.8%	7.4%	2.4%	0.0%	7.4%
	Central	9.9%	9.4%	0.5%	0.0%	9.4%
	Northeast	12.7%	16.3%	-3.6%	0.0%	16.3%
	Southeast	28.4%	31.3%	-2.9%	0.0%	31.3%
	South Central	19.8%	14.8%	5.0%	0.0%	14.8%
	Southwest	19.4%	20.8%	-1.4%	0.0%	20.8%
Density	Rural	30.3%	25.9%	3.8%	-0.6%	26.5%
	Urban	69.7%	74.1%	-3.8%	0.6%	73.5%
Education	High school diploma or less	18.5%	28.9%	-27.7%	-17.3%	46.2%
	Some college (includes two-year degree, technical degree, Associate's)	40.2%	32.0%	13.6%	5.4%	26.6%
	College degree (Four-year college graduate)	22.6%	26.3%	5.5%	9.2%	17.1%
	Graduate work	18.7%	12.8%	8.6%	2.7%	10.1%
*The data source used for comparison was the American Community Survey 5-year Estimate, 2012-2016.						

The research team will follow a rigorous data collection process, utilizing the quality control techniques noted above. First, the Center's research team will work closely with each client to design survey questions that respondents can understand; ensure that sections flow well; look for possible context effects; and consider respondent burden (question complexity, length, etc.). Sixteen standard demographic questions will also be included so that the researchers can complete robust data analysis; these questions examine: age, sex, ethnicity,

race, education, income, political affiliation, political ideology, geography (zip code, county, region, rural/urban status), number of children and adults in the household, household size, and marital status.

The Center for Survey Research utilizes Marketing Systems Group (MSG) to recruit respondents who had previously signed up to participate in a variety of web survey panels in exchange for nominal compensation paid by the panel provider. These web survey panels were constructed utilizing a double opt-in recruitment technique where the potential respondents provide their email address to indicate interest in participating in the panel. An automated email is then sent to the email address provided to ensure that the owner of the email is aware of the intended use. Once the email owner confirms their interest in participating, the individual becomes part of the panel. After providing information on a wide variety of demographic topics to establish a panelist profile, they are eligible to participate in future surveys. This system guarantees that all respondents have provided explicit consent to be contacted for marketing and/or research purposes in compliance with all applicable federal laws.

MSG will send emails to potential respondents that include a customized web survey link, programmed in the CSR's Qualtrics web survey account, with a respondent identifier. Upon successful completion of the survey, the respondent will be redirected back to the appropriate web panel provider. As a result of this setup, only members of the research team have access to the survey results and the research team will never have access to any personally identifiable information of any of the survey respondents (names, emails, etc.). Regular panel maintenance, completed in collaboration with MSG, will be performed throughout the fielding period to manage quotas and ensure that high-quality respondents are being utilized in the survey.

A "soft-launch" will be performed where a working draft of the survey instrument will be pre-tested with a small sample of respondents before full survey distribution begins. The pre-test process will ensure that the survey programming is functioning as intended. Pre-testing also increases the likelihood that the questions provide accurate data while decreasing the likelihood of collecting unusable data. Assuming that the pre-test is error-free, it will be incorporated into the final dataset. The survey will field until all completed surveys are received (1,000 for the Lion Poll and 500 for Lion Poll Lite).