



Institute of Governmental Studies
126 Philosophy Hall
University of California, Berkeley
Berkeley, CA 94720
Tel: 510-642-6835
Email: igs@berkeley.edu

Survey Methods of the Berkeley IGS Poll

The Berkeley IGS Poll is a regularly scheduled non-partisan survey of California public opinion conducted by the Institute of Governmental Studies (IGS) at the University of California, Berkeley. A component of the University of California system's flagship Berkeley campus, IGS is the oldest organized research unit in the UC system and the oldest public policy research center in the state. Each poll is conducted to assess public opinion throughout California on important public policy issues facing the state and the nation on behalf of the University.

Data collection

The poll is administered online by distributing email invitations to stratified random samples of the state's registered voters. Each invitation asks voters to participate in a non-partisan survey conducted by the University and provides a link to the IGS website where the questionnaire is housed. The invitation is distributed under University of California letterhead and is signed by the Institute's co-directors, Professor Eric Schickler and Associate Professor G. Cristina Mora. It describes the broad purposes of the poll, its sponsorship, how long the survey is likely to take, and the fact that all email addresses were obtained from publicly accessible information contained on the state's voter rolls.

A consent form is also appended, which respondents are required to complete before opening the link to the survey questionnaire. The consent form explains that responses to the survey will remain anonymous and that all personally identifiable information about them will be purged from the data file and replaced with a unique identification number during data processing. Respondents are also provided with the contact information of the IGS co-directors,

as well as that of the UC Berkeley Committee for the Protection of Human Subjects, should they have any questions or concerns about the legitimacy of the poll.

To maximize survey participation, up to three email invitations are distributed to voters at different intervals over an approximately one-week data collection period. An opt-out link is provided at the bottom of each invitation for voters not wishing to receive reminder emails from IGS about the survey. While participation is voluntary, to encourage participation, voters are also told that by responding to the survey they will be entered into a drawing to win a gift card worth \$200.

The low cost of sampling voters by an email invitation approach enables the poll to efficiently retrieve large samples of the California registered voter population, typically on the order of 6,000 - 8,000 or more. These large sample sizes enable the statewide findings to be subdivided and reliably examined across each of eight major geographic regions of the state, as well as across a wide range of political and demographic subgroups of the registered voter population. Samples of this size also lend greater opportunities to apply detailed weighting targets to align the sample to population characteristics of California registered voters at both the state and regional level.

Questionnaire development

Berkeley IGS Poll director Mark DiCamillo takes the lead in developing each survey questionnaire, in consultation with the IGS co-directors. In addition, when developing questions on state election contests and major public policy issues facing California, he consults with reporters and editors of the *Los Angeles Times*, which provides the poll with annual funding and support. In return the *Times* is granted the right of first release to all reports prepared for public release by the Berkeley IGS Poll at the conclusion of each survey.

Upon request, university faculty members, campus institutes, as well as non-profit centers, foundations, and government agencies can also pay for the right to add questions to the poll to further their own research on public policy topics of their own choosing. While each party is responsible for developing the questions to be included on the poll, their proposed questions are

also reviewed by the poll director to ensure that they are posed in a balanced and unbiased manner.

Once the survey questionnaire has been finalized, it is translated into Spanish, and following this, both the English and Spanish language versions are programmed onto the Qualtrics platform for online administration. The programs are thoroughly tested on both laptop and mobile devices to review their appearance on the screens, and to ensure that the survey's skip pattern logic and item randomizations are working properly.

Sampling procedures

Stratified random samples of California registered voters with email addresses are provided to IGS by Political Data, Inc. (PDI), a leading supplier of voter lists in California. The PDI listings are derived from voter information originating from official state and county voter registration records. PDI's listings are regularly updated to ensure that the samples they provide to IGS are current and reflect the state's ever-changing voter population. In addition, because the poll is conducted using a registration-based sampling methodology, all survey participants are known to be registered voters, which is an advantage in polling voters on election-related issues.

Because IGS has found that response rates to email invitation surveys vary in relation to a voter's age and gender, before drawing its samples the poll director provides PDI with instructions for stratifying its statewide sample listings into different age and gender-based segments, and to then select voters randomly within each strata. Proportionately greater numbers of voter listings are selected for lower responding voter segments, while proportionately fewer listings are drawn for the higher responding segments. This procedure enables the survey to retrieve a balance of respondents across all age and gender subgroups.

To ensure a proper representation of the state's non-English-speaking Latino voter population, IGS also draws random samples from the pool of California voters who have requested receiving their registration and voting materials sent to them in Spanish. These voters are then sent their email invitations in Spanish and given the opportunity to complete the survey questionnaire in either Spanish or English.

IGS also sometimes augments its statewide sample with additional listings targeting voters who live within a particular geographic area of the state or within a specific demographic or political voter subgroup, such as among Central Valley voters, Blacks, or registered Republicans. When sample augmentations of targeted subgroups like these are included in the poll, weights are devised to return the oversampled segments back to their actual shares of the state's overall registered voter population during data processing.

Data processing

Sample drawn from the state's voter rolls contain a wealth of important demographic information about each voter, including their age, gender, county of residence, and party registration, as well as their history of voting in past elections. This information is merged into each respondent's survey record during the data processing phase. The data transferred from the voter file is then checked against the testimony given by survey respondents to ensure that the information they are providing is accurate and consistent with the information of that voter on their voter registration record. Where obvious inconsistencies are found, these cases are deleted from the data file.

Voter file matching procedures like these enable the poll to avoid the vulnerabilities associated with the inclusion of fraudulent responses from professional polltakers, bots, click farms, or from sources other than the randomly selected voter. The method also ensures that only one response can be obtained from each voter listing.

To protect the anonymity of survey respondents, at the conclusion of the data processing phase, each voter's name, email address, and all other personally identifiable information about the voter derived from their voter rolls are purged from the data file and replaced with a unique and anonymous identification number during the survey's data processing phase.

Sample weighting

At the conclusion of data collection and processing, sample adjustment weights are applied to bring the survey sample into alignment with known demographic and geographic benchmarks of the state's registered voter population. This is done using raking procedures, also referred to as sample balancing or iterative proportional fitting. The raking process can be thought of as a multidimensional poststratification procedure because the weights are post-stratified to one set (or dimension) of control totals, and these adjusted weights are post-stratified to another dimension. This process is iterated until the control totals for all dimensions are simultaneously satisfied, resulting in the production of a single final weighting assignment for each respondent.

Weighting targets are derived from databases of the California registered voter population as reported to IGS by PDI from their voter listing files, as well as from other government sources, such as the California Secretary of State and the U.S. Census Bureau's Current Population Survey, Voting Supplement for California. The weighting targets are applied at the state level, and whenever possible, at the regional level as well. For a typical poll, the following variables are included among the sample weighting targets: age, gender, race/ethnicity, educational attainment, dominant language, geographic region of the state, urbanity of their residential area, party registration, and voting participation in past elections, as well as combinations of these variables. In designing the weights, these targets incorporate important interactions to help address problems of differential non-response.

Tabulations and estimating sampling error

At the conclusion of the data processing phase, detailed tabulations are then prepared reporting the weighted results to each survey question both statewide and across a wide range of geographic, political, and demographic subgroups of the sample. When polling on election contests, additional tabulations are prepared filtering the survey results to include only those who are considered most likely to vote in an upcoming election. IGS typically defines likely voters from respondent testimony about their stated interest in and/or intention to vote in the upcoming election during the survey, coupled with each respondent's history of voting in recent past elections as reported on their voting record.

The tabulations display both the unweighted and weighted sample sizes of each voter segment under investigation. The unweighted sample sizes are helpful in estimating the sampling error associated with the results reported for each voter subgroup since they are used to estimate the variance of survey results derived from random sampling of a population. However, because surveys employing sample stratification and weighting procedures like those used by the Berkeley IGS Poll deviate from a simple random sampling of a population, the sampling error estimates associated with findings from the poll must account for these “house effects” and typically result in somewhat larger sampling error range than those derived from simple random sampling formulae. Nevertheless, because of the Berkeley IGS Poll’s large sample sizes, estimates of the sampling error associated with results based on each poll’s overall statewide sample typically fall with the range of +/- 2 to 3 percentage points at the 95% confidence level.

Reporting of the survey release and data archiving

Shortly after the completion of each survey, the poll director takes the lead in analyzing the survey results and in preparing press release-style reports for public distribution by IGS. Each report is printed on Berkeley IGS Poll letterhead and posted onto the Berkeley IGS Poll website through the University’s open-access publication platform, eScholarship, and are publicly accessible at <https://igs.berkeley.edu/research/berkeley-igs-poll>.

Poll releases are subdivided by topic and typically include a summary of the main poll findings on that topic, accompanied by tables comparing the survey results across key population subgroups or to past measures on the same topic. A fact sheet is also appended providing details about how the poll was conducted, the dates of data collection, and the exact wording and order in which the questions reported were asked to conform with the guidelines of the American Association for Public Opinion Research’s Transparency Initiative.

After the conclusion of each study, the source data file and its supporting documentation are forwarded to UC Data, the University’s main depository of social science research data, where

they can be accessed by students, faculty, and others for scholarly review and secondary analysis.