



Institute of Governmental Studies
102 Philosophy Hall, #2370
University of California
Berkeley, CA 94720-2370
Tel: 510-642-1473
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 California voter opinions on important public policy issues facing the state on behalf of the University.

Data collection

The poll is administered online by distributing email and text message invitations to stratified random samples of registered voters across California. The invitations ask 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. It also describes the broad purposes of the poll, its sponsorship, how long the survey is likely to take, and the fact that each voter's contact information was obtained from publicly accessible information contained on the state's voter registration rolls. The invitation is signed by IGS co-directors, Professors Eric Schickler and G. Cristina Mora, who serve as the poll's principal investigators for the University.

A consent form is also appended, which respondents are required to complete before beginning the survey. The consent form gives voters additional information about the survey explaining that their responses 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 to ensure the anonymity of all

survey respondents. 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 poll.

An initial email invitation is distributed to the sample of voters via Qualtrics. A reminder email is then forwarded via Qualtrics to those who have not responded several days later. An opt-out link is provided at the bottom of each invitation for voters not wishing to receive further correspondences from IGS about the survey. A final reminder is then sent to voters not responding to either email via text message. 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 relatively low cost of sampling voters by means of email and text messages enables IGS to efficiently retrieve large samples of the participants in its polls, typically on the order of 5,000 - 6,000 or more. Samples of this size enable the survey findings to be subdivided and reliably examined across eight major geographic subregions, as well as across a wide range of political and demographic subgroups of the California registered voter population. The survey's large sample sizes also lend greater stability to the statistical weights that are applied to the sample to align it to characteristics of the state's registered voter population.

Sampling procedures

The poll employs a registration-based sampling approach, whereby voters are selected from listings originating from the state's voter registration rolls, which is a public record. Sample listings are provided to IGS by Political Data, Inc. (PDI), a leading supplier of voter lists to the California research community. PDI is also responsible for and serves as the distributor of the poll's text messages during data collection.

To ensure a proper representation of the state's non-English-speaking Latino voter population, IGS also draws random samples of voters from the pool of registered voters who have requested that their official voting materials be sent to them in Spanish by the county Registrars. Data collection for the poll is sometimes expanded to be

administered in Chinese, Vietnamese and Korean languages. When doing so, random samples of voters are also drawn from the poll of registered voters who requested that their voting materials be sent to them in these languages.

PDI's database of the California registered voter population is regularly updated at the county level to ensure that the samples provided to IGS are current and reflect the state's ever-changing voter population. Because the poll's sample listings originate from the state's voter rolls, all survey participants are known to be registered voters.

Because IGS has found that response rates to email and text message invitations can vary in relation to a voter's age, gender, or other demographic factors, before drawing its samples, IGS provides PDI with instructions for stratifying its statewide voter file by age and gender, and to then randomly select voters within each stratum. This is done so that proportionately greater numbers of registered voters can be drawn for lower responding voter segments, while proportionately fewer listings are drawn from the higher responding segments. This procedure balances the number of completed surveys obtained across age and gender segments.

IGS sometimes augments its samples with additional listings targeting voters who live within a particular geographic area of the state or within specific demographic subgroups. When sample augmentations of targeted subgroups are included, weights are developed to adjust the oversampled segments so that the final weighted sample aligns each segment to its actual share of the registered voter population.

Questionnaire development

Berkeley IGS Poll director Mark DiCamillo typically takes the lead in developing the survey questionnaire, working in consultation with the IGS co-directors, and with reporters and editors of the *Los Angeles Times*, which serves as the poll's media partner. University faculty members, other campus institutes, as well as non-profit centers and foundations are also given the option of adding questions to the polls to further their own research on public policy topics.

Once the survey questionnaire has been finalized, it is forwarded to the University's Committee for Protection of Human Subjects (CPHS), which is responsible for reviewing and approving human subjects research on the Berkeley campus. Once approval from CPHS has been received, the survey materials are translated and are then programmed by IGS staff onto the Qualtrics software platform for online administration.

Before launching the survey, the programs are thoroughly tested to ensure that the survey's logic and randomizations are working properly and to review their appearance to voters who may be accessing the survey on different mobile and computer devices.

Data processing

Because the state's registered voter rolls contain a wealth of publicly available information, such as voters' age, gender, county and political jurisdiction of residence, party registration, and history of voting in past California elections, this information is merged into each respondent's survey record during the data processing.

To ensure that the information provided by those responding to the survey are the voters invited to participate, responses to several key demographic questions included on the survey questionnaire are compared to the demographic data about each voter included on their voting record. Where inconsistencies are found, these cases are deleted from the survey data file. Duplicate responses to each survey invitation are also not permitted.

Weighting Procedures

At the conclusion of data processing, sample adjustment weights are applied to align the demographic, geographic and political characteristics of those responding to the survey with population characteristics of the state's registered voters both overall and by region. This is done using raking procedures, also referred to as sample balancing. Raking is an iterative process in which survey estimates are controlled to their marginal population distributions. 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 then post-stratified to another dimension. This process is iterated until the control totals for all dimensions are simultaneously satisfied. This results in the creation of a final single weighting factor for each respondent. As a final step, the weighting assignments are trimmed so that no single respondent is assigned an abnormally large weight.

Weighting targets of demographic, geographic and political characteristics of the state's registered voters are derived from population estimates reported to IGS by PDI from its database of the state's registered voter file, and when appropriate, from other sources, such as the California Secretary of State, local Registrars of Voters, or the U.S. Census Bureau's Current Population Survey. The variables typically included in the development of the weighting targets include age, gender, race/ethnicity, English language proficiency, tenure, educational attainment, party registration, voting participation in past elections, geographic region and urbanicity of their residence, or combinations of these variables. In designing the weights, targets incorporate important interactions to help address problems of differential non-response.

Tabulation of the results

Once the weighting factors have been created, detailed tabulations of the survey results are created reporting the results to each substantive survey question in total and across a wide range of geographic, political, and demographic subgroups of the registered voter populations. These tabulations also display the unweighted and weighted sample sizes of each subgroup to permit the calculation of estimates of the sampling error applicable to the results of each subgroup.

When polling on election contests, additional sets of tabulations are prepared filtering the survey results among voters who are considered most likely to vote in an upcoming election. IGS typically defines likely voters from respondent testimony as to their

stated interest or intention to vote in the upcoming election, together with their history of voting in recent past elections as indicated on their voting record.

Reporting

The survey results are then analyzed and summarized in press release-style reports and are forwarded to the *Los Angeles Times*, the poll's media partner for initial public release, and are then later posted onto the *Berkeley IGS Poll* website at <https://igs.berkeley.edu/research/berkeley-igs-poll>, for general public distribution via eScholarship, the University's open-access publication platform.

Each press release prepared by IGS includes a summary of the poll's main findings and is accompanied by tables examining the survey results across key population subgroups or compares them to past measures on the same topic, when applicable. A fact sheet is also appended, providing details about how the poll was conducted, the dates of data collection, and the wording and order in which the questions being reported were asked in the survey, to conform with the guidelines of the American Association for Public Opinion Research's Transparency Initiative.

Data archiving

After the results from each survey have been distributed, the source data file from each poll is archived onto the University's Dataverse platform, so it can be accessible to students, faculty, and others for scholarly review and secondary analysis.