

Institute of Governmental Studies

124-126 Moses Hall

University of California

Berkeley, CA 94720

Tel: 510-642-6835

Email: igs@berkeley.edu

**Berkeley IGS Poll**

**Survey of California Registered Voters**

June 2019

Codebook 2019-02

**Data file citation:**

June 2019 Berkeley IGS Poll [machine-readable data file] Institute of Governmental Studies, University of California, Berkeley, 2019, Study number 2019-02.

***Overview of the June 2019 Berkeley IGS Poll***

Universe: California registered voters

Data Collection Period: June 4-10, 2019

Method of Data Collection: Online through email invitations distributed to registered voters statewide.

Languages of administration: English and Spanish

Sampling Method: Registration-based sampling of voters with email addresses from the state voter rolls.

Number of Cases: 4,435 registered voters

### About the Institute of Governmental Studies

The Institute of Governmental Studies (IGS) is an interdisciplinary organized research unit that pursues a vigorous program of research, education, publication and public service. A component of the University of California (UC) system’s flagship Berkeley campus, it is the oldest organized research unit in the UC system and the oldest public policy research center in the state.

IGS conducts periodic surveys of public opinion in California on matters of politics and public policy through its *Berkeley IGS Poll*. The poll, which is disseminated widely, seeks to provide a broad measure of contemporary public opinion, and to generate data for subsequent scholarly analysis.

***Survey Procedures***

The findings in this report are based on a *Berkeley IGS Poll* completed by the Institute of Governmental Studies (IGS) at the University of California, Berkeley. The poll was administered online in English and Spanish June 4-10, 2019 among 4,435 registered voters statewide.

The survey was administered by distributing email invitations to stratified random samples of the state’s registered voters. Once thequestionnaire and email invitations had been finalized, they were translated into the Spanish and reviewed for cultural appropriateness. Each email invited voters to participate in a non-partisan survey conducted by IGS and provided a link to the IGS website where the survey was housed. Reminder emails were distributed to non-responding voters over a seven-day period. An opt out link was provided at the bottom of each invitation for voters not wishing to participate or not wanting to receive future emails from IGS about the survey.

Samples of registered voters with email addresses were provided to IGS by Political Data, Inc., a leading supplier of registered voter lists in California. The email addresses of voters were derived from information contained on the state’s official voter registration rolls. The overall sample of registered voters with email addresses was stratified in an attempt to obtain a proper balance of survey respondents across major segments of the registered voter population by age, gender and race/ethnicity.

To protect the anonymity of survey respondents, voters’ email addresses and all other personally identifiable information were purged from the data file and replaced with a unique and anonymous identification number during data processing. At the conclusion of the data processing phase, post-stratification weights were applied to align the sample to population characteristics of the state’s overall registered voter population. Likely voters were then identified based on each voter’s stated intention to vote in next year’s primary election and factoring in their history of voting in past elections.

The sampling error associated with the results from the survey are difficult to calculate precisely due to the effects of sample stratification and the post-stratification weighting. Nevertheless, it is likely that the results from the overall registered voter sample are subject to a sampling error of approximately +/- 2.5 percentage points at the 95% confidence level. Results based on subgroups of this population would be subject to larger margins of sampling error.

Press releases and poll reports were then prepared from the survey results and were publicly released under *Berkeley IGS Poll* letterhead. For a complete listing of stories issued by the *Berkeley IGS Poll* go to <https://igs.berkeley.edu/igs-poll/berkeley-igs-poll>.

***About registration-based sampling***

Lists of citizens who are registered to vote in California are compiled and regularly updated by local County Registrars and the California Secretary of State. The lists contain a wealth of information about each voter, including their name, address, county, city, zip code and other geographic identifiers, gender, date of birth (age), party registration, voting participation, as well as other political and demographic descriptors.

More than six million registered voters in California now include their email address on their voting record when registering or re-registering to vote to vote, and the number of voters choosing to do so is increasing rapidly year by year. When drawing samples of registered voters with email addresses, the poll first stratifies the lists of voters sent an email invitation in an attempt retrieve a proper balance of voters across major segments of the registered voter population, such as by age, gender, region, and party registration.

Use of the registered voter data file as a sampling frame for the *Berkeley IGS Poll* offers a number of advantages when polling registered voters in California compared to other alternative sample methods.

First, it provides a random sampling of voters, unlike many online surveys that use a non-random sample and then must adjust it to match a voter model.

Second, since the voter file is derived from official state and county voting records, all respondents included in its samples are by definition registered to vote and which congressional district they’re registered in. The list also contains accurate information about the voter’s party registration, age, gender, and in many cases, race and ethnicity. This is superior to alternative approaches that must rely on respondent testimony of registration status, which is often unreliably reported, is used to establish post-stratification weighting parameters to align the samples to parameters of the registered voter population in each district.

Third, the file also provides accurate information about each voter’s participation in past elections in California, which is useful in identifying which registered voters are most likely to vote in an upcoming election. In addition, since each voting record also contains the voter’s age, gender, party registration, and city and county of residence, this information.

Finally, this method allows the poll to easily query people in languages other than English and to show respondents visual displays, such as the text of the ballot language for an initiative. For this poll, respondents were allowed to take the survey in English or Spanish.

***Sampling issues related to email surveys of the state’s registered voters***

The biggest potential problem is that only about one-third of voters include an email address in their voter file, and that one-third isn’t randomly distributed.

Certain groups, Latino voters, for example, are less likely to have an email in the voter file. This is also a challenge faced by telephone surveys that use the voter file, since not everyone has a working number on file, but the problem is bigger with emails.

To deal with potential under-representation of some groups, the sample of registered voters selected to participate in the survey is stratified across all major segments of the registered voter population, such as by age, gender, race and ethnicity in an attempt to retrieve a balance of responses from voters.

That means distributing a larger share of invitations to voter segments who are under-represented, such as Latino voters who are Spanish speakers. Stratification adds some complications to the calculation of the sampling error applicable to the survey results, but results in a more representative sample.

Once the survey results come in, the poll takes an additional step to ensure that the sample accurately represents the population at large, by weighting the data to reflect known demographics of the state’s registered voter population.