CITIZEN DATA

VOTE BY MAIL & ABSENTEE PROJECTIONS

Background

Citizen Data is currently working to project expected vote by mail, also known as absentee voting, behavior this November, with the primary goal of supporting election administrators with strategic resource allocation in light of these unprecedented national circumstances. Citizen is developing statistical models to predict the likelihood that individuals in key states — including Ohio, Florida, Georgia, Michigan, Wisconsin, and more — will vote absentee for the general election this year.

States were chosen for this project because they met the following criteria:

- typically see a minority of votes cast by absentee ballot in federal elections, and thus are likely to experience a significant shift this year and urgently need to adapt processes for more voters by mail and dropoff; and/or
- have both rural areas and large metropolitan cities where it's difficult to accommodate a highly concentrated population so they experience crowded polling places with long lines; and
- have already had a primary election or are going to have it soon, so that 2020 election turnout data would be available

This memorandum outlines the preliminary findings for our models, as well as our next steps for additional weighting and research.

Data and Method: Overview

Our absentee models combine historical data with recent survey responses to understand how Americans have acted in the past and how they intend to in the future in light of COVID uncertainties. Citizen conducted large-sample (N=4,000) surveys for each state, and matched the results of those surveys to its in-house national voter file. Then, Citizen weighted the dataset against the historical data to predict absentee vote turnout in the 2020 general election.

Results

As of late July, Citizen's modeling projected overall turnout rates and likely absentee turnout for five states, listed below:

Florida

- 9,799,932 of 13,620,535 (71.9%) registered voters will vote in the November election
- 42.8% (4,192,603) of likely Florida voters intend to vote absentee, either by mail or by returning their ballot at a secure dropbox in the November election

Georgia

- 4,074,136 of 6,764,824 (60.2%) registered voters are projected to vote in the November election
- 65.7% (2,674,719) of likely voters in Georgia are likely to vote absentee, either by mail or by returning their ballot at a secure dropbox in the November election

Michigan

- 4,777,619 of 7,191,963 (66.4%) registered voters are likely to vote in the November election
- 43.0% (2,054,291) of likely Michigan voters are projected to vote absentee, either by mail or by returning their ballot at a secure dropbox in the November election

Ohio*

- 5,705,038 of 7,399,732 (77.1%) registered voters will vote in the November election
- 71.3% (4,071,755) of likely Ohio voters will vote by mail in the November election

Wisconsin

- 2,444,761 of 4,506,975 (54.2%) registered voters are likely to vote in the November election
- 34.3% (839,096) of likely voters in Wisconsin are projected to vote absentee, either by mail or by returning their ballot at a secure dropbox in the November election

For most states, absentee turnout proved to vary significantly by party:

Florida

- 65.8% of Florida Democrats;
- 21.4% of Republicans; and
- 39.9% of non-partisans intend to vote by mail or by returning their ballot at a secure dropbox.

Georgia

- 83% of Georgia Democrats;
- 50.4% of Republicans; and
- 61.6% of non-partisans intend to vote by mail or by returning their ballot at a secure dropbox.

Michigan

- 61.1% of Michigan Democrats;
- 24.3% of Republicans; and
- 17% of non-partisans intend to vote by mail or by returning their ballot at a secure dropbox.

Wisconsin

- 48.2% of Wisconsin Democrats;
- 15.2% of Republicans; and

* The data that underlies Citizen's projections for Ohio resulted from a survey fielded in late June, while all other state projections are based on data from late July, which may explain projected differences in overall and partisan sentiment.

- ➢ 30.9% of non-partisans intend to vote by mail or by returning their ballot at a secure dropbox.

Ohio*

- ➢ 76.5% of Ohio Democrats;
- ➢ 79.6% of Republicans; and
- ➢ 49.6% of non-partisans intend to vote by mail or by returning their ballot at a secure dropbox.

Trends and Insights

Across all states, overall turnout and absentee voting rates will be higher than those seen in the 2016 general election. However, not all states will see the same levels of increase. The greatest increases relative to the 2016 general election will be in Ohio (322%),* Michigan (86%) and Florida (80%). The lowest increases are projected to be in Wisconsin (12.5%) and Georgia (20.8%).

On average, Democrats are much more likely to vote absentee (67.9%) than Republicans (37.6%). The only state where this is not true is Ohio, where Republicans will vote at a higher rate (79.6%) than Democrats (76.5%). Ohio is also the only state where white voters' absentee voting rates will exceed the absentee rates of racial minorities. In all other states, Hispanic, Black, and Asian voters will see higher rates of absentee turnout than white voters.

Election administrators in all states should anticipate increases in absentee voting. In Ohio and Georgia, the vast majority of turnout is likely to be absentee. Florida and Michigan could see absentee turnout rates that represent nearly an equal share to in-person turnout. Although we currently project a smaller proportion of Wisconsin voters to vote absentee, absentee turnout in the presidential primary was high, and Citizen will continue to monitor Wisconsin --- and all other states --- to assess whether absentee voter turnout is likely to be higher than currently projected.

Future Model Updates

We project higher rates of overall turnout than the 2016 general election, and, in general, Citizen's projections are higher than previously seen rates of absentee voting in each state. Though we weighted the responses to account for reasonable levels of response bias, it's important to note that this prediction was reached by surveying voters, who sometimes overestimate their voting intentions.

While Citizen is confident in its projections given the data inputs, the situation is fluid and dynamic, and future updates and improvements to the model will be necessary. This is an unprecedented and anxious time for Americans, and many factors that will affect how Americans turn out to vote absentee, including but not limited to:

- the varied language people use to reference understand vote by mail;
- unpredictable public-health related uncertainty; and
- future projections of unprecedented unemployment rates nationally.

* The data that underlies Citizen's projections for Ohio resulted from a survey fielded in late June, while all other state projections are based on data from late July, which may explain projected differences in overall and partisan sentiment.

While these factors were incorporated into these initial models Citizen will continue to update its projections for Florida, Georgia, Michigan, Ohio, and Wisconsin as these factors evolve with even greater emphasis. Further, Citizen will continue to improve its method as it receives additional input from experts in the field.

Conclusion

Due to the health and economic impacts of the COVID-19 pandemic, election administrators should anticipate notably higher degrees of voting absentee by mail and dropoff in 2020 than in previous elections. Still, our country's current situation is unprecedented and evolving, meaning that Citizen's model projections will need to be updated dynamically in coming weeks.

Technical Appendix

Survey Data Collection

Citizen conducted large-sample (N=4,000) surveys for each states, beginning with Ohio active voters between 6/29/2020 to 7/02/2020, and followed by active voters in Florida, Georgia, Michigan, and Wisconsin between 7/21/20 to 7/26/20. Voters in the samples were required to have voted in at least one election since and including the 2016 general election or to have been newly registered. The survey respondents were selected to closely match the age and other demographic distribution of their respective state electorate, and were sampled evenly across Congressional Districts. After the survey was completed, Citizen matched each respondent to its dynamic in-house national voter file.

Likely Voter Modeling

Before modeling the likelihood that an individual is likely to vote by mail or dropoff, we first had to predict whether the individual was likely to vote at all. To do this, Citizen generated a dataset that reflected what the voter file would have looked like in 2016 prior to the general election. We trained a model to predict voter likelihood in the 2016 election using an ensemble of machine learning methods. We then applied that predictive model to the 2020 voter file and generated a likelihood between 0 and 1 that each voter would vote.

Likely Absentee Voter Modeling

Citizen generated usable data points and tags from the survey responses that we could use for modeling an individual's likelihood of voting absentee. Specifically, we accepted individual's answers that they were "Likely" or "Very Likely" to vote by mail or dropoff as an intention to vote absentee and considered all other voters as unlikely to vote absentee. We then eliminated all individuals from the survey file who responded that they had voted in the preceding primary election, but in fact had not.

Using an ensemble of machine learning methods, we ultimately trained a model that predicted whether an individual would vote absentee against the dataset resulting from the survey. We then applied that predictive model to 2020 "Likely Voters" as determined by the first stage of the model and generated a likelihood between 0 and 1 that each voter would vote absentee.

* The data that underlies Citizen's projections for Ohio resulted from a survey fielded in late June, while all other state projections are based on data from late July, which may explain projected differences in overall and partisan sentiment.