



CITIZEN DATA

OHIO ABSENTEE VOTING PROJECTIONS

Background

Citizen Data is currently working to project expected 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, Pennsylvania, Wisconsin, and more — will vote absentee for the general election this year.

Ohio was chosen for this project because the state:

- typically sees around 20% of votes cast by absentee ballot in federal elections;
- has both large metropolitan cities and rural areas; and
- held its primary election in April (meaning the state's presidential primary voter turnout data is available to be incorporated in predicting its population's likelihood to vote by absentee ballot)

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

Data and Method: Overview

Our Ohio absentee model combines historical data with recent survey responses to understand how Ohioans have acted in the past and how they intend to in the future in light of COVID uncertainties. Citizen conducted a large-sample (N=4,000) survey and matched the results of that survey to its in-house national voter file. Then, Citizen weighted the dataset against the historical data to predict absentee vote turnout in the 2020 Ohio general election.

Context

In the 2016 general election, approximately 20% of Ohio voters voted absentee. In the 2020 presidential primary election — which was among the first conducted amid the height of the pandemic — approximately 92% of voters voted by absentee ballot.

Results

As of early July, Citizen's modeling projected:

- 5,705,038 of 7,399,732 (77.1%) registered Ohio voters will vote in the November election
- 4,071,755 of 5,705,038 (71.4%) likely voters intend to vote absentee, either by mail or by returning their ballot at a secure dropbox in the November election

Ohio absentee turnout is projected to transcend partisanship:

- 76.5% of Democratic voters intend to vote by mail or by returning their ballot at a secure dropbox
- 79.6% of Republican voters intend to vote by mail or by returning their ballot at a secure dropbox
- 49.6% of non-partisan voters intend to vote by mail or by returning their ballot at a secure dropbox

Ohio absentee voting is projected to be consistently high across racial strata:

- 56.8% of Asian voters;
- 72.4% of White voters;
- 67.6% of Black voters; and
- 62.0% Hispanic voters intend to vote by mail or by returning their ballot at a secure dropbox

Rates of voting absentee are projected to be lowest among the lowest earners in Ohio and highest among the highest earners:

- 58.9% of those earning below \$25,000 per year are projected to vote absentee; and
- 76.4% of those earning over \$200,000 per year are projected to vote absentee

Modeling and Methodology Details

Survey Data Collection

Citizen conducted a large-sample (N=4,000) survey among Ohio active voters between 6/29/2020 to 7/2/2020. Voters in the sample 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 the Ohio 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.

Key Takeaways and Future Model Updates

Citizen projects uniformly high rates of absentee vote turnout across the state of Ohio. We also project slightly higher rates of overall turnout than the 2016 general election. This high rate of absentee turnout is shared across racial, income, and education demographic strata. It is projected to be consistent across parties, but to be slightly higher among older populations than younger populations.

Citizen’s Ohio projection is significantly higher than previously seen rates of absentee voting in Ohio. 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 have a tendency to overestimate their voting intentions. At the same time, our results are plausible relative to a very high absentee turnout rate in the 2020 Ohio primary (92%).

While Citizen is confident in its projections given the data inputs, future updates and improvements to the model are necessary. This is an unprecedented and anxious time for Americans, and many factors that will affect how Ohionans turn out to vote were not included in our pilot modeling process, including but not limited to:

- the varied language people use to reference understand vote by mail;
- unpredictable public-health related uncertainty;
- future projections of the unemployment rate in Ohio; and
- many other factors (see Appendix A for a full list of research questions in progress).

Citizen will continue to update its projections for Ohio as these factors update. Further, Citizen will continue to improve its method as it receives additional input from partners and from experts in the field.

Conclusion

Due to the health and economic impacts of the COVID-19 pandemic, Ohio election administrators should anticipate significantly higher degrees of voting absentee by mail and dropoff in 2020 than in previous elections (with the exception of the 2020 primary earlier this year, which saw absentee turnout at rates higher than what our model projects for the general election). 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.

APPENDIX A: Research Questions in Progress

- *How polarizing is voting by mail along party lines, and how might that impact potential voter turnout?*

- *What programs for voter education and mobilization could affect voter turnout based on various levels of funding?*
- *What role may anti-vote-by-mail efforts have on voter turnout?*
- *How several variables could impact voter turnout, including:*
 - *the level of coronavirus apprehension*
 - *the restrictions implemented by social distancing measures*
 - *Any changes in state policy*
 - *Any USPS delays in mailing and receiving absentee applications and ballots*
 - *Lack of voter confidence that their ballot will be received and counted*
 - *An unprecedented unemployment rate*
 - *The lack of universal education and understanding of the nuances of absentee voting by mail*