CITIZEN DATA

GEORGIA 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, Wisconsin, and more — will vote absentee for the general election this year.

Georgia was chosen for this project because the state:

- has both large metropolitan cities and rural areas; and
- held its presidential primary election in June (meaning the state's presidential primary voter turnout data will be available to be later incorporated in predicting its population's likelihood to vote by absentee ballot); and
- experienced issues with in-person voting in its primary, leading to extremely long lines and other challenges

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

Data and Method: Overview

Our Georgia absentee model combines historical data with recent survey responses to understand how Georgians 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 Georgia general election.

Context

In the 2016 general election, approximately 58.2% of Georgia votes were absentee, out of the total 3,820,881 total votes cast.

Results

As of late July, Citizen's modeling projected:

- 4,074,136 of 6,764,824 (60.2%) registered Georgia voters are projected to vote in the November election
- 2,674,719 of 4,074,136 (65.7%) 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

Georgia absentee turnout projections proved to vary by party:

- ➢ 83% of Democratic Georgia voters are projected to vote by mail or by returning their ballot at a secure dropbox
- ➢ 50.4% of Republican Georgia voters are projected to vote by mail or by returning their ballot at a secure dropbox
- ➢ 61.6% of non-partisan Georgia voters are projected to vote by mail or by returning their ballot at a secure dropbox

Georgia absentee voting is projected to be a significant percentage across racial strata:

- ➢ 31.9% of Asian voters;
- ➢ 39.1% of White voters;
- ➢ 43.9% of Black voters; and
- ➢ 22.5% Hispanic voters are likely to vote by mail or by returning their ballot at a secure dropbox

Rates of voting absentee are projected to be highest among the highest earners in Georgia:

- ➢ 78.1% of those earning more than $250,000 per year are projected to vote absentee; and
- ➢ 59.6% of those earning between $25,000 to $50,000 per year are projected to vote absentee

Modeling and Methodology Details

Survey Data Collection

Citizen conducted a large-sample (N=4,000) survey among Georgia active voters between 7/24/2020 to 7/26/2020. 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 Georgia 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

We project rates of overall turnout higher than the turnout rate of the 2016 general election, and with a slightly higher rate of absentee turnout, as well. These rates of increased absentee turnout are shared across racial and income demographic strata. Absentee turnout rate is projected to be vary significantly by party identification, and we predict that the rate will be slightly higher among older populations than younger populations.

Citizen's Georgia projection is slightly higher than previously seen rates of absentee voting in Georgia. 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 Georgians 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;
- future projections of the unemployment rate in Georgia; and
- many other factors (see Appendix A for a full list of research questions in progress).

While these factors were incorporated into this initial model, Citizen will continue to update its projections for Georgia 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, Georgia election administrators should anticipate slightly 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.

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*