



# CITIZEN DATA

## **Post-Election Vote by Mail Analysis**

*Published: December 4, 2020*

### **Project Overview**

In spring of 2020, the coronavirus was hitting — and sweeping — our nation. As Americans across the country sheltered in place and began adjusting to the “new normal” of masks and a sluggish economy, some began to wonder how this global pandemic would impact the 2020 election.

As part of our non-partisan mission, Citizen Data seeks to produce actionable data and insights in the public interest. Early into the pandemic, our team — along with others who ultimately partnered with us on this effort — recognized that the logistics of in-person voting would present safety concerns and challenges if COVID was to continue into the fall, and that voting by mail would likely be implemented in the majority of states. However, we also identified a knowledge gap: because the circumstances were unprecedented, election administrators and non-partisan non-profits were left without the critical information they needed to plan ahead.

Over the course of the spring, summer, and fall, our team partnered with [Stanford-MIT Project on a Healthy Election](#), [National Vote at Home Institute](#), and more in a robust effort to predict, project, analyze, and understand vote method behavior nationwide and in key states including Iowa, Georgia, Texas, Pennsylvania, North Carolina, Michigan, Florida, and Wisconsin, among others. The goal was to provide election officials and non-partisan groups the data necessary to allocate resources strategically and help ensure a smooth, secure, and representative election.

In addition to bespoke election and voter behavior research, we [conducted a series of surveys nationwide and in key states](#) to representative and weighted samples of registered voters. Our team also [coded and produced predictive models](#) anticipating voter behavior by vote method. Ahead of the election, we compared our projections to 2016 actual voter data in order to demonstrate how much turnout and vote by mail returns would increase; then, post-election, we compared our projections to 2020 actual voter data in order to assess our model accuracy and identify areas for improvement and further precision in the future. With the benefit of a few weeks having passed between the election and now, with all states having their election results certified, our team is now releasing this comprehensive analysis of our findings from the months leading up to the election and beyond.

Our key findings presented throughout this analysis include:

- **Our projections for overall voter turnout numbers and vote by mail turnout were highly accurate; 95.1% and 94.5% accurate on average in key states, respectively. (page 2)**
- **While presidential polling data was notoriously inaccurate across the board, our team's survey data predicting vote method was much more accurate, relatively: in most key states, the final percentage of estimated vote method drawn from our survey data fell within 2 to 3 percentage points of the actual vote method numbers. (page 4)**
- **Democrats were much more likely to plan to vote by mail, and ultimately, they did vote by mail in numbers much higher than non-partisans and Republicans. In fact, Democrats were almost exactly twice as likely to vote by mail than Republicans. (page 6)**
- **There is strong support in USPS among voters broadly, and that support rose steadily as we moved closer to Election Day. (page 9)**
- **Factors such as increased convenience and transparency or COVID cases increasing in the area made these voters say they were much more likely to vote by mail; the USPS experiencing delays that would affect their ability to deliver ballots made voters 41.4% less likely to plan to vote by mail. (page 10)**
- **In the days and weeks leading up to Election, our team predicted that election counts and calls would go more smoothly than some had anticipated; it did. (page 11)**

Citizen Data's full body of work surrounding vote by mail, containing the insights outlined above and much more, can be found [here](#).

### **Turnout & Vote Method: Projections vs. Actuals**

Citizen's modeled projections for overall voter turnout in key states were highly accurate, with 67% of projections landing at 95 to 100% accurate and 89% landing at 90 to 100% accurate. Our most accurate model for voter turnout was in Iowa, where our projection was 99.3% accurate and our final turnout prediction was just 11,192 voters less the final total turnout numbers.

	Total Turnout*	Citizen Projected Total Turnout	Citizen Projection Accuracy
Iowa	1,688,310	1,677,118	99.3%
Georgia	4,998,566	4,901,561	98.1%
Texas	11,317,911	11,021,087	97.4%
Pennsylvania	6,924,179	6,697,642	96.7%
North Carolina	5,524,801	5,250,269	95.0%
Ohio	5,922,202	5,622,311	94.9%
Michigan	5,539,302	5,154,483	93.1%
Florida	11,067,456	10,276,003	92.8%
Wisconsin	3,297,434	2,881,451	87.4%

\*Data updated as of 12/2/2020 and compiled/confirmed by National Election Pool.

Taking a deeper dive into our projections for voters who returned a mail ballot by mail or at a secure dropoff location, our models remained highly predictive. Every one (100%) of our models were over 84% accurate and over half of our models were over 90% accurate. In Pennsylvania and Georgia, two highly competitive states at the presidential level, our models were within just .03% and 1.3%, respectively, of being 100% accurate when compared to actual results. This is particularly remarkable for Pennsylvania, which proved challenging to predict for several reasons, including their historically low level of mail ballots (<5% in 2016), late election policy changes, known legal efforts to challenge some mail ballots, and no early in person voting.

	VBM Returns*	Total Projected VBM	Citizen Projection Accuracy
Pennsylvania	2,616,012	2,616,849	99.9%
Georgia	1,322,524	1,304,862	98.7%
Florida	4,855,677	5,118,024	94.9%
Iowa	577,631	634,522	91.0%
North Carolina	1,001,401	895,689	89.4%
Ohio**	3,512,372	2,993,156	85.2%

Wisconsin	1,306,092	1,552,517	84.1%
-----------	-----------	-----------	-------

*\*Data compiled/confirmed by each states' respective Board of Elections and updated as of 12/2/2020, with the exception of Iowa VBM return data, which was last updated on 11/9/2020.*

*\*\*Note: Ohio's Board of Elections included absentee votes that were sent in both by mail and submitted in-person. While our model was highly accurate as-is, we believe it would be more accurate if we were able to draw upon state data of strictly VBM returns.*

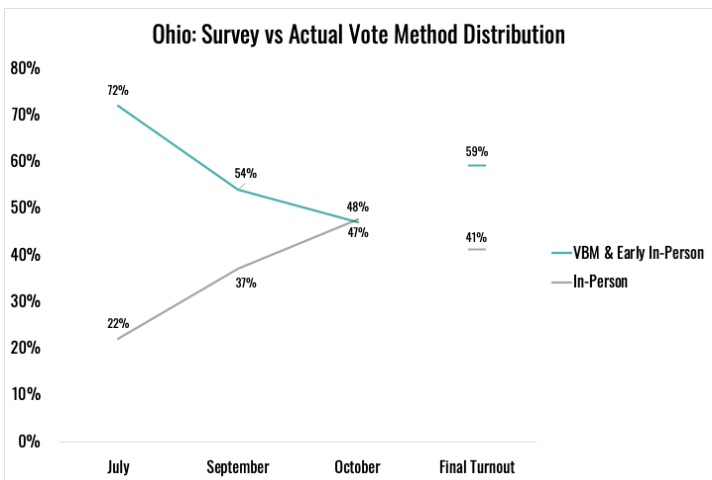
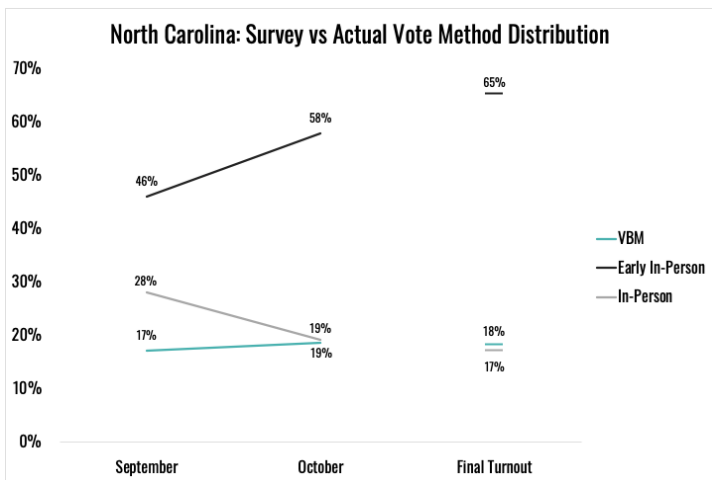
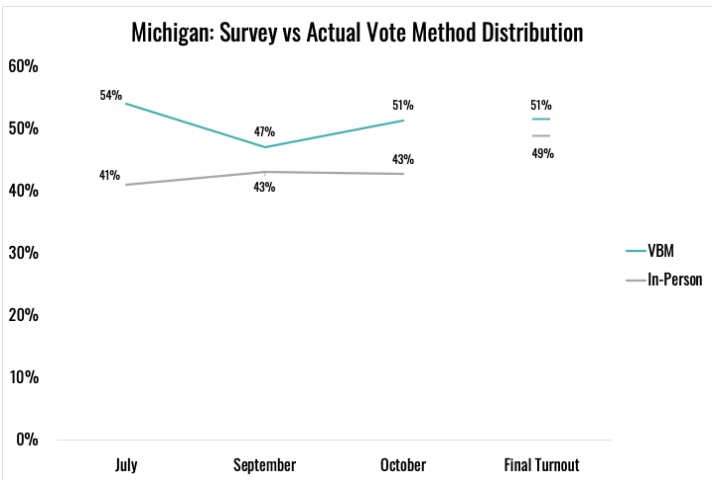
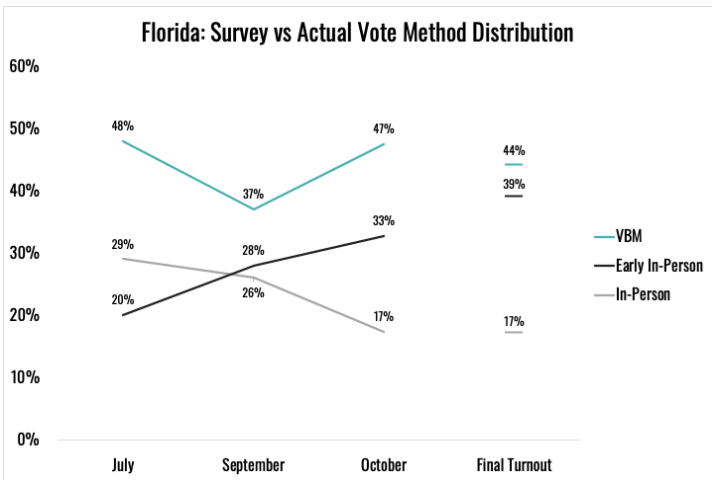
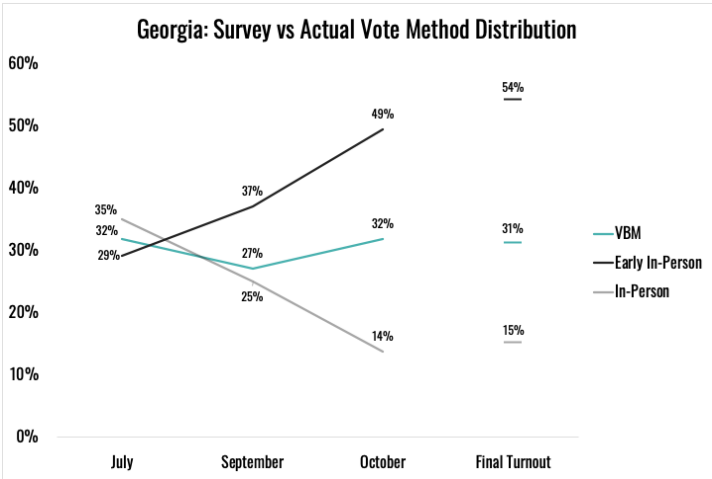
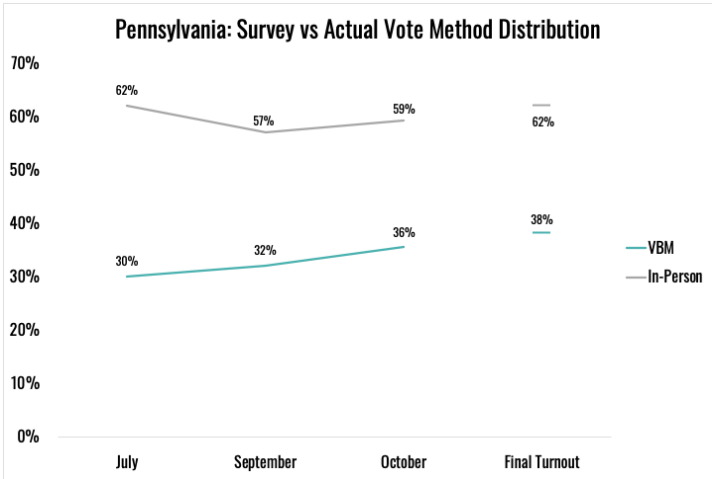
## Vote Method Survey Results: State-by-State Breakdown

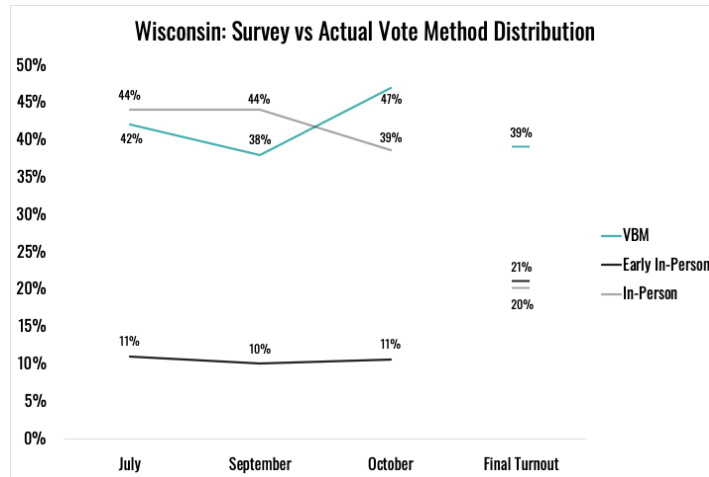
While small-sample public opinion polling [revealed deficiencies](#) again this election cycle, the surveys Citizen fielded to measure intent to vote by vote method were incredibly useful tools for measuring vote by mail, early in-person, and Election Day turnout with high degrees of accuracy. Our team executed benchmark surveys of this nature in July, September, and October, with levels of accuracy around vote method increasing as we moved closer to Election Day.

In most key states, the final percentage of estimated vote method drawn from our survey data fell within 2 to 3 percentage points of the actual vote method numbers. Our survey work was most accurate in states like Michigan, which correctly anticipated 51% of likely voters voting by mail, and Georgia, which anticipated 32% of likely voters would vote by mail (ultimately, 31% did).

We believe that our survey work measuring vote method was largely accurate because: 1) our surveys were larger than an average statewide public opinion poll; each state's completes were N=1,500 to N=3,000 with sample pulled from our national voter file and carefully weighted for representation and 2) while vote method did demonstrate stark partisan breakdowns, questions around logistical plans to vote are relatively less polarizing than other political survey questions, thus eliciting more honest answers from respondents.

In terms of how plans to vote changed among likely voters in key states, our team observed some broad trends over the course of the benchmark surveys. In states where early in-person voting was an option, such as Georgia, Florida, North Carolina, and Wisconsin, voters' intent to vote early in-person increased by an average of 10.5% between July and October. Trends of plans to vote in-person varied by state, but broadly decreased across the board as we moved closer to the Election. In regards to vote by mail specifically, voter intent to vote this way remained relatively stable in key states, with the notable exception of Ohio.

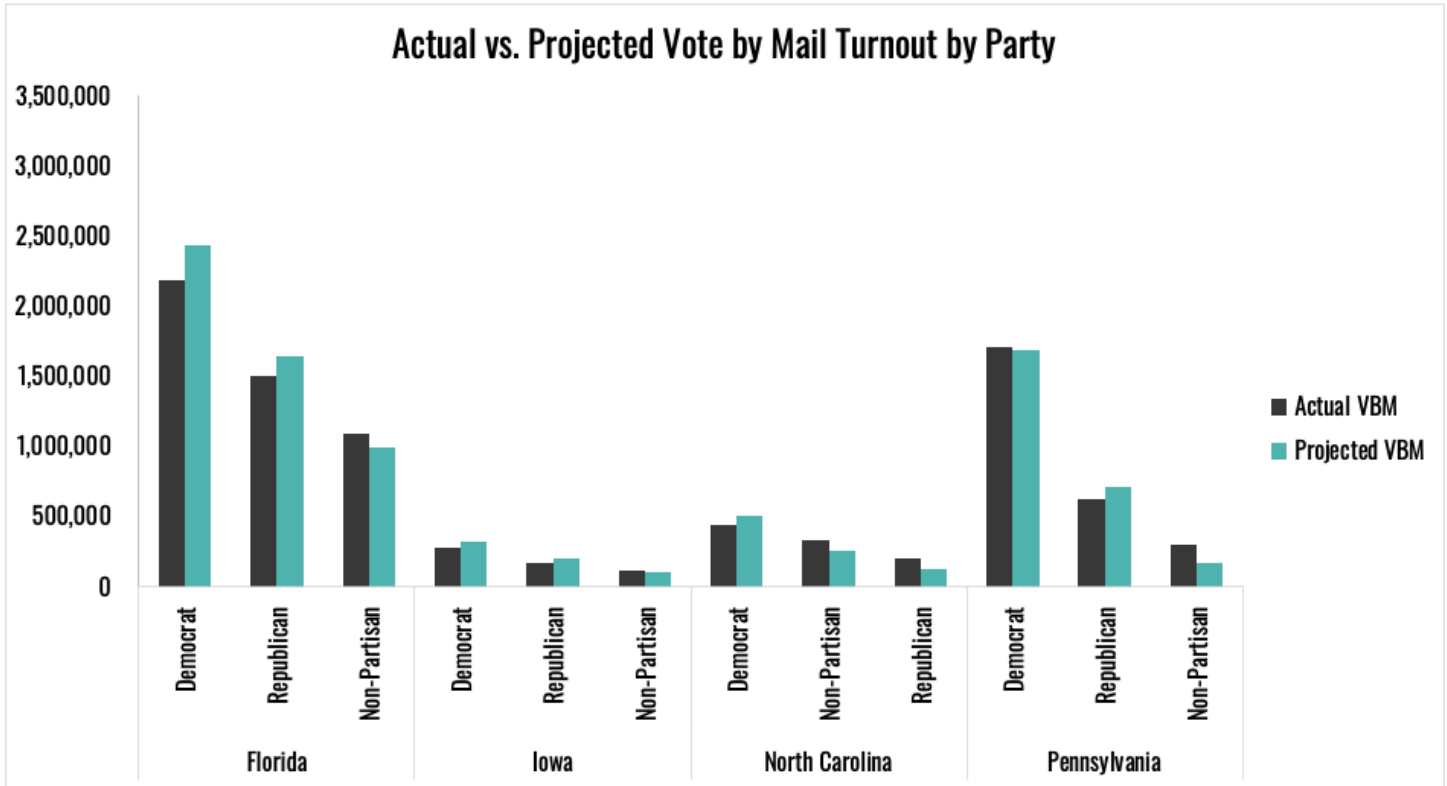




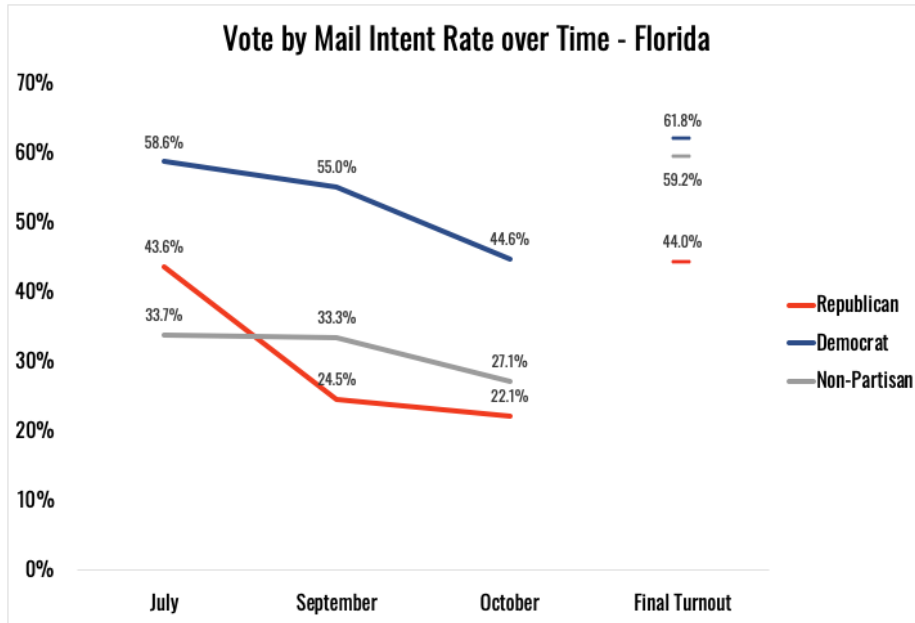
### Turnout & Vote Method: Partisan Divides

Across the key states we examined, Democrats were much more likely to plan to vote by mail, and ultimately, they did vote by mail in numbers much higher than non-partisans and Republicans. In fact, Democrats were almost exactly twice as likely to vote by mail than Republicans.<sup>1</sup> In some states such as Pennsylvania, Democrats voted by mail at a ratio of 3:1 when compared with Republicans. This stark partisan breakdown of vote by mail voters vs. early in-person and/or Election Day voters was one that our team observed and published early on in the project. This observation proved critical in setting expectations for election officials and the media to inform the public about the unique nature of the vote count this year.

<sup>1</sup> For states in which party registration data is available: Iowa, Florida, Pennsylvania, and North Carolina.



As our country moved closer to Election Day, we saw that voters overall, but Republicans specifically, were much less likely — 118.4% less so, on average — to say they would vote by mail in survey responses. In Florida, for example, 58.6% of likely Democratic voters stated in July intending to vote by mail, as opposed to 43.6% of Republican voters. By October, those numbers had decreased to 44.6% to 22.1% respectively, indicating a general trend away from intent to vote by mail, but still showing that intent was much stronger among Democrats than Republicans.



We know that one contributing factor to this was that the more voters overall heard about vote by mail, the less likely they were to plan to do so (especially among Republicans). Further research is needed to definitively conclude causes of this phenomenon, but reasons could include the President’s rhetoric surrounding this voting method, fear of fraud or security issues, and/or simple unfavorability towards it when compared to in-person options.

<i>“How has hearing about vote by mail impacted your likelihood to vote by mail?”*</i>	More Likely	Less Likely	No Change
Texas	11.8%	52.3%	33.3%
Georgia	16.7%	48.0%	30.9%
Florida	21.4%	40.5%	34.4%
Wisconsin	19.4%	40.4%	37.8%
Pennsylvania	19.6%	46.5%	30.9%
North Carolina	12.3%	51.9%	34.0%
Michigan	24.2%	38.2%	35.5%
Ohio	17.5%	50.3%	30.0%
Total	17.5%	46.5%	33.2%



*\*Survey question asked October 17-20 via IVR/P2P to N=1,000 likely voters in Florida, Georgia, North Carolina, Michigan, Ohio, Wisconsin, Texas, and Pennsylvania. Results were weighted and matched to Citizen's voter file of all registered voters.*

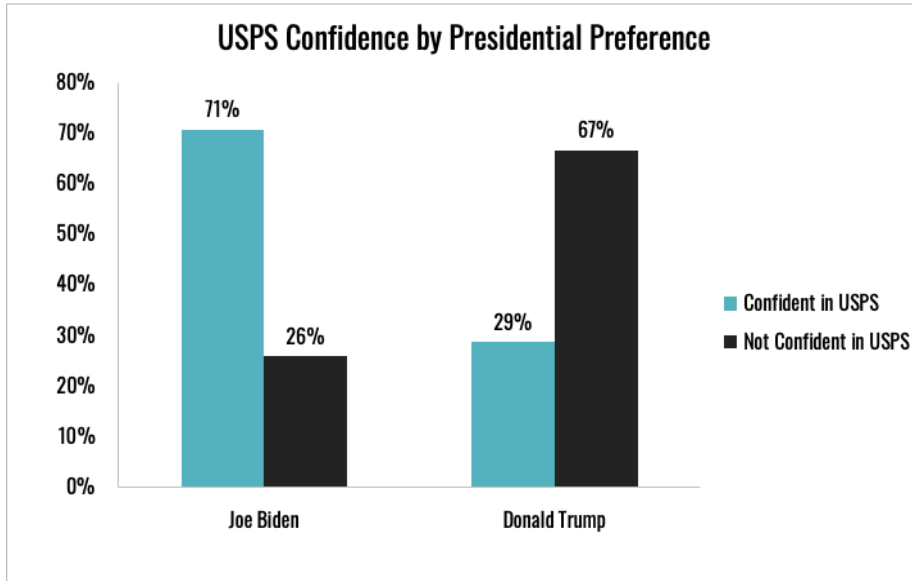
## **Trust in USPS as a VBM Factor**

Our team investigated trust in the United States Postal Service (USPS) as a factor of intent to vote by mail. We established two major findings.

First, there is strong support in USPS among voters broadly, and that support rose steadily as we moved closer to Election Day. In a September national survey, 52% of total likely voters said that they were confident in the USPS's ability to deliver ballots to voters and to election officials in time, including 89% of voters who were planning to return their ballot by mail and 68% of voters who planned to drop off. In surveys conducted in key states in September and October, confidence rose an average of 13% in key states, and final confidence levels closest to the election were at 56.2% and higher across the board.

<b>USPS Confidence</b>	<b>September Survey</b>	<b>October Survey</b>
Florida	49.9%	63.4%
Georgia	48.7%	60.8%
Michigan	56.3%	64.1%
North Carolina	41.6%	59.0%
Ohio	56.5%	57.0%
Pennsylvania	50.0%	56.2%
Wisconsin	57.0%	66.1%
Total	47.6%	60.6%

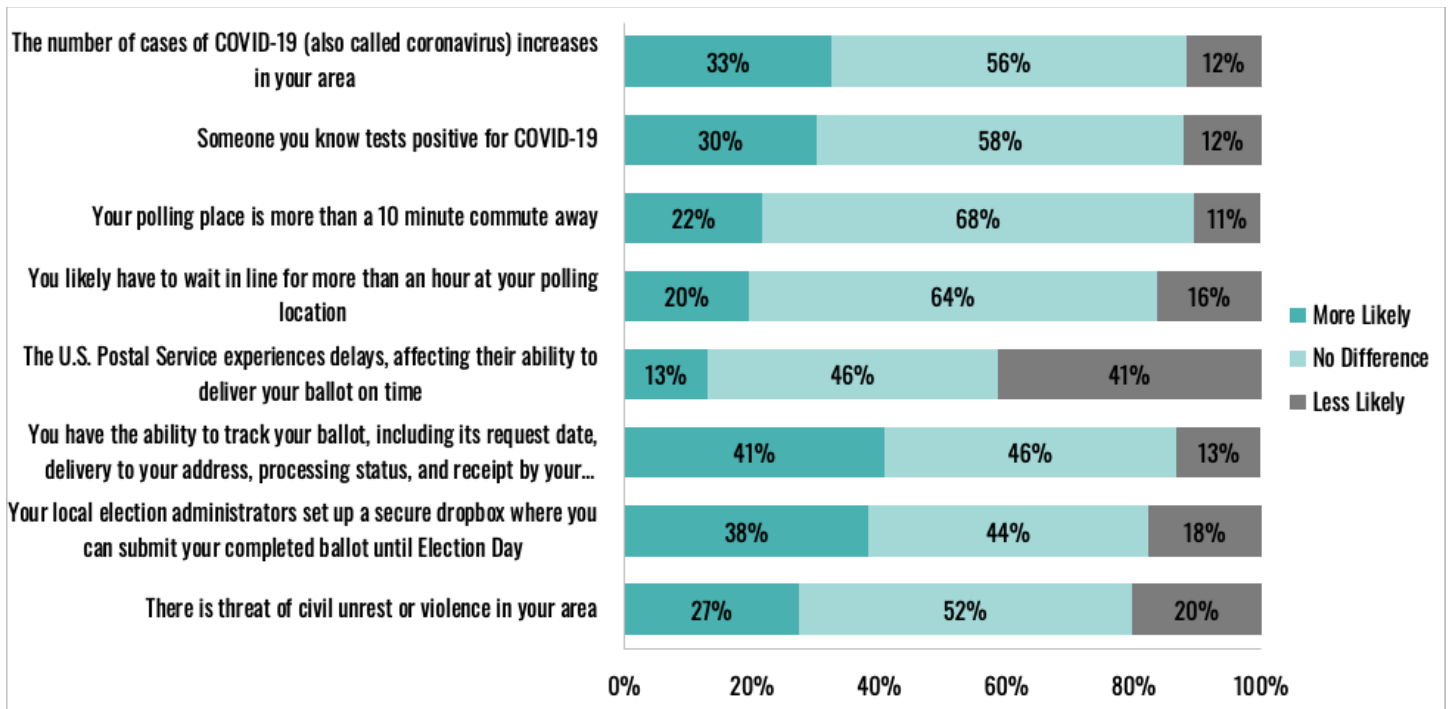
However — and secondly — confidence in USPS looked much different among both Democrats and Republicans and Biden and Trump supporters. Our team consistently found that both Democrats and Biden supporters had much higher levels of confidence in USPS. Looking at confidence in USPS among supporters of presidential candidates in our national survey fielded in September, 71% of voters planning to vote for Joe Biden were confident in the USPS while only 29% of Donald Trump voters were. In fact, 67% of those planning to vote for Trump were decidedly not confident in the USPS.



### **VBM & Current Events: A Dynamic Landscape**

Our team recognized that many dynamic and rapidly-evolving factors would affect how voters planned to vote and attitudes towards vote by mail. To the best of our ability, our team considered these factors when predicting, projecting, and analyzing voter turnout and vote by mail.

In a national N=30,000 survey conducted in September, we asked respondents from a representative sample of likely voters nationwide about these various factors. Factors such as increased convenience and transparency or COVID cases increasing in the area made these voters say they were much more likely to vote by mail (40.7 and 32.% more, respectively). Notably, the USPS experiencing delays that would affect their ability to deliver ballots on top made voters 41.4% less likely to plan to vote by mail.



### Vote by Mail Impact on Election: Analysis

In the days and weeks leading up to Election, our team [predicted](#) that election counts and calls would go more smoothly than some had anticipated; it did.

Most voters in battleground states had already voted prior to Election Day, meaning that Election Day turnout would (and did) lag when compared to prior years. Accordingly, election administrators were well-prepared for high overall turnout and a surge in mail-in ballots. Efforts such as our team’s, in addition to many other initiatives through broad coalitions of non-partisan non-profits, supported and further prepared these officials with the information they needed to plan.

Our team’s predictions largely held up in the key states we studied most closely, including Florida, which reported results early, and Arizona, Wisconsin, and Michigan, which were largely clear the day after the Election. And, though it felt longer for some, the official call in favor of a Biden victory was made by [media outlets and the Associated Press](#) on the Saturday morning after the election, November 7th — just three days after polls closed.

Despite many in the media predicting blue or red “mirages,” our team [anticipated](#) that the results were more likely to be a mixed bag in real time wherein absentee, early votes, and Election Day votes were

reported nearly simultaneously. That prediction held true (with the exception of states like Pennsylvania, where an anticipated “red mirage” did occur due to delays in counting mail-in votes there).

While the Election (referred to some as “Election Week”) went smoothly overall, our prediction of a more “normal” night was undercut by the widespread inaccuracy of presidential candidate public polling, which exaggerated Biden’s support in all battleground states. Accordingly, one of our team’s greatest learnings is to put even more trust into our sophisticated data-driven modeling (which relies on actual voter behavior data as a foundation) than small-sample public opinion polls from traditional political vendors.

The prediction of a “normal” night was also undercut by the President’s unexpected and immediate challenges to the results, which continue as of the publication of this memorandum. Nonetheless, accessible data and insights about election results have proven instrumental to ensuring the integrity of the election in court, among election officials and certifiers and the public, to the best extent possible.

## **Project Methodology**

### *Surveys*

We began by conducting large-sample N=4,000 surveys via IVR/P2P in key states: on June 29 to July 2 for Ohio and between July 24 to 26 for Florida, Georgia, Michigan, and Wisconsin. Voters in the samples for each state 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 each state’s electorate, and were sampled evenly across congressional districts. After each state survey was completed, we matched each respondent to our national voter file.

On September 3 to 8, we then surveyed 30,000 likely voters nationwide via IVR/P2P. This survey data was matched to Citizen’s voter file of all registered voters, and weighted. The breakdown for key states measured in the survey was as follows: N=2,000 for Florida, N=1,016 for Georgia, N=954 for Michigan, N=1,004 for North Carolina, N=1,106 for Ohio, N=1,202 for Pennsylvania, and N=578 for Wisconsin.

Lastly, between October 17 to 20, our team surveyed N=1,000 likely voters in Florida, Georgia, North Carolina, Michigan, Ohio, Pennsylvania, Wisconsin, and Texas, via IVR/P2P. Results were weighted and matched to Citizen’s voter file of all registered voters.

### *Predictive Models*

To create our predictive models for vote method in key states, we first created a custom turnout model tailored to each state. For each state, we predicted each voter's likelihood to vote in the 2020 election by training an ensemble of machine learning methods. We then predicted based on voters in each specific state from the 2016 and 2018 general elections.

Then, we created separate models to predict the likelihood that those individuals who had not yet requested a ballot would do so. For individuals likely to request a ballot according to the model, we assumed they would return that ballot by mail at a rate equal to the mean of that group's turnout likelihood. For those who had already requested a ballot, we assumed that that group would vote by mail at a rate equal to the mean of that group's turnout likelihood.

Our initial model projections were first released in mid-September. From that point forward, Citizen continuously refreshed, updated, and published its models, drawing upon the most recently-available vote by mail ballot return data. When our team's direct feed of ballot return data indicated that an individual voter had returned their ballot, their return probability became 100%. When it was indicated that an individual voter had requested a ballot, their request probability also became 100%, overriding the predicted request probability, and their return probability became the probability at which the individual would turn out to vote. The numbers indicated above represent the latest of those updates.