

Does Vote-by-Mail Cause Voters to Gather Information About Politics?*

James Szewczyk
Department of Political Science
Emory University
james.szewczyk@emory.edu

June 28, 2019

Abstract

In this paper, I examine the effects of vote-by-mail on voter behavior and voter knowledge. I argue that vote-by-mail electoral systems result in a more informed electorate, because voters have additional time with their ballot and access to resources to conduct research about races on the ballot that they know nothing about. I present the results of two empirical studies that support this prediction. First, I find that all-mail elections in Utah cause a 6.368 percentage point decrease in straight ticket voting. This is consistent with the logic that voters spend more time with their ballots when voting by mail relative to when they are voting at a polling place. Second, I estimate the effects of vote-by-mail on voter knowledge using an original repeated cross-sectional survey that was fielded during the 2018 general election in California. The research design exploits the implementation of the California Voters Choice Act (VCA), which resulted in five counties in the state switching to an election system in which all voters in the counties are sent a mail-in ballot. I find that the VCA causes an increase in voter knowledge and an increase in time that voters spend gathering information about the election. However, the reform does not affect the prevalence of political discussion or levels of knowledge about the party identification and ideology of candidates.

*I thank the MIT Election Data and Science Lab and the Madison Initiative of the William Flora Hewlett Foundation for generously funding this research. I am also grateful to Tom Clark, Zac Peskowitz, Adam Glynn, participants at SPSA 2019, and participants at the 2019 State Politics and Policy Conference for helpful feedback on this project.

1 Introduction

Scholars have extensively documented the low levels of political knowledge in the American public (e.g., Delli Carpini and Keeter 1996, Zaller 1992). A lack of information about politics among citizens can generate a number of negative consequences for representative democracy. For example, changing information levels in the electorate has implications for voters' choices at the ballot box as well as representation in government and policy outcomes. Indeed, increasing citizens' knowledge about politics affects the behavior and preferences of voters. When information is provided to voters about candidates and parties prior to their ballots being cast, voters adjust their choices at the ballot box (Fowler and Margolis 2014, Klein and Baum 2001). Moreover, institutions that facilitate access to information about politics affect elected official behavior and policy outcomes (Snyder and Strömberg 2010). As a result, institutional reforms that affect how voters gather political information will affect the quality of a democratic government.

In this paper, I argue that an electoral reform, vote-by-mail, changes the behavior of voters in a way that increases the likelihood that they gather information about politics. Consequently, voters who reside in a jurisdiction that implements the use of mail-in ballots will be more informed than voters who cast their ballots at a polling place. Voters are able to obtain relevant political information when voting by mail because they have additional time with their ballots when voting. Voters can use this extra time to access resources on the Internet and consult their voter information guides to conduct research about the election. On the other hand, voters do not have the ability to use these information sources when voting at a polling place if they come across a race on the ballot that they do not know anything about.

Previous research suggests that vote-by-mail will cause voters to spend more time filling out their ballots and be more informed. For example, Hanmer and Traugott (2004) argue that this will be the case when Oregon became the first state to implement all-mail elections. Further, other work exploits the staggered implementation of vote-by-mail in Washington and provides evidence in line with the argument that voters gather more information about politics when they receive their ballots through the mail. Szewczyk (2018) shows that there are changes in policy outcomes and elected official behavior that reflect a more informed electorate after the switch to vote-by-mail in the state. The results of this study indicate that there is a decrease in taxing and spending as well as a decrease in negative findings on accountability audits in Washington's municipalities after vote-by-mail is implemented. In addition, vote-by-mail causes a decrease in ballot roll-off on statewide ballot measures (Szewczyk 2018) and on other statewide races (Marble 2018) in

Washington. This is interpreted as evidence in support of the argument that voters take the time to gather political information when voting by mail.

Despite this previous research on how vote-by-mail will cause increases in voters' knowledge about elections and other work that examines the relationship between the reform and political information¹, previous research has been unable to directly estimate the effects of vote-by-mail on voter knowledge. I expand our knowledge of this important electoral reform in this paper by directly answering the following question: does vote-by-mail cause voters to gather more information about politics than they would when voting at a polling place? In particular, I present the results of two empirical studies that take advantage of the staggered roll-out of all-mail election systems in Utah and California.

In order to establish that vote-by-mail does indeed change how people vote, in the first study I estimate the effects of the implementation of all-mail elections in Utah on the prevalence of straight ticket voting. Voters in Utah have the option to make a single mark on their ballot to cast a vote for the member of a given political party in each individual race on the ballot; that is, cast a straight ticket vote. I argue that if citizens use more time to fill out their ballots when voting by mail, they will be less likely to use this straight ticket option because they will consider each race individually. Exploiting within-county changes in the use of compulsory vote-by-mail in Utah over time, I find that the electoral institution causes a 6.368 percentage point decrease in straight ticket votes. This is consistent with my theory of the relationship between vote-by-mail and voter behavior.

In the second study, I directly estimate the effects of vote-by-mail on voter behavior and voter knowledge. I exploit a policy intervention in California, the Voter's Choice Act (VCA), which resulted in five counties in the state switching to an all-mail electoral system during the 2018 primary and general elections. This provides exogenous variation in vote-by-mail use among Californians. I field a repeated cross-sectional survey of eligible voters during the 2018 general election in California that measures voter knowledge about statewide races on the ballot and voter behavior at two points in time: first, prior to when voters were sent mail-in ballots; second, after the election took place. I use the survey data to directly test my theory of how vote-by-mail affects voter knowledge and how citizens gather information about politics. I estimate the effects of the VCA on the outcomes of interest with a two sample two-stage least squares (TS2SLS) esti-

¹For example, Meredith and Malhotra (2011) show how early voting options, like vote-by-mail, can result in individuals casting their votes differently than if they voted on election day when new information about candidates is revealed late in the campaign.

mation strategy. TS2SLS allows me to combine information from the California voter file and my original survey to examine the relationship between vote-by-mail and voter behavior.

In the process of generating the TS2SLS estimates I calculate three quantities of interest. First, the first-stage model uses the California voter file data and reveals the effect of the VCA on the probability that registered voters use a mail-in ballot. I find that the VCA increases the likelihood an individual uses vote-by-mail by 0.115. Second, I estimate the direct effect of implementing the VCA on voter behavior and knowledge with the original California survey data. These are the reduced form estimates and are interpreted as the intent-to-treat (ITT) effect or the net effect of a county implementing the VCA on the outcomes of interest. Third, I estimate the impact of actually using a mail-in ballot with 2STSLS. The 2STSLS estimates can be simply calculated as the ratio of the coefficients of interest from the reduced form and first-stage models. This quantity is the local average treatment effect (LATE) of using vote-by-mail and can be interpreted as the impact of everyone in a jurisdiction switching from voting at a polling place to using a mail-in ballot. The reduced form and TS2SLS estimates indicate that the VCA caused changes in voter behavior. Using a mail-in ballot results in voters increasing the amount of time they spent conducting research about the election. In addition, implementing the VCA caused an increase in the likelihood that voters are able to answer questions about the ballot measures that appeared on the ballot correctly. Vote-by-mail does not appear to affect the probability of political discussion or the types of information sources that voters report to use.

2 Political Information, Elections, and Vote-by-Mail

Voters in the United States must make decisions about many races when voting. Elected officials ranging from dog catcher and county coroner to mayor and governor regularly appear on the ballot. In addition, in a number of states and local jurisdictions there are systems of direct democracy where politicians or the public place measures on the ballot. These ballot measures are often complex and ask voters to make decisions about technical policy issues. As a result, in order to make informed choices when casting their ballots, voters must have high levels of knowledge about policy issues and candidates across multiple levels of government.

However, voters have many competing demands for their time and the ballot that they cast is rarely pivotal in deciding the outcome of an election (Downs 1957). The proliferation of entertainment options

also gives voters the ability to avoid political news, which provides knowledge of upcoming elections (Prior 2007). Therefore, when the average voter shows up to vote at a polling place on election day it is unlikely that she has sufficient knowledge to make an informed decision about each race on the ballot. Indeed, previous research finds that voters may not have enough political knowledge to meet this ideal of being sufficiently informed. Voters are often unable to correctly identify the policy positions of candidates (Freeder, Lenz and Turney 2019), and they may not make voting decisions that are most in line with their policy preferences when casting their ballots for prominent electoral contests that are heavily covered by the media, like presidential elections (Bartels 1996, Lau, Andersen and Redlawsk 2008). So, it is not surprising that citizen know even less about elections for statewide office and local elections (Binder et al. 2016).

In light of these limitations, how do voters make decision while voting? Scholars have long recognized that voters use a variety of information shortcuts when voting. Ballots provide informational cues to voters that may aid them when making decisions (e.g., Matson and Fine 2006). For example, markings on the ballot identifying which candidate is the incumbent (Krebs 1998) and labels indicating candidate party identification (Schaffner, Streb and Wright 2001) provide informational shortcuts to voters. These heuristics may make up for the lack of knowledge that voters have about politics and cause them to make vote choices more in line with their preferences (e.g., Lupia 1994).

There are a number of drawbacks to voters relying solely on heuristics when casting their ballots. Most notably, sometimes heuristics do not appear on the ballot. For example, in many local elections in the United States citizens are not provided with the partisan affiliation of candidates when voting. In addition, during primary elections in which voters must choose between a set of candidates from the same political party a candidate's party affiliation does differentiate among candidates. Second, some simple informational shortcuts may require a base level of political knowledge and sophistication that some voters lack (Delli Carpini and Keeter 1996). Finally, voters use shortcuts like a candidate's race, ethnicity, and beauty, which could penalize candidates from racial minority groups (Crowder-Meyer, Gadarian and Trounstine 2019) and inflate the vote shares of more attractive candidates (Jones and Price 2017). These effects are exacerbated when voters are asked to complete cognitively complex tasks and when information levels are low (Crowder-Meyer et al. 2018, Jones and Price 2017).

It is apparent that many voters in the United States do not have high enough levels of political knowledge to make informed decisions at the ballot box, and there is, consequently, much room for an improvement in the amount of information that citizens obtain about politics before voting. In this paper, I argue that vote-

by-mail elections can help solve this issue and result in voters being more informed than they would when voting at a polling place. Vote-by-mail affects how voters gather information about politics for a variety of reasons. When casting a ballot in a vote-by-mail election a voter has an extended period of time with her ballot and access to information about candidates and propositions that she knows nothing about. In addition, voters do not need to take on the higher cost of learning about what is going to be on the ballot before they go to vote, as they would when voting at a polling place. Because of these lower costs of obtaining information (Martinelli 2006), individuals who use vote-by-mail will seek out more information about politics than they would when voting at a polling place. And, voters will have higher levels of political knowledge when voting by mail compared to when they vote at a polling place. I now expand on the mechanisms that result in the additional acquisition of political information when elections are held using mail-in ballots.

Most importantly, voters have additional time with their ballot when voting by mail. In jurisdictions that use vote-by-mail in the United States, ballots are typically sent to voters about thirty days before an election. This time gives individuals who receive their ballots in the mail the ability to seek out information about politics. However, when a voter goes to a polling place she may be confronted with races that she knows nothing about in the voting booth. At this point, the voter does not have the ability to conduct additional research about the election. This additional time with the ballot is useful because voters can use it to seek out new political information. Indeed, mailing a ballot to voters may serve as a form of a nudge (Thaler and Sunstein 2009) that encourages a voter that would otherwise be uninformed to gather information about the issues and races on the ballot. Because there may be many low salience races during a given election, a voter may not even know what she does not know until she sees her ballot. If voters learn that they do not know anything about a particular election, they may be encouraged to conduct research about the race if they have the opportunity. When voting by mail a voter has an extended amount of time with her ballot and the resources to gather this relevant information about politics.

There are a number of information sources that voters have access to and that they can use to obtain information about politics when they are filling out their ballot at home. One of the main tools available to voters in many states is a voter information guide. States that have high levels of vote-by-mail usage often mail these guidebooks directly to voters. A voter information guide provides a wide variety of types of information to voters. For example, the voter information guide in California contains biographical information about candidates for federal and statewide office as well as summaries and detailed descriptions of propositions that appear on the ballot. When a voter is casting her ballot in the comfort of her own home,

she has the ability to consult the voter information guide about candidates and ballot measures that she knows nothing about.

Voters also have access to the Internet when casting a mail-in ballot, which gives them an avenue to uncover a great deal of political information. Notably, individuals have access to candidates' websites, which will have relevant information about the election. A voter can also use a search engine, like Google, to look up more details about the policy positions of and background information on candidates. There are a variety of other sources of information that voters can exploit, either through the Internet or in print, when they do not have to vote in a polling place. For example, citizens can consult local news coverage to gather information about low salience races when voting with a mail-in ballot. Further, special interest groups, like the League of Women Voters and the Sierra Club, regularly produce pamphlets and websites with recommendations of which candidates to vote for and which ballot measures to vote in favor of.

Finally, mailing a ballot to a voter's home may lead to a more social voting experience and encourage political discussion, which will increase voter information levels (Eveland 2004). Although scholars have argued that vote-by-mail decreases the social pressure aspect of voting² (Funk 2010), I propose that the electoral institution introduces a new social dimension to the voting process. That is, when voting by mail individuals may be encouraged to discuss the election with a spouse or a roommate. This may result in the person that voter lives with sharing new information with her or it can encourage the housemates to seek out political information together because they receive their ballots at the same time. In an editorial in *New York Magazine*, a mail-in ballot voter in California describes his social voting process during the 2018 primary election as follows:

So this weekend, [my wife] and I plan to sit down with our mail ballots and our handy-dandy state and county voter guides...We will have the leisure to make our way through the incredibly extensive list of statewide candidates facilitated by California's jungle nonpartisan top-two primary...That means there are 32 U.S. Senate candidates and 27 gubernatorial candidates on the June 5 ballot. It's really not something you want to first encounter alone, in a voting booth (Kilgore 2018).

This quote summarizes the intuition behind how vote-by-mail results in a substantial change in how citizens vote. Voters are faced with an excessively large number of decisions on election day. It is easier to face these

²However, vote-by-mail does not eliminate the possibility that an individual may still be asked if he voted, which drives voter turnout due to disutility from lying about voting (DellaVigna et al. 2017).

choices when a voter is in her home, because she has access to information and can discuss the candidates and issues with other individuals as she fills out her ballot.

Based on this theoretical discussion, I argue that vote-by-mail changes the voting experience and encourages individuals to gather information about politics. I hypothesize that when voters cast their ballot in a vote-by-mail election they will be more informed than they would be when voting at a polling place. This is because voters will spend more time filling out their ballots and will use this time to conduct research about the races that they are voting on. I test these predictions by exploiting policy interventions in Utah and California that resulted in the staggered roll-out of all-mail elections these states. In the next section, I present the results of the study that examines the effects of mail-in ballot use on straight ticket voting in Utah. Then, in the following section I report the estimates of the effects of vote-by-mail on individual-level voter behavior and knowledge in California.

3 Vote-by-Mail Decreases Straight Ticket Voting in Utah

I first build on previous work that finds vote-by-mail causes a decrease in ballot roll-off (Marble 2018, Szewczyk 2018) by examining how all-mail elections affect straight ticket voting. The purpose of this analysis is to convincingly establish that implementing vote-by-mail causes changes in how citizens vote before directly looking at the effects of mail-in ballot use on voter knowledge. Straight ticket voting refers to the phenomenon in which citizens cast their votes for members of the same political party for all of the races that appear on the ballot. I expect that vote-by-mail will cause a decrease in straight ticket voting. This is because when casting a mail-in ballot citizens spend additional time voting and gather more information than they would when voting at a polling place. If voters are spending an increased amount of time with their ballots when filling it out in the comfort of their own home, they will be more likely to consider each candidate on the ballot individually and not simply rely on party identification as a heuristic.

There is a large body of work that studies the determinants of straight ticket voting in democracies. While some previous research suggests that split ticket voting, voting for multiple political parties across different races, is the result of voters behaving strategically (Carsey and Layman 2004, Fiorina 1996) or not having strong partisan preferences (Beck et al. 1992), other work focuses on how electoral systems impact split ticket voting. These studies find that reforms that simplify the voting process results in decreases in straight ticket voting (Barnes, Tchintian and Alles 2017). Further, institutions that expand access to political

information, like the availability of local news, lead to increases in split ticket voting (Moskowitz 2018). Relatedly, I argue that vote-by-mail eases the process of gathering political information and will cause a decrease (increase) in straight (split) ticket voting.

I test this hypothesis using data from Utah. The state of Utah presents an excellent context to examine the effects of vote-by-mail on split ticket voting for two reasons. First, Utah is one of nine states that have an straight ticket vote option on the ballot. This allows voters to cast their vote for all of the candidates in a given political party with a single mark on the ballot rather than having to mark each individual contest in a given election. Therefore, I can directly examine how vote-by-mail changes the rate in which this straight ticket option is chosen without having to rely on survey data to approximate ticket splitting. Second, Utah has incrementally introduced a system of all-mail elections in the state. Counties are able to decide if and when they would like to implement the compulsory vote-by-mail, which has resulted in the staggered implementation of the reform. After a county implements all-mail elections each registered voter in the county is mail a ballot for each election. This allows me to identify the effects of all-mail elections using within-county changes in electoral systems over time.

3.1 Data

I estimate the effects of switching to all-mail elections on straight ticket voting in Utah using data gathered from the state's Lieutenant Governor's Office, which is the office that governs elections in the state. I obtained information on the year that each county in the state implemented all-mail elections from direct correspondence with election officials in the Lieutenant Governor's Office. The first county in the state to run an all-mail election was Duchesne in 2012. By 2018, all of Utah's counties conducted elections entirely by mail except for Carbon and Emery Counties. Table C.1 in the Online Appendix displays the year that each county in the state switched to compulsory vote-by-mail.

I also collect data on voter turnout and the number of straight ticket votes cast for each party by county in even-year general elections from 2008 to 2018 from the Lieutenant Governor's Office.³ Using this information I calculate the percentage of straight ticket votes cast in each county from 2008 to 2018. This variable, $PercentStraightTicket_{ct}$, is the total number of straight ticket votes cast for all political parties divided by the total number of votes cast in the election. On average, 29.92% of voters in Utah cast straight

³The data begins in 2008 because this is the first year that the number of straight ticket votes cast are available on the Lieutenant Governor Office's website.

ticket votes over the time period of interest. Figure 1 displays a plot of the percentage of straight ticket votes cast over time. It is important to note that this is likely an underestimate of the number of straight ticket votes cast in Utah. The count of straight ticket votes comes from the total number of citizens who mark the straight ticket vote option on the ballot, but voters may also vote for all of the candidates from a given political party by marking each race individually. More detailed summary statistics of the data are displayed in Table B.1 in the Online Appendix.

Finally, I construct variables that measure the time-varying demographic characteristics of Utah’s counties with the U.S. Census American Community Survey 5-Year Estimates. These variables include controls for counties’ total population, race, gender, education, and median age.

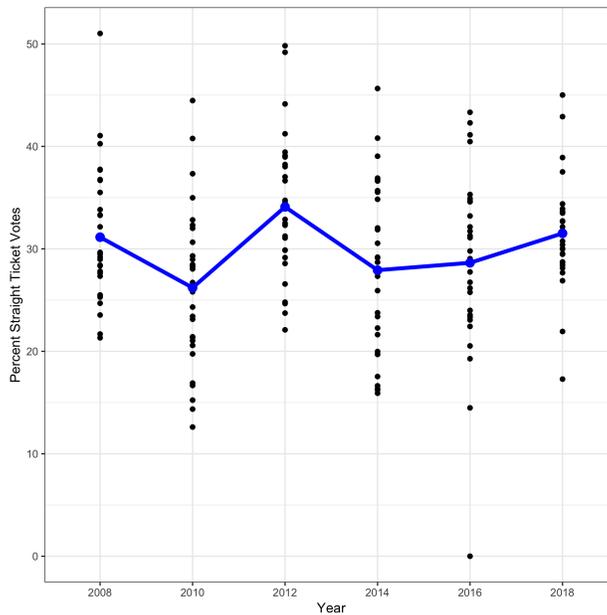


Figure 1: Plot of the percentage of straight ticket votes cast in each county in Utah from 2008 to 2018. Blue line and points show the statewide average over time.

3.2 Results

I estimate the effects of vote-by-mail on the percentage of straight ticket votes by estimating models of the following form,

$$PercentStraightTicket_{ct} = \beta AllMail_{ct} + \alpha_c + \gamma_t + X'_{ct} \Theta + \alpha_c \times t + \epsilon_{ct} \quad (1)$$

where the dependent variable is $PercentStraightTicket_{ct}$, the percentage of straight votes cast in county c in year t . $AllMail_{ct}$ is a dummy variable equal to one if county c held its elections entirely by mail in year t . The models also include county, α_c , and year, γ_t fixed effects. In some specifications I add a vector of time varying county demographic control variables⁴, \mathbf{X}_{ct} , in addition to county time trends, $\alpha_c \times t$. The estimate of β is quantity of interest that is estimated from this model: the effect of all-mail elections on the percentage of straight-ticket votes. β is interpreted as the causal effect of conducting elections entirely by mail under the assumption that the timing of the switch to all-mail elections is independent of unobserved variables after controlling for time invariant county specific factors, year specific shocks that are constant across counties, time varying county-level demographic characteristics, and county time trends. The research design improves on previous research that examines the effects of vote-by-mail on straight-ticket voting. For example, Hanmer and Traugott (2004) make cross-sectional comparisons between the rates of straight-ticket voting in Oregon before and after the state implemented all-mail elections. The staggered implementation of all-mail elections in Utah allows me to more credibly identify the effects of all-mail elections by comparing changes in the rate of straight-ticket voting in jurisdictions that did implement all-mail elections to changes in jurisdictions that did not implement the reform.

The results of the models are displayed in Table 1.⁵ The model in Column 1 includes county and year fixed effects and Columns 2 and 3 add demographic controls and county time trends to the specification. The results indicate that implementing all-mail elections caused a substantial decrease in the number of straight ticket votes cast in Utah. Depending on the specification, I find that there is a 5.237 percentage point to 6.368 percentage point decrease in the percentage of straight ticket votes after the switch to vote-by-mail. Not only are the point estimates statistically significant ($p < 0.01$), but this is a substantively large reduction in straight ticket voting. The substantive magnitude of the effects is slightly larger than Moskowitz's (2018) finding that access to local news increases split ticket voting by 4-5 percentage points. Further, these effects of all-mail elections on split ticket voting in Utah have occurred during a time of increased nationalization and polarization of elections (Abramowitz and Webster 2016).

⁴The control variables include total population, median age, percent Black, Native America, Hispanic, Asian, high school graduate, college graduate, and female. The data is obtained from the U.S. Census American Community Survey 5 Year Estimates from 2010 to 2017.

⁵In Table D.1 in the Online Appendix I display the results of leads and lags models to determine if future switches to all-mail elections cause changes in the percentage of straight ticket votes. I do not find that this is the case, which provides confidence in the causal interpretation of these results.

Table 1: The Effects of Vote-by-Mail on Straight Ticket Voting in Utah

	<i>Dependent variable:</i>		
	Percent Straight Ticket Votes		
	(1)	(2)	(3)
All-Mail Election	-5.237*** (1.451)	-5.476*** (1.525)	-6.368*** (1.942)
County Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	No	Yes	Yes
County Time Trends	No	No	Yes
Observations	174	174	174
Adjusted R ²	0.655	0.668	0.677

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors clustered by county.

These findings, of course, do not rule out the possibility that voters in Utah choose to mark each race on the ballot for the candidate of a given political party rather than choosing the straight party option after the switch to vote-by-mail. That being said, voters are less likely to engage in split ticket voting when they are not able to use a straight ticket option (Darcy and Schneider 1989, Reynolds and McCormick 1986), so these estimates do likely indicate legitimate increases in ticket splitting after the implementation of vote-by-mail. In all, the results present evidence consistent with the argument that voters spend more time filling out their ballots when voting by mail relative to when they are voting at a polling place. Instead of only marking their ballots once in order to vote for a political party’s entire slate of candidates, voters make decisions about each individual race separately when voting by mail. This provides compelling evidence that vote-by-mail induces substantial changes in the way the citizens vote.

4 Vote-by-Mail Increases Time Spent Gathering Information and Voter Knowledge in California

The previous section shows that switching to all-mail elections in Utah caused a decrease in the percentage of voters who choose the straight party option on the ballot. What explains these findings? As discussed, I propose that voters are spending additional time with their ballots in order to gather information about politics when voting by mail. I test this hypothesis by estimating the effects of a reform that caused California citizens in some counties to use a mail-in ballot for the first time during the 2018 general election.

4.1 The Context of California and Data Overview

Vote-by-mail is a widely used form of voting in California, because voters can choose to opt into permanent absentee voter status and receive a mail-in ballot for each election. In addition, counties can choose to conduct elections entirely by mail in precincts that have fewer than 250 registered voters. Alpine County, for example, has been conducting all of its elections entirely by mail since at least 2000 because their precincts are sufficiently small. Figure 2 displays the percentage of votes in California’s counties that were cast using a mail-in ballot in each even-year general election from 2000 to 2018. The blue line displays the statewide average over time. As can be seen in the figure, there has been a dramatic increase over recent years in the percentage of ballots cast by mail in California. In 2000 only 28.4% of voters voted with an absentee or mail-in ballot, but by 2018 74.1% of voters in California had eschewed traditional polling places in favor of vote-by-mail. Further, there was a major reform signed into law on September 29, 2016 that greatly expanded the prominence of vote-by-mail in California. Senate Bill No. 450, also known as the California Voter’s Choice Act (VCA), dramatically overhauled the administration of elections in California and provides a much more flexible voting experience.

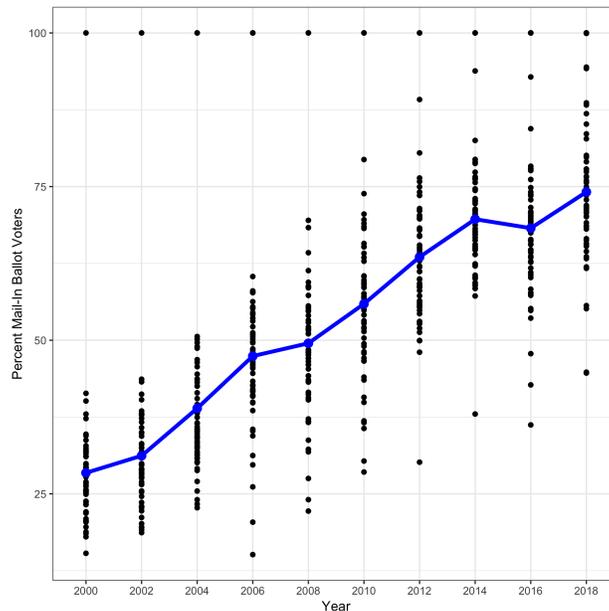


Figure 2: The percentage of ballots cast by mail in California general and midterm elections from 2000 to 2018.

4.1.1 Background on the California Voters' Choice Act

Under the VCA system, all voters in a county that administer the reform, regardless of whether or not they have chosen to be a permanent absentee voter, are delivered a ballot 28 days prior to the election through the mail. After a voter makes her selections and fills out the ballot, it can be returned through the mail, at a county ballot drop box, or at a county vote center. Ballot drop boxes and county vote centers replace polling places and are located throughout each county. The number of these drop boxes and vote centers are determined with a formula that is based on the population of each county. Figure 3 shows a description of the process of voting by mail under the VCA from an informational flier for voters in a county that implemented the reform in 2018. In addition to voting with her mail-in ballot, a voter in a VCA county can choose to go to any vote center in her county to obtain a ballot in person, fill it out, and return it at the same time. Even though each voter is still sent a ballot in the mail, this method of voting under the VCA approximates the experience of voting at a polling place. In addition, voters that do not reside in a county that implemented VCA can still opt to use a mail-in ballot, as they could prior to the VCA being passed by the state legislature.

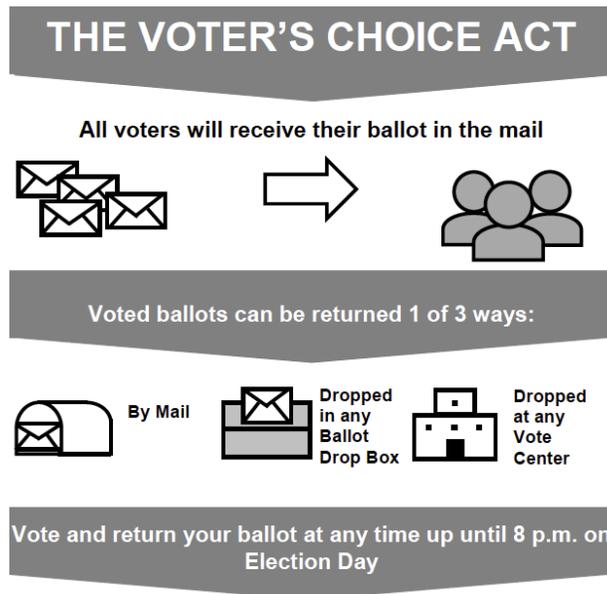


Figure 3: Description of voting methods under the California Voter's Choice Act. (Source: Sacramento County Website)

Importantly, the entire state did not simultaneously switch to the VCA. The choice of whether or not a county implements the reform is up to the county's board of supervisors. Fourteen counties in the state were given permission to switch to the new system in 2018 and the remaining counties in the state can begin to

switch in 2020. This excludes Los Angeles county, which will not be permitted to change electoral systems until 2024. Of the fourteen counties given permission, five counties in the state opted into the new vote-by-mail system for the 2018 primary and general elections while nine counties chose to wait to potentially implement the reform during a future election.



Figure 4: Map of California with counties that implemented the Voter's Choice Act shaded in gray.

There are a number of reasons as to why counties who were able to chose not implement the VCA in 2018. Most notably, the debate surrounding the reform fell along partisan lines in the state legislature⁶ as well as among county supervisors with Democrats being more supportive of the policy than Republicans. In fact, the all-Republican board of supervisors in Orange County refused to even consider implementing the reform in 2018 (Wildermuth 2017). Election administration officials also cited the upfront costs as a drawback of switching to the new system. Despite the fact that the VCA will save money on election administration in the long run, some counties were concerned with the initial startup costs of implementing the reform, like replacing outdated voting machines. Even counties that eventually decided to switch to the VCA faced conflict over the price of the reform. For example, the board of supervisors in Nevada county only approved the reform after the estimated cost of implementing the VCA was lowered from \$303,000 to \$258,000 (Riquelmy 2018). In addition to funding issues, county officials were concerned with the potential

⁶For the final vote in the State Assembly, 46 Democrats and only 1 Republican voted in favor of the bill. In the State Senate, 25 Democrats and 2 Republicans voted for the bill.

for abuse of a vote-by-mail based electoral system as well as potential negative consequences of closing local polling places that citizens have grown accustomed to (Wildermuth 2017).

Figure 4 displays a map of California with counties that implemented the VCA during the 2018 general and primary elections shaded in gray. The five counties that implemented the electoral reform are Madera, Napa, Nevada, Sacramento, and San Mateo. The source of exogenous variation in this study is that voters who reside in one of these counties that switched to the VAC in 2018 will be sent a mail-in ballot to their place of residence. This is essentially an source of encouragement for individuals to uptake the treatment of interest: using a mail-in ballot to vote in the 2018 general election. As a result of voters' ability to still be able to choose their voting method, individuals in California do not have to comply with this encouragement to be treated with vote-by-mail, which I discuss further when presenting the empirical strategy. Nevertheless, this is a shock to election administration that resulted in many new individuals choosing to use a mail-in ballot. I exploit this by administering a repeated cross-sectional survey among California voters during the 2018 general election that includes a battery of questions measuring political knowledge and voter behavior.

4.1.2 Survey Data

The survey instrument asks respondents questions about their objective knowledge of these statewide races that were on the ballot during the general election. In addition, to measure how voters' experiences change when using vote-by-mail, I ask subjective questions about how informed they feel during the election cycle and how they conducted research about races on the ballot. The survey only asks about statewide elections in order to hold the campaign context constant across respondents that live in different counties. A copy of the full survey questionnaire is located in Section A of the Online Appendix.

Before discussing the survey design in more detail, I briefly describe the races on the ballot that the survey questions were based on. On November 6, 2018 in California there were elections for governor, lieutenant governor, and treasurer among other statewide executive positions. The most high profile race on the ballot was that for governor, which pitted Democrat Gavin Newsom against Republican John Cox. Meanwhile due to California's top-two primary system two Democrats were up for election for lieutenant governor, Eleni Kounalakis and Ed Hernandez, while the candidates for treasurer were Fiona Ma, a Democrat, and Greg Conlon, a Republican. In addition, California citizens decided on a total of 12 statewide ballot measures during the 2018 general election. Two measures dealt directly with tax issues. Proposition 5 proposed to remove a requirement for the transfer of property tax bases that would expand the ability of

older and disabled homeowners to transfer their property tax base to a new property. This may decrease revenue for school and local governments. Proposition 6 sought to repeal a transportation tax law from 2017 that increased fuel and vehicle taxes. There was also a measure, Proposition 9, that met all requirements to appear on the ballot but was removed by the California Supreme Court. The measure proposed to split California into three different states. Finally, Proposition 10 and Proposition 11 concerned rent control policies enacted by local governments and breaks for private sector employees, respectively.

The survey was administered by the survey research firm Civiqs. The Civiqs panel provides a number of advantages over similar survey research firms. First, the online platform and large panel allows Civiqs to quickly gather data. This is important in this context as I am interested in obtaining a snapshot of voter information levels at particular points of the election cycle. Second, Civiqs has a large number of panel members who reside in California. This gives me the ability to collect a sufficient number of responses from individuals who live in counties that switched to the VCA in 2018, which only accounts for about 7% of the state's total population.

The survey consisted of two waves that asked two different cross-sections of respondents to answer the same survey questionnaire at two different points in time.⁷ The final data set used in the analysis includes a sample of 2,253 eligible voters in California who were surveyed during the first wave of the survey and 1,166 eligible voters who were interviewed during the second wave, for a total of 3,419 respondents.⁸ The first wave of the survey was fielded from October 1, 2018 to October 4, 2018 and the second wave was fielded from November 7, 2018 to November 12, 2018. Importantly, the first wave of the survey was administered prior to when ballots were mailed to voters and the second wave was administered after election day. I choose to field the second wave after election day, because many vote-by-mail voters choose to deliver their ballots to a vote center on election day.

I measure voter information levels and behavior using a number of different variables that are generated from individuals' responses to the survey questions. To ease comparison across the dependent variable, they are rescaled to range between zero and one. The first set of dependent variables capture how voters gather information about politics. First, I generate the measure *Time Spent Researching* using a question

⁷There are slight changes made to verb tenses on the questionnaire for the second wave to reflect that it was fielded after the election took place.

⁸The number of observations used in the analysis differs slightly from the total number of respondents in the data that I received from Civiqs. Respondents are dropped for three reasons. First, they are dropped if their county of residence is not in California. Second, respondents matched to the TargetSmart data are removed if they are not registered to vote in California. Third, there were a small number of respondents younger than 18 who are removed from the sample.

that asks respondents how much time they spent conducting research about the election. Higher values of this variable indicate that voters spent more time conducting research. Second, I ask respondents to identify whether or not they used any of the following sources of information: the Voter Information Guide, a local newspaper, or an Internet search engine. *Information Sources* is equal to one if a respondent reported using any of these sources of information. Third, the variable *Political Discussion* is equal to one if a respondent has discussed politics with someone that she lives with and zero otherwise. I only observe this variable for individuals who report that they do not live alone, because I am particularly interested in how vote-by-mail shapes interactions about politics with the people that voters live with.

I then take the eleven questions that ask about knowledge of statewide elections and create an index of the number of correct responses for each individual that took the survey. These questions ask respondents to identify the ideological position of candidates for governor, the party ID of candidates for lieutenant governor and treasurer, and correct details about ballot measures. The resulting variable, *Knowledge Index*, is the average number of questions that a respondent answered correctly. Next, I create three variables that indicate whether or not respondents answered all of the questions correctly in the three categories of knowledge questions. These variables, *Ballot Measure Correct*, *Party ID Correct*, and *Ideology Correct*, are equal to one if a respondent answered all of the questions regarding ballot measures, party identification, and ideology correct, respectively, and zero otherwise. The final dependent variable is a self-assessment of how informed voters feel about the candidates for statewide executive office and statewide ballot measures. This variable, *Knowledge Self-Assessment*, is the average response from two separate questions asking about candidates and ballot measures.

Table 2: Summary Statistics of Dependent Variables

Statistic	N	Mean	St. Dev.	Min	Max
Time Spent Researching	3,419	0.306	0.307	0.000	1.000
Information Sources	3,419	0.613	0.487	0	1
Discussion	2,434	0.810	0.393	0	1
Knowledge Index	3,419	0.523	0.273	0.000	1.000
Ballot Measure Correct	3,419	0.180	0.384	0	1
Party ID Correct	3,419	0.167	0.373	0	1
Ideology Correct	3,419	0.737	0.440	0	1
Knowledge Self-Assessment	3,419	0.357	0.306	0.000	1.000

Summary statistics of the dependent variables across both waves of the survey and among individuals who do and do not live in VCA counties are displayed in Table 2. The data also includes information on the following demographic characteristics of respondents: age, party identification, education, race, and gender. Summary statistics of the full set of covariates in the dataset are displayed in Table B.2 in the Online Appendix. I now briefly discuss the summary statistics of the dependent variables. First, the average of *Time Spent Researching* is 0.306, which is equivalent to individuals reporting that they spent fewer than five minutes conducting research about the 2018 general election. Over half of the respondents report using one of the *Information Sources* that were asked about and over three quarters of individuals report engaging in *Political Discussion*. Next, the mean of the *Knowledge Index* dependent variable indicates that respondents were able to answer about half of the survey questions correctly. The mean values of *Party Identification Correct* and *Ballot Measure Correct* are 0.167 and 0.180, respectively. This makes it apparent that it was difficult for respondents to correctly identify the political party that candidates for relatively low salience races, like treasurer and lieutenant governor, and to answer questions about ballot measures. On the other hand, 73.7% of respondents can identify which side of the ideological spectrum the candidates for governor belong on. Finally the average of *Knowledge Self-Assessment* is 0.357. This translates into respondents saying that they feel moderately to somewhat uninformed about the statewide elections and ballot measures.

4.1.3 Voter File Data

In order to examine how the VCA affected the propensity of voters to use mail-in ballots, I use individual-level voter file data from the California Secretary of State. The data is an extract of VoteCal, California's statewide voter registration database, taken in April of 2019. The publicly available files from the database contain two essential pieces of information: first, there is a list of all registered voters in the state; second, there is a list of who voted and the method (i.e. polling place or mail-in ballot) that each individual used to cast their ballot in each election for which there is data available. For the analyses presented in this paper, I use the voter file data from 2016 and 2018. The central reason for only using these two years of data is to match the number of waves of the survey. In addition, the VoteCal database only consistently contains voter registration and vote history information starting in 2016, which is the first year that counties were required to report this information to the state. For years prior to 2016 there is uneven coverage in VoteCal across counties. Using the information on the method individuals used to vote, I create the variable *UseMailBallot*, which is a dummy variable equal to one if an individual used a mail-in ballot to vote in a given election.

In addition to the indicator for the use of a mail-in ballot, I am able to obtain information on the gender, party identification, age, and race of individuals from the voter file. Age is constructed using each respondent's dates of birth. Gender and party identification are coded using information in the voter file that citizens provide when they register to vote. To predict the race of registered voters I match their surnames to the Census Bureau's Surname List and use information on county of residence, age, and party identification to calculate the probability they belong to each of the following racial groups: White, Black, Hispanic, Asian, and other (Imai and Khanna 2016). I identify an individual as being a member of the racial group that has the highest probability after this procedure is completed.

4.2 Estimation Strategy

There are a number of potential pitfalls when estimating the effects of vote-by-mail, or any electoral reform, on voter behavior. For example, one could test the hypothesized relationship using data from a state that allows registered voters to opt into using a mail-in ballot by comparing the behavior of voters who use a mail-in ballot to those that choose to vote at a polling place. This would lead to positively biased estimates of the effects of vote-by-mail if individuals who choose to use the voting method are more likely to conduct research about an election and be more informed about politics. It is very likely that this is the case, because opting into using vote-by-mail suggests that the individual is an active voter and cares enough about politics to take this additional step after or while registering to vote. Using a related research design, a researcher could make cross-sectional comparisons between individuals in jurisdictions that have compulsory vote-by-mail and those that do not have all-mail elections. If jurisdictions adopt all-mail elections because their citizens care about and are more informed about politics than jurisdictions that do not, this research design would also lead to positively biased estimates of the effects of vote-by-mail.

My estimation strategy overcomes both of these issues. First, I exploit the implementation of the Voter's Choice Act in California, which provides exogenous variation in the use of vote-by-mail. Second, because the survey instrument was fielded at multiple points in time I can identify the effects of using a mail-in ballot with within-county changes in the use of all-mail elections over time. This allows me to hold the time invariant characteristics of counties constant. I identify the effects of vote-by-mail on the outcome variables of interest with the two-sample two-stage least squares (TS2SLS) estimator developed by Angrist and

Krueger (1992; 1995).⁹ This method has been used in a variety of applications ranging from examining the relationship between teen drinking and education outcomes (Dee and Evans 1997; 2003) to understanding the effects of public assistance on food insecurity (Borjas 2004).

TS2SLS is an instrumental variables model that can be used in instances in which the endogenous and outcome variables of interest are not observed in the same data set, but there are two data sets that contain at least one of these pieces of information along with the instrumental variable. In the context of this study, the endogenous treatment variable is a voter actually using a mail-in ballot to vote and the exogenous instrument is living in a county that is assigned to the VCA. The outcome variables are the measures of voter knowledge and behavior. The two datasets used in the analysis are the voter file data and the survey data. In the voter file data, I observe if individuals live in a VCA county (the instrument) and if they used a mail-in ballot during the 2016 and 2018 general elections (the endogenous variable). However, the voter file data does not have any measures of voter knowledge. On the other hand, the survey data contains the instrumental variable (living in a VCA county) and the outcome variables (measures of voter behavior and knowledge), but there is no measure of whether or not respondents used a mail-in ballot in the two most recent elections.¹⁰ TS2SLS provides a framework in which to combine information from both of these data sources to estimate three distinct quantities of interest and characterize the relationship between these estimates: the first-stage, the reduced form, and the TS2SLS estimates. I now turn to a discussion of the three estimating equations and the interpretation of the main quantities of interest.

4.2.1 First-Stage

I estimate the first-stage models using the California voter file data with the following equation,

$$UseMailBallot_{ict} = \beta_1 VCA_{ict} + X'_{ict}\pi_1 + u_{1c} + v_{1t} + \epsilon_{1ict} \quad (2)$$

where $UseMailBallot_{ict}$ is an indicator for whether or not individual i in county c used a mail-in ballot in year t . VCA_{ict} is a dummy variable equal to one for individuals who live in a VCA county in the time period after the reform was implemented. β_1 is the estimate of the effect of the VCA on the probability that

⁹This estimator is what Angrist and Krueger (1995) refer to as split-sample instrumental variables. I use the terminology from Inoue and Solon (2010) who refer to it as the two-sample two-stage least squares estimator and show that it is preferable to the related two-sample instrumental variables estimator (Angrist and Krueger 1992).

¹⁰In the survey data, I only observe whether or not respondents from the second wave of the survey used a mail-in ballot during the 2018 general election. The survey does not ask first-wave respondents about their use of mail-in ballots in previous elections because it is unlikely that respondents will be able to accurately report voting method in previous elections.

a registered voter uses a mail-in ballot and is the coefficient of interest from this model. X_{ict} represents a vector of individual-level exogenous control variables. The control variables include measures of race, gender, age, party identification, and permanent absentee status. Finally, u_{1c} and v_{1t} are county and time period fixed effects, respectively. For all of the analyses, I report robust standard errors clustered by county.

4.2.2 Reduced Form (Intent-to-Treat Effect)

I next estimate the reduced form, which is the effect of the VCA on the behavioral outcomes of interest. The reduced form can be interpreted as the intent-to-treat effect (ITT) of the VCA on the dependent variables.¹¹ In other words, it is the net effect of implementing the VCA on voter knowledge in the context of the 2018 general election in California. I estimate the reduced form using the survey of eligible voters in California with the following equation,

$$Y_{ict} = \beta_2 VCA_{ict} + X'_{ict}\pi_2 + u_{2c} + v_{2t} + \epsilon_{2ict} \quad (3)$$

Y_{ict} represents the variables measuring how voters gather information about politics and their knowledge of races that appeared on the ballot. The right hand side of the reduced form equation takes the same form as that for the first-stage equation. That being said, for all of the reduced form models I estimate a model with an extended set of control variables, X_{ict} . In particular, I add controls for respondents' levels of education. This set of controls is not included in the main specification because I do not observe it in the voter file used for the first-stage model, and all control variables must be the same for both the reduced form and first-stage models when estimating an instrumental variables model. The coefficient of interest from this model is β_2 , which is the effect of implementing the VCA on voter knowledge and behavior while controlling for time invariant county factors, year specific shocks that are constant across counties, and individual-level demographic characteristics of respondents.

In this context, the ITT effect will underestimate the effect of actually using a mail-in ballot on the outcomes of interest. This is because there is a great deal of noncompliance with the VCA as a mechanism that causes individuals to use a mail-in ballot. There are three ways that an individual can not comply with their treatment status: first, an individual in a VCA county can cast their ballot at a county vote center; second, an individual in a non-VCA county can use a mail-in ballot because they are a permanent absentee

¹¹This can also be interpreted as the difference-in-differences estimate under the parallel trends assumption.

voter or live in a precinct with fewer than 250 registered voters; third, an individual can choose not to vote. Figure 5 shows that noncompliance did indeed take place during the 2018 general election in California. Using data on mail-in ballot usage from the California Secretary of State's Office, Figure 5 reveals the percentage of voters in VCA counties and in non-VCA counties that used vote-by-mail. 93.1% of voters who live in VCA counties chose to use a mail-in ballot in 2018 and, thus, complied with their treatment status. However, among voters who do not reside in VCA counties 72.3% opted for vote-by-mail and received the treatment while only 27.7% of voters in these counties went to a traditional polling place. As a result, simply comparing changes in the outcome variable between VCA and non-VCA counties over time, as is done in the reduced form models, may underestimate the impact of vote-by-mail on voter knowledge.

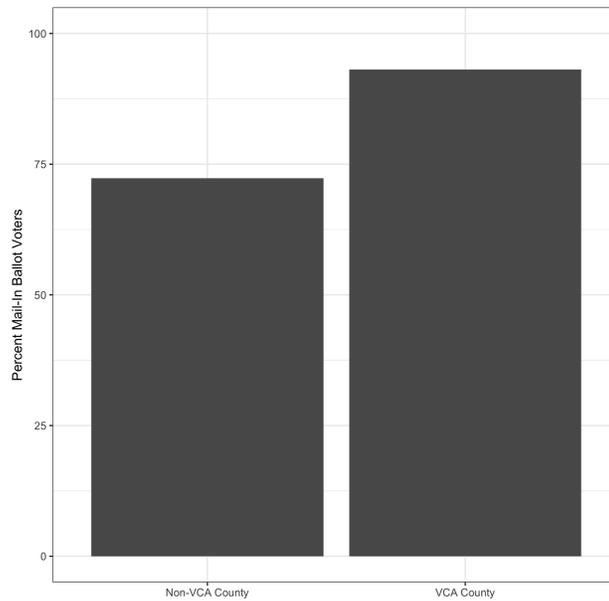


Figure 5: Percentage of voters in California using a mail-in ballot during the 2018 general election in VCA and non-VCA counties.

4.2.3 TS2SLS (Local Average Treatment Effect)

Although the reduced form identifies the ITT effect, or the net effect of the VCA, I am also interested the impact of actually using a mail-in ballot on voter behavior and knowledge. I estimate this quantity of interest with the second-stage of the TS2SLS model. The model is exactly identified so, as observed by Dee and

Evans (1997; 2003), the TS2SLS estimate is simply the ratio of the coefficients of interest from the reduced form and first-stage, like in the standard 2SLS setting.¹²

$$\beta_{TS2SLS} = \frac{\beta_2}{\beta_1} \quad (4)$$

Recall that β_2 is the estimated effect of the VCA on the probability that an a registered voter uses a mail-in ballot and β_1 is the estimated effect of the VCA on the measures of voter behavior and knowledge. Combining these two estimates, β_{TS2SLS} is interpreted as the effect of using a mail-in ballot among individuals that were caused to use a mail-in ballot by the VCA. This is the Local Average Treatment Effect (LATE). This quantity can also be interpreted as the estimated effect of all individuals in a jurisdiction switching from using a polling place to vote to using vote-by-mail to cast their ballots. The cluster-robust standard errors for the TS2SLS estimates are calculated using the method described in Pacini and Windmeijer (2016).¹³

The causal interpretation of β_{TS2SLS} , along with the estimates of β_1 and β_2 , requires that selection into adopting the VCA in 2018 is independent of the outcome variables of interest after controlling for the time invariant characteristics of counties and year specific shocks. I argue that this is plausible because the main determinants of the implementation of the VCA in 2018, which are population and partisanship according to news coverage, are unlikely to change drastically over time. In addition to this assumption, in order for the estimate of β_{TS2SLS} to be unbiased the exclusion restriction must hold: that the VCA only affects voter’s knowledge levels and how they gather information about politics through the use of a mail-in ballot. It is indeed unlikely that the implementation of the VCA itself would have an effect on voter knowledge and how they gather information about the election except for through the channel of using vote-by-mail.

4.3 Results

In this section I present the results of the California study. I first discuss the first-stage before presenting the reduced form and second-stage models. The first-stage estimates are displayed in Table 3. Column 1 shows the estimated effect of the VCA on mail-in ballot use without any individual-level demographic controls while in Column 2 I add these additional control variables to the model. The preferred first-stage estimate used to construct the TS2SLS estimates includes the set of control variables. I estimate that the VCA caused

¹²Alternatively, the TS2SLS estimates can be produced using the cross-sample fitted values from the first-stage.

¹³I use the estimator for the variance of the TS2SLS estimates shown in Equation 12 (Pacini and Windmeijer 2016). As suggested for cluster-robust standard errors, I substitute estimates of the cluster-robust variances from my first-stage and reduced form models for the robust variances in the equation.

a .107 to .115 increase in the probability that a voter used a mail-in ballot. Importantly, this is a very strong instrument, and considering the large sample size, it is not surprising that the results are highly statistically significant for both specifications of the model. That being said, this is only a moderate increase in the likelihood that vote-by-mail is used. This is probably the case because so many individuals in California used mail-in ballot prior to the implementation of the VCA, so there is a limit on the potential magnitude of the increase. It is also important to note that this suggests, as we will see in the next set of the results, that the TS2SLS estimates will be substantially larger than the reduced form estimates.

Table 3: First-Stage: The Effects of the VCA on VBM Ballot Use

	<i>Dependent variable:</i>	
	Use VBM	
	(1)	(2)
VCA	0.107*** (0.013)	0.115*** (0.009)
County Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Controls	No	Yes
Observations	30,157,797	30,157,797
Adjusted R ²	0.066	0.411

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors clustered by county.

The first-stage results shows that the VCA had a positive impact on the use of mail-in ballots among registered voters in California. What is the impact of this uptake in vote-by-mail on the behavior of voters and how informed they were about the races on the ballot during the 2018 general election in California? To answer this question, I turn to a discussion of the reduced form and TS2SLS models. The results of the models that estimate the effects of the VCA on voter behavior are displayed in Table 4. Each row of the table shows the results of three different models for a given dependent variable. All of the dependent variables are standardized such that they range between zero and one. Columns 1 and 2 reveal the reduced form estimates with two different sets of control variables. The basic control variables include ace, gender, age, party identification, and permanent absentee status. The models with extended control variables add measures of educational attainment to this specification. Column 3 displays the results of the TS2SLS models with the basic set of control variables. Recall that the TS2SLS estimates will be equal to the ratio of the coefficients from the reduced form and first-stage (Table 3, Column 2).

In the first row of Table 4 I show the effects of the VCA and being a vote-by-mail ballot voter on the amount of time that respondents spent conducting research about the election. The estimates are positive and remain at least marginally significant ($p < 0.1$) across all of the specifications of the model. With the reduced form models I estimate that vote-by-mail causes a 0.022 to 0.024 increase in the dependent variable. According to the TS2SLS model, however, I estimate that using a mail-in ballot causes a significantly larger 0.188 increase in the measure of the amount of time that voters spend conducting research about the election. To aid in the substantive interpretation of these estimates, it is important to note that there are five response categories to the question asking respondents how long they spent conducting research about the election: fewer than 5 minutes, five minutes to one hour, one to two hours, three to 4 hours, and more than four hours. A 0.25 increase on the zero to one scale indicates that a respondent moves from one category to the next. Therefore, the substantive size of TS2SLS estimate is only slightly smaller than a unit increase in the amount time voters expend gathering information about the election. The most common response to this question from survey respondents was the first category, so the effect of using a mail-in ballot for the typical respondent would be causing them to increase the amount of time they spent conducting research about the election from less than five minutes to about 5 minutes to less than an hour.

Let us turn next to the dependent variable that measures whether or not a respondent used the Official California Voter Information Guide, an online search engine, or a local newspaper to gather information about the 2018 California general election. Across all three specifications of the model, the coefficient on the independent variables of interest are positive, which indicates an increase in the probability that a voter uses one of these sources of information. However, the estimates never achieve standard levels of statistical significance. Finally, examining the bottom row of Table 4 indicates the estimates for the discussion dependent variable are actually negative, which is in the opposite direction of what I expected. These estimates are also quite imprecisely estimated.

The survey data suggests that voters spend more time conducting research about an election when voting by mail relative to when they vote at a polling place. However, there is little evidence that the electoral reform causes other changes in the behavior of voters, like the types of information sources that they use or their propensity to discuss politics with individuals that they live with. Despite the null effects on these latter two outcomes, these findings corroborate my interpretation of the previous results in which I find that vote-by-mail causes a decrease in straight-ticket voting in Utah. I argue that this decrease in straight-ticket voting, along with the effects of vote-by-mail on ballot roll-off documented in other studies (Marble 2018,

Table 4: Reduced Form and TS2SLS: The Effects of the VCA and Vote-by-Mail on Voter Behavior

<i>Dependent Variable:</i>	<i>Reduced Form Estimates:</i>		<i>TS2SLS Estimates:</i>
	VCA County		Use VBM
	(1)	(2)	(3)
Time Spent Researching	0.022* (0.013)	0.024* (0.013)	0.188* (0.112)
Information Sources	0.028 (0.021)	0.033 (0.021)	0.244 (0.181)
Discussion	-0.035 (0.042)	-0.031 (0.041)	-0.301 (0.303)
Time Fixed Effects	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes
Controls	Basic	Extended	Basic

Note: Robust standard errors clustered by county in parentheses. Each cell represents the results from a separate model. Reduced form estimates display the coefficient on the variable indicating if a respondent lives in a VCA county. TS2SLS estimates display the estimated effect of using a mail-in ballot from the second-stage model. Basic control variables include race, gender, age, party identification, and permanent absentee status. Extended control variables add education to this set of controls. N = 3,419 for rows one and two. N = 2,434 for row three. *p<0.1; **p<0.05; ***p<0.01.

Szewczyk 2018), is the result voters spending more time and using more care filling out their ballots when voting by mail. Indeed, the survey of voters in California suggests that vote-by-mail causes an increase in the amount of time spent gather information.

I now examine if this increase in the time spent conducting research about the election translates into an increase in the amount of knowledge that voters have about the election. Table 5 reveals the estimated effects of the VCA on voter knowledge. In the first row of Table 5 I show the effects of the VCA and using a mail-in ballot on the knowledge index, which is simply the average number of factual questions about the races on the ballot that voters could correctly answer. The reduced form estimates (Columns 1 and 2) are very small, the effect size ranges from 0.003 to 0.006, and imprecisely estimated. As expected, the TS2SLS estimate (Column 3) is slightly larger but statistically insignificant.

Next, let us look at the models examining if respondents were able to answer questions identifying the ideological position and party identification of candidates correctly. These two dependent variables are dummy variables equal to one if a respondent is able to answer all of the questions in each of these categories correctly. The ideology correct variable is constructed using two questions on the ideological position of

the two candidates from governor. The party ID correct variable is constructed using four questions asking which political party each candidate belongs to in the races for lieutenant governor and treasurer. The results indicate that using vote-by-mail does not have an effect on the ability of respondents to recall the ideology and party identification of candidates for statewide executive office. The point estimates from both the reduced form and TS2SLS estimates are negative, substantively small, and very imprecisely estimated.

Table 5: Reduced Form and TS2SLS: The Effects of the VCA and Vote-by-Mail on Voter Knowledge

<i>Dependent Variable:</i>	<i>Reduced Form Estimates:</i>		<i>TS2SLS Estimates:</i>
	VCA County		Use VBM
	(1)	(2)	(3)
Knowledge Index	0.003 (0.012)	0.006 (0.012)	0.029 (0.104)
Party ID Correct	-0.008 (0.027)	-0.006 (0.027)	-0.070 (0.232)
Ideology Correct	-0.017 (0.013)	-0.016 (0.013)	-0.149 (0.116)
Ballot Measure Correct	0.041** (0.021)	0.043** (0.021)	0.355** (0.181)
Knowledge Self-Assessment	0.013 (0.016)	0.016 (0.015)	0.116 (0.135)
Time Fixed Effects	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes
Controls	Basic	Extended	Basic

Note: Robust standard errors clustered by county in parentheses. Each cell represents the results from a separate model. Reduced form estimates display the coefficient on the variable indicating if a respondent lives in a VCA county. TS2SLS estimates display the estimated effect of using a mail-in ballot from the second-stage model. Basic control variables include race, gender, age, party identification, and permanent absentee status. Extended control variables add education to this set of controls. N = 3,419. *p<0.1; **p<0.05; ***p<0.01.

Despite the null effects of the VCA on the overall knowledge index in addition the party ID correct and ideology correct dependent variables, I find that the implementation of the VCA causes voters to be more likely to correctly answer questions about the statewide ballot measures that were voted on during the 2018 general election in California. The ballot measure correct dependent variable is equal to one if respondents correctly answered all of the survey questions about ballot measures that they actually had to vote on during the election. Examining the fourth row of Table 5, the reduced form models reveal that living

in a VCA county causes a 0.041 to 0.043 increase in the probability that respondents answer questions about ballot measures from the 2018 California general election correctly. The TS2SLS estimates indicate that among the group of individuals that were caused to use vote-by-mail by the VCA, using a mail-in ballot increases the likelihood that an individual is able to answer the ballot measure survey questions by 0.355 points. Thus, there is clear evidence that implementing the VCA caused voters in California to be more informed when casting their vote on propositions that appeared on the ballot. Finally, the last set of results in row five of Table 5 indicate that the VCA and using a mail-in ballot does not have an effect on the voters' self-assessment of their knowledge about the election.

Why does vote-by-mail appear to increase knowledge of ballot measure but not the party ID or ideology of candidates? One possible explanation for these findings is that the ballot measures are likely the most complicated issues that respondents are asked about in the survey. There were many measures being voted on during the 2018 general election in California that dealt with policies ranging from rent control to property and transportation taxes. So, it is possible that these more complicated tasks that voters are faced with is where we should expect to find vote-by-mail to have an effect on levels of voter knowledge. Conversely, the questions regarding the ideological placement of candidates asked about the two candidates for governor. As discussed, the two candidates for governor were a Democrat and a Republican so voters may be able to easily distinguish between these two candidates.

4.4 Robustness Checks and Additional Results

In this section I present a series of placebo tests and robustness checks to bolster confidence in the main results. First, I estimate a placebo test that exploits the fact that there was a group of counties that was eligible to adopt the VCA in 2018 but chose not to do so and a group of counties that was not eligible to adopt the VCA in 2018. For this placebo test, I categorize individuals who live in eligible non-adopting counties as being in the treatment group while individuals in counties that were ineligible are in the control group. Individuals in counties that adopted the VCA are excluded from the analysis. If it is actually the implementation of the VCA that is causing changes in voter behavior and not some type of unaccounted for difference between counties that were and were not eligible for the VCA, then these placebo tests should produce null results. The results indicate that this is indeed the case. Tables E.4 and E.5 in the Online Appendix reveal that the coefficients on the variables of interest are close to zero and not statistically significant across all of the dependent variables.

Second, an alternative way to conceptualize the treatment that individuals get in an all-mail election is receiving a mail-in ballot. That is, it is possible the citizens gain information about the election from getting the ballot in the mail even if they choose not to vote. For example, a non-voter could look at the ballot when he receives it in the mail and learn about which parties candidates belong to without ever filling out and returning the ballot. To quantify the magnitude of this alternative treatment, I estimate TS2SLS models where the treatment variable is an indicator for whether or not an individual receives a mail-in ballot. The first-stage models, Table E.1 in the Online Appendix, indicate that the VCA causes an 0.318 increase in the probability that a registered voter receives a mail-in ballot. The results of the TS2SLS models are displayed in Tables E.2 and E.3. I find that the magnitude of the effect on the outcomes of interest are substantively smaller than the main results but they maintain similar levels of statistical significance. This also rules out an alternative explanation that the effects of vote-by-mail are simply due to the fact that the reform causes an increase in voter turnout (Gerber, Huber and Hill 2013). Shineman (2016) finds that mobilizing citizens to vote results in them gathering information about the races in the election. The fact that receiving a ballot causes increases in time spent researching and knowledge of ballot measures indicate that the mobilization effects of vote-by-mail are not exclusively driving these results.

Third, I expect that the effects of vote-by-mail on voter behavior and knowledge will be driven by individuals who are less enthusiastic about politics than highly engaged voters. This is because individuals who care a lot about an election are likely to be motivated to gather political information regardless of whether or not they receive a ballot in the mail. In Figure E.1 in the Online Appendix I display the results of the reduced form models with respondents split into quartiles based on their level of enthusiasm and interest in the election. The political interest measure comes from TargetSmart, so this analysis is restricted to the subset of respondents who are matched to the TargetSmart voter file. I find evidence that the effects of the VCA are driven by low enthusiasm voters. For example, for the time spent researching, ballot measure correct, and knowledge self-assessment dependent variables, the estimates for respondents in the lowest quartile of political interest are substantively larger than the estimates for respondents in the other quartiles and statistically significant ($p < 0.05$). Further, for this group of voters the effect of the VCA on the knowledge index is marginally significant ($p < 0.1$). I discuss additional details of this analysis in Section E.5 in the Online Appendix.

Fourth, I construct post-stratification weights for the survey data to make the sample more representative of the California population as a whole. There are substantial differences in the demographic characteristics

of the population compared to the survey sample, as illustrated in Table E.7, so one may be concerned that the estimated effects do not represent the effects among the population. However, I find that the reduced form estimates with the weighted survey data (Tables E.8 and E.9) are substantively similar to the main reduced form estimates, though the results are more statistically uncertain. In Section E.4 of the Online Appendix I describe the weighting procedure and discuss the results in more detail.

Finally, I also estimate the effects of the VCA on ballot roll-off using county-level administrative data from the California Secretary of State's Office. If voters are spending more time with their ballots and gathering more information about politics than they would when voting about politics, then there should also be a decrease in roll-off for lower salience races on the ballot. Table E.6 in the Online Appendix displays the effects of the VCA on ballot roll-off at the county level for the following races: ballot measures, governor, lieutenant governor, and treasurer. I find that the VCA causes a 0.794 percentage point decrease in ballot roll-off on statewide ballot measures. This effect is even larger after accounting for the percentage of voters in each county that cast their ballots by mail prior to the implementation of the VCA. On the other hand, it appears that the VCA did not have a statistically significant decrease in roll-off for the races for governor, lieutenant governor, and treasurer. This is consistent with the results of the survey in which I find that the VCA only increased knowledge about statewide ballot measures. It is important note, however, that the sample size for the estimates with these races is smaller than that for the statewide ballot measures because elections for statewide executive office only occur every four years while ballot measures typically appear in each even-year general election.

5 Conclusion

In this paper, I examine the effects of vote-by-mail on voter knowledge and how voters gather information about politics. In the first empirical section of the paper, I find that all-mail elections cause a decrease in the percentage of straight ticket votes cast in Utah. Then, I present the results of a study that exploits a policy intervention in California which resulted in five counties in the state implementing a new electoral system in which all voters in the county are mailed a ballot to their residence. I field an original survey that measures how citizens' knowledge and behavior changes over the election cycle. I find that vote-by-mail causes an increase in voter knowledge and an increase in some behaviors that suggest voters in VCA counties are gathering more information about politics.

These findings expand our understanding about vote-by-mail. Previous research on the reform tends to focus on how the electoral reform affects voter turnout (e.g., Magleby 1987, Southwell and Burchett 2000, Gerber, Huber and Hill 2013). Further, although some have suggested that the usage of mail-in ballots may also cause voters to become more informed (Marble 2018, Szewczyk 2018), scholars have been unable to estimate the effects of vote-by-mail on voter knowledge. I contribute to our knowledge on vote-by-mail by showing that the institutions does indeed result in voters gathering more information about politics than they would when voting at a polling place.

References

- Abramowitz, Alan I and Steven Webster. 2016. "The Rise of Negative Partisanship and the Nationalization of US Elections in the 21st Century." *Electoral Studies* 41:12–22.
- Angrist, Joshua D and Alan B Krueger. 1992. "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples." *Journal of the American statistical Association* 87(418):328–336.
- Angrist, Joshua D and Alan B Krueger. 1995. "Split-Sample Instrumental Variables Estimates of the Return to Schooling." *Journal of Business & Economic Statistics* 13(2):225–235.
- Barnes, Tiffany D, Carolina Tchintian and Santiago Alles. 2017. "Assessing Ballot Structure and Split Ticket Voting: Evidence from a Quasi-Experiment." *The Journal of Politics* 79(2):439–456.
- Bartels, Larry M. 1996. "Uninformed votes: Information Effects in Presidential Elections." *American Journal of Political Science* pp. 194–230.
- Beck, Paul Allen, Lawrence Baum, Aage R Clausen and Charles E Smith. 1992. "Patterns and Sources of Ticket Splitting in Subpresidential Voting." *American Political Science Review* 86(4):916–928.
- Binder, Michael, Matthew Childers, Andrew Hopkins and Colleen Hampsey. 2016. "In Voters Minds, Are All Politics Really Local? Comparing Voters Knowledge of National and Local Politics." *Working Paper* .
- Borjas, George J. 2004. "Food Insecurity and Public Assistance." *Journal of Public Economics* 88(7-8):1421–1443.
- Carsey, Thomas M and Geoffrey C Layman. 2004. "Policy Balancing and Preferences for Party Control of Government." *Political Research Quarterly* 57(4):541–550.
- Crowder-Meyer, Melody, Shana Kushner Gadarian and Jessica Trounstine. 2019. "Voting Can Be Hard, Information Helps." *Urban Affairs Review* .
- Crowder-Meyer, Melody, Shana Kushner Gadarian, Jessica Trounstine and Kau Vue. 2018. "A Different Kind of Disadvantage: Candidate Race, Cognitive Complexity, and Voter Choice." *Political Behavior* pp. 1–22.

- Darcy, Robert and Anne Schneider. 1989. "Confusing Ballots, Roll-Off, and the Black Vote." *Western Political Quarterly* 42(3):347–364.
- DeBell, Matthew and Jon A Krosnick. 2009. "Computing Weights for American National Election Study Survey Data." *Ann Arbor, MI, and Palo Alto, CA* .
- Dee, Thomas S and William N Evans. 2003. "Teen Drinking and Educational Attainment: Evidence From Two-Sample Instrumental Variables Estimates." *Journal of Labor Economics* 21(1):178–209.
- Dee, Thomas Sean and William N Evans. 1997. *Teen Drinking and Educational Attainment: Evidence From Two-Sample Instrumental variables (TSIV) estimates*. Vol. 6082 National Bureau of Economic Research.
- DellaVigna, Stefano, John A List, Ulrike Malmendier and Gautam Rao. 2017. "Voting to Tell Others." *The Review of Economic Studies* 84(1):143–181.
- Delli Carpini, Michael X and Scott Keeter. 1996. *What Americans Know About Politics and Why it Matters*. Yale University Press.
- Downs, Anthony. 1957. "An Economic Theory of Political Action in a Democracy." *Journal of Political Economy* 65(2):135–150.
- Eveland, Jr., William P. 2004. "The Effect of Political Discussion In Producing Informed Citizens: The Roles of Information, Motivation, and Elaboration." *Political Communication* 21(2):177–193.
- Fiorina, Morris P. 1996. *Divided Government*. Allyn and Bacon.
- Fowler, Anthony and Michele Margolis. 2014. "The Political Consequences of Uninformed Voters." *Electoral Studies* 34:100–110.
- Freeder, Sean, Gabriel S Lenz and Shad Turney. 2019. "The Importance of Knowing "What Goes with What": Reinterpreting the Evidence on Policy Attitude Stability." *The Journal of Politics* 81(1):274–290.
- Funk, Patricia. 2010. "Social Incentives and Voter Turnout: Evidence from the Swiss Mail Ballot System." *Journal of the European Economic Association* 8(5):1077–1103.
- Gerber, Alan S., Gregory A. Huber and Seth J. Hill. 2013. "Identifying the Effect of All-Mail Elections on Turnout: Staggered Reform in the Evergreen State." *Political Science Research and Methods* 1(01):91–116.

- Hanmer, Michael J and Michael W Traugott. 2004. "The Impact of Voting by Mail on Voter Behavior." *American Politics Research* 32(4):375–405.
- Imai, Kosuke and Kabir Khanna. 2016. "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records." *Political Analysis* 24(2):263–272.
- Inoue, Atsushi and Gary Solon. 2010. "Two-Sample Instrumental Variables Estimators." *The Review of Economics and Statistics* 92(3):557–561.
- Jones, Todd R and Joseph Price. 2017. "Information and the Beauty Premium in Political Elections." *Contemporary Economic Policy* 35(4):677–683.
- Kilgore, Ed. 2018. "The Joy of Voting by Mail." *New York Magazine* .
URL: <http://nymag.com/intelligencer/2018/05/voting-by-mail-is-growing-and-should-grow-more.html?gtm=bottom; m=top>
- Klein, David and Lawrence Baum. 2001. "Ballot Information and Voting Decisions in Judicial Elections." *Political Research Quarterly* 54(4):709–728.
- Krebs, Timothy B. 1998. "The Determinants of Candidates' Vote Share and the Advantages of Incumbency in City Council Elections." *American Journal of Political Science* 42:921–935.
- Lau, Richard R, David J Andersen and David P Redlawsk. 2008. "An Exploration of Correct Voting in Recent US Presidential Elections." *American Journal of Political Science* 52(2):395–411.
- Lupia, Arthur. 1994. "Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." *American Political Science Review* 88(01):63–76.
- Magleby, David B. 1987. "Participation in Mail Ballot Elections." *The Western Political Quarterly* 40(1):79–91.
- Marble, William. 2018. "Mail Voting Reduces Ballot Roll-Off: Evidence from Washington State." *Working Paper* .
- Martinelli, César. 2006. "Would Rational Voters Acquire Costly Information?" *Journal of Economic Theory* 129(1):225–251.

- Matson, Marsha and Terri Susan Fine. 2006. "Gender, Ethnicity, and Ballot Information: Ballot Cues in Low-Information Elections." *State Politics & Policy Quarterly* 6(1):49–72.
- Meredith, Marc and Neil Malhotra. 2011. "Convenience Voting Can Affect Election Outcomes." *Election Law Journal* 10(3):227–253.
- Moskowitz, Daniel J. 2018. Local News, Information, and the Nationalization of US Elections. Technical report Working Paper, <https://bit.ly/2W15rjp>.
- Mummolo, Jonathan and Erik Peterson. 2017. "How Content Preferences Limit the Reach of Voting Aids." *American Politics Research* 45(2):159–185.
- Pacini, David and Frank Windmeijer. 2016. "Robust Inference for the Two-Sample 2SLS Estimator." *Economics letters* 146:50–54.
- Prior, Markus. 2007. *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. Cambridge University Press.
- Reynolds, John F and Richard L McCormick. 1986. "Outlawing "Treachery": Split Tickets and Ballot Laws in New York and New Jersey, 1880-1910." *The Journal of American History* 72(4):835–858.
- Riquelmy, Alan. 2018. "Nevada County Supervisors Approve Voters Choice Act, Ushering in New Way to Vote." *The Union* .
URL: <https://www.theunion.com/news/local-news/nevada-county-supervisors-approve-voters-choice-act-ushering-in-new-way-to-vote/>
- Schaffner, Brian F., Matthew Streb and Gerald Wright. 2001. "Teams Without Uniforms: The Nonpartisan Ballot in State and Local Elections." *Political Research Quarterly* 54(1):7–30.
- Shineman, Victoria Anne. 2016. "If You Mobilize Them, They Will Become Informed: Experimental Evidence that Information Acquisition is Endogenous to Costs and Incentives to Participate." *British Journal of Political Science* pp. 1–23.
- Snyder, James M. and David Strömberg. 2010. "Press Coverage and Political Accountability." *Journal of Political Economy* 118(2):355–408.

- Southwell, Priscilla L. and Justin I. Burchett. 2000. "The effect of all-mail elections on voter turnout." *American Politics Research* 28(1):72–79.
- Szewczyk, James. 2018. "How Electoral Institutions Affect Political Accountability: Evidence from All-Mail Elections." *Working Paper* .
- Thaler, Richard H. and Cass R. Sunstein. 2009. *Improving Decisions About Health, Wealth, and Happiness*. Penguin Books.
- Wildermuth, John. 2017. "Partisan Rift Opens Over Vote-by-Mail Law." *San Francisco Chronicle* .
URL: <https://www.sfchronicle.com/politics/article/Partisan-rift-opens-over-vote-by-mail-law-11290232.php>
- Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge university press.

Online Appendix

A Survey Questionnaire

The following questionnaire is from the first wave of the survey. For the second wave, verb tenses and wording were slightly changed to reflect the fact that the survey was in the field *after* the election took place.

Question 1: How liberal or conservative do you think Gavin Newsom, candidate for governor is, on a scale from 1-7?

Response categories labeled as follows: 1 = Extremely Liberal; 2; 3; 4 Moderate; 5; 6; 7 = Extremely Conservative; I Don't Know

Question 2: How liberal or conservative do you think John Cox, candidate for governor is, on a scale from 1-7?

Response categories labeled as follows: 1 = Extremely Liberal; 2; 3; 4 Moderate; 5; 6; 7 = Extremely Conservative; I Don't Know

Question 3: Eleni Kounalakis is a candidate for lieutenant governor of California. Which political party do you think she is a member of?

Response categories labeled as follows: Democratic; Republican; Green; Libertarian; No party preference; I Don't Know

Question 4: Ed Hernandez is a candidate for lieutenant governor of California. Which political party do you think he is a member of?

Response categories labeled as follows: Democratic; Republican; Green; Libertarian; No party preference; I Don't Know

Question 5: Greg Conlon is a candidate for treasurer of California. Which political party do you think he is a member of?

Response categories labeled as follows: Democratic; Republican; Green; Libertarian; No party preference; I Don't Know

Question 6: Fiona Ma is a candidate for treasurer of California. Which political party do you think she is a member of?

Response categories labeled as follows: Democratic; Republican; Green; Libertarian; No party preference; I Don't Know

Question 7: Proposition 5 is a ballot measure that will change the requirements for certain property owners to transfer their property tax base to a new property. If this measure passes, will it increase or decrease property tax revenue for school and local governments?

Response categories labeled as follows: Increase revenue; Decrease revenue; I Don't Know

Question 8: Proposition 6 proposes to repeal a 2017 transportation tax law. If this measure passes, will there be an increase or decrease in fuel taxes in California?

Response categories labeled as follows: Increase in fuel taxes; Decrease in fuel taxes; I Don't Know

Question 9: Proposition 9, a ballot measure that proposes to divide California into three states, was removed from the ballot for the 2018 general election in California. Who decided to remove the measure?

Response categories labeled as follows: The California state legislature; The California Supreme Court; The Governor of California; The United States Supreme Court; I Don't Know

Question 10: Proposition 10 concerns rent control policies put in place by local governments. If this measure passes, will it increase or decrease the power of local governments to enact rent control policies?
Response categories labeled as follows: Increase power to enact rent control policies; Decrease power to enact rent control policies; I Don't Know

Question 11: Proposition 11 pertains to breaks from work for private sector employees. If this measure passes, will private sector emergency ambulance employees have to remain on-call during work breaks?
Response categories labeled as follows: Yes; No; I Don't Know

Question 12: For the races for statewide executive office on the ballot during the 2018 general election in California, how informed do you feel about the candidates and issues?
Response categories labeled as follows: Very uninformed; Moderately uninformed; Somewhat uninformed; Somewhat informed; Moderately informed; Very informed

Question 13: For the statewide ballot measures being voted on during the 2018 general election in California, how informed do you feel about the issues?
Response categories labeled as follows: Very uninformed; Moderately uninformed; Somewhat uninformed; Somewhat informed; Moderately informed; Very informed

Question 14: In the past month, about how much time do you think that you have spent doing research about the races on the ballot for the 2018 California general election?
Response categories labeled as follows: Fewer than 5 minutes; 5 minutes - 1 hour; 1 - 2 hours; 3 - 4 hours; More than 4 hours

Question 15: In the past month, have you used the Official California Voter Information Guide to gather information about the races on the ballot for the 2018 California general election?
Response categories labeled as follows: Yes; No

Question 16: In the past month, have you used an online search engine, like Google, to gather information about the races on the ballot for the 2018 California general election?
Response categories labeled as follows: Yes; No

Question 17: In the past month, have you used local newspapers, either online or in print, to gather information about races on the ballot for the 2018 California general election?
Response categories labeled as follows: Yes; No

Question 18: In the past month, have you discussed the 2018 California general election with anyone that you live with or do you live alone?
Response categories labeled as follows: Yes; No; I Live Alone

B Summary Statistics

Table B.1: Summary Statistics of Utah Data

Statistic	N	Mean	St. Dev.	Min	Max
Percent Straight Ticket Votes	174	29.919	7.803	0.000	51.022
All-Mail Election	174	0.351	0.479	0	1
Total Population	174	97,018.300	213,180.200	702	1,106,700
Percent Black	174	0.488	0.439	0.000	2.047
Percent Native American	174	2.955	8.754	0.000	49.178
Percent Asian	174	0.793	0.758	0.000	3.890
Percent Hispanic	174	8.354	3.911	0.550	17.911
Percent High School Graduate	174	29.478	6.476	14.467	43.320
Percent College Graduate	174	16.228	5.631	7.085	32.748
Percent Female	174	49.059	1.533	39.459	52.452
Median Age	174	33.430	5.351	24.300	53.100

Table B.2: Summary Statistics of Survey Data

Statistic	N	Mean	St. Dev.	Min	Max
Time Spent Researching	3,419	0.306	0.307	0.000	1.000
Information Sources	3,419	0.613	0.487	0	1
Discussion	2,434	0.810	0.393	0	1
Knowledge Index	3,419	0.523	0.273	0.000	1.000
Ballot Measure Correct	3,419	0.180	0.384	0	1
Party ID Correct	3,419	0.167	0.373	0	1
Ideology Correct	3,419	0.737	0.440	0	1
Knowledge Self-Assessment	3,419	0.357	0.306	0.000	1.000
VCA County*Time	3,419	0.152	0.359	0	1
Permanent Absentee Voter	3,419	0.306	0.461	0	1
Democrat	3,419	0.374	0.484	0	1
Republican	3,419	0.249	0.432	0	1
Independent	3,419	0.300	0.458	0	1
Female	3,419	0.464	0.499	0	1
Male	3,419	0.527	0.499	0	1
Age	3,419	0.640	0.480	0	1
White	3,419	0.040	0.196	0	1
Black	3,419	0.190	0.392	0	1
Hispanic	3,419	0.029	0.167	0	1
Asian	3,419	0.009	0.093	0	1
Some High School	3,419	0.058	0.233	0	1
High School Degree	3,419	0.335	0.472	0	1
Some College	3,419	0.330	0.470	0	1
Bachelor's Degree	3,419	0.255	0.436	0	1

Table B.3: Summary Statistics of Voter File Data

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Use VBM	30,157,797	0.429	0.495	0	0	1	1
Receive VBM Ballot	30,157,797	0.601	0.490	0	0	1	1
VCA County*Time	30,157,797	0.002	0.042	0	0	0	1
Permanent Absentee Voter	30,157,797	0.600	0.490	0	0	1	1
Democrat	30,157,797	0.441	0.496	0	0	1	1
Republican	30,157,797	0.249	0.433	0	0	0	1
Independent	30,157,797	0.262	0.440	0	0	1	1
Female	30,157,797	0.172	0.377	0	0	0	1
Male	30,157,797	0.149	0.356	0	0	0	1
Age	30,157,797	49.842	18.869	16	34	64	257
White	30,157,797	0.588	0.492	0	0	1	1
Black	30,157,797	0.059	0.236	0	0	0	1
Hispanic	30,157,797	0.261	0.439	0	0	1	1
Asian	30,157,797	0.092	0.288	0	0	0	1

C The Transition to Vote-by-Mail in Utah

Table C.1: Year of First All-Mail Election in Utah Counties

County	Year of Switch
Beaver	2014
Box Elder	2018
Cache	2014
Carbon	
Daggett	2018
Davis	2014
Duchesne	2012
Emery	
Garfield	2014
Grand	2014
Iron	2016
Juab	2016
Kane	2016
Millard	2016
Morgan	2016
Piute	2018
Rich	2016
Salt Lake	2016
San Juan	2014
Sanpete	2014
Sevier	2014
Summit	2016
Tooele	2016
Uintah	2016
Utah	2018
Wasatch	2016
Washington	2016
Wayne	2014
Weber	2016

Notes: Data obtained from Utah Lieutenant Governor's Office. Empty cells indicate that county has not yet held all-mail elections.

D Utah Robustness Checks

Table D.1: The Effects of Vote-by-Mail on Straight Ticket Voting in Utah, Leads and Lags Models

	<i>Dependent variable:</i>	
	Percent Straight Ticket Votes	
	(1)	(2)
All-Mail Switch $t + 2$	0.053 (1.367)	0.183 (1.359)
All-Mail Switch $t + 1$	2.171 (1.920)	2.054 (2.016)
All-Mail Switch t	-3.501 (2.286)	-3.752 (2.428)
All-Mail Switch $t - 1$	-2.325 (2.560)	-2.541 (2.553)
All Mail Switch $t - 2$	-2.782 (3.607)	-3.255 (3.657)
County Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Controls	No	Yes
Observations	174	174
Adjusted R ²	0.651	0.663

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors clustered by county.

E California Robustness Checks and Additional Results

E.1 TS2SLS Estimates of the Effects of Receiving a VBM Ballot

Table E.1: First-Stage: The Effects of the VCA on Receiving a VBM Ballot

	<i>Dependent variable:</i>	
	Receive VBM Ballot	
	(1)	(2)
VCA	0.310*** (0.009)	0.318*** (0.00002)
County Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Controls	No	Yes
Observations	30,157,797	30,157,797
Adjusted R ²	0.053	0.998

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors clustered by county.

Table E.2: TS2SLS: The Effects of Receiving VBM Ballot on Voter Behavior

<i>Dependent Variable:</i>	<i>TS2SLS Estimates:</i>
	Recieve VBM Ballot
	(1)
Time Spent Researching	0.068* (0.040)
Information Sources	0.088 (0.065)
Discussion	-0.109 (0.109)
Time Fixed Effects	Yes
County Fixed Effects	Yes
Controls	Basic

Note: Robust standard errors clustered by county in parentheses. Each cell represents the results from a separate model. Reduced form estimates display the coefficient on the variable indicating if a respondent lives in a VCA county. TS2SLS estimates display the estimated effect of using a mail-in ballot from the second-stage model. Basic control variables include race, gender, age, party identification, and permanent absentee status. Extended control variables add education to this set of controls. N = 3,419 for rows one and two. N = 2,434 for row three. *p<0.1; **p<0.05; ***p<0.01.

Table E.3: TS2SLS: The Effects of Receiving VBM Ballot on Voter Knowledge

<i>Dependent Variable:</i>	<i>TS2SLS Estimates:</i> Recieve VBM Ballot (1)
Knowledge Index	0.011 (0.038)
Party ID Correct	-0.025 (0.084)
Ideology Correct	-0.054 (0.042)
Ballot Measure Correct	0.128** (0.064)
Knowledge Self-Assessment	0.042 (0.049)
Time Fixed Effects	
County Fixed Effects	Yes
Controls	Basic

Note: Robust standard errors clustered by county in parentheses. Each cell represents the results from a separate model. Reduced form estimates display the coefficient on the variable indicating if a respondent lives in a VCA county. TS2SLS estimates display the estimated effect of using a mail-in ballot from the second-stage model. Basic control variables include race, gender, age, party identification, and permanent absentee status. Extended control variables add education to this set of controls. N = 3,419. *p<0.1; **p<0.05; ***p<0.01.

E.2 Placebo Test: Non-Adopting Eligible Counties v. Ineligible counties

Table E.4: Reduced Form: The Effects of the VCA and Vote-by-Mail on Voter Behavior (Placebo Test)

<i>Dependent Variable:</i>	<i>Reduced Form Estimates:</i>	
	VCA County	
	(1)	(2)
Time Spent Researching	-0.007 (0.034)	-0.009 (0.037)
Information Sources	0.026 (0.053)	0.023 (0.056)
Discussion	0.006 (0.038)	0.005 (0.034)
Time Fixed Effects	Yes	Yes
County Fixed Effects	Yes	Yes
Controls	Basic	Extended

Note: Robust standard errors clustered by county in parentheses. Each cell represents the results from a separate model. Reduced form estimates display the coefficient on the variable indicating if a respondent lives in a VCA county. Basic control variables include race, gender, age, party identification, and permanent absentee status. Extended control variables add education to this set of controls. N = 2,348 for rows one and two. N = 1,685 for row three. *p<0.1; **p<0.05; ***p<0.01.

Table E.5: Reduced Form: The Effects of the VCA and Vote-by-Mail on Voter Knowledge (Placebo Test)

<i>Dependent Variable:</i>	<i>Reduced Form Estimates:</i>	
	VCA County	
	(1)	(2)
Knowledge Index	-0.013 (0.017)	-0.013 (0.020)
Party ID Correct	0.008 (0.040)	0.005 (0.041)
Ideology Correct	0.031 (0.030)	0.030 (0.030)
Ballot Measure Correct	-0.004 (0.065)	-0.006 (0.066)
Knowledge Self-Assessment	-0.003 (0.025)	-0.004 (0.027)
Time Fixed Effects	Yes	Yes
County Fixed Effects	Yes	Yes
Controls	Basic	Extended

Note: Robust standard errors clustered by county in parentheses. Each cell represents the results from a separate model. Reduced form estimates display the coefficient on the variable indicating if a respondent lives in a VCA county. Basic control variables include race, gender, age, party identification, and permanent absentee status. Extended control variables add education to this set of controls. N = 2,348. *p<0.1; **p<0.05; ***p<0.01.

E.3 The Effects of the VCA on Ballot Roll-Off

Table E.6: The Effects of the VCA on Ballot Roll-Off

	<i>Dependent variable:</i>							
	Ballot Measures		Ballot Roll-Off				Treasurer	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VCA County	-0.794*** (0.185)	-1.595** (0.715)	0.102 (0.213)	0.803 (1.141)	-3.851 (2.642)	1.678 (13.241)	-0.746 (0.513)	-0.154 (2.464)
VCA County* Percent Mail-In Ballots (Lagged)		0.011 (0.008)		-0.009 (0.013)		-0.075 (0.152)		-0.008 (0.035)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	406	406	232	232	232	232	232	232
Adjusted R ²	0.732	0.731	0.369	0.365	0.853	0.852	0.653	0.651

Note:

*p<0.1; **p<0.05; ***p<0.01
Robust standard errors clustered by county.

E.4 Survey Weights

In this paper, I am interested in drawing inferences about the California population as a whole. Because the sample from Civiqs is an opt-in online convenience sample one may be concerned that the sample is not representative of the residents of California. There is evidence that this is the case. In Table E.7 I show the percentage of individuals in the California population and the percentage of survey respondents who belong to a variety of demographic groups. The difference between these two percentages is displayed in the fourth column of Table E.7. If the difference is negative it indicates that there are more members of that demographic group in the population while if it is positive it means there are more survey respondents that identify as being a part of that group.

This exercise highlights some significant differences between the California population and the survey sample. The survey sample is older, more educated, and more likely to live in a county that switched to the VCA in 2018. There are 19.40% fewer individuals in the 35 to 64 age range and 17.31% fewer individuals older than 65 in the population compared to the respondents. Further, among respondents whose highest level of education is a high school degree or lower, there is a 4.93% difference between the population and the sample. Notably, there are nearly 25% more respondents from VCA counties in the survey sample

compared to California as a whole. This is due to an oversampling of individuals who live in counties that switched to the VCA in 2018 in order to ensure that I could estimate the effects of the reform.

Table E.7: Comparison of Survey Sample and California Population

Variable	Population	Survey (unweighted)	Survey (weighted)	Difference (unweighted)	Difference (weighted)
35-64	38.43	57.83	53.87	19.40	15.45
Age 65 and Over	13.21	30.51	19.86	17.31	6.66
Age Under 35	24.99	11.66	26.27	-13.33	1.28
Asian	14.12	2.84	8.84	-11.28	-5.28
Black	5.81	4.06	5.93	-1.74	0.13
Mixed Race	4.71	5.05	5.28	0.34	0.58
Native American or Hawaiian	0.75	1.37	0.90	0.63	0.15
Other Race	13.67	3.08	10.04	-10.59	-3.63
White	60.56	83.59	69.00	23.03	8.45
College	13.58	33.56	28.49	19.99	14.91
High School or Less	11.66	6.72	15.82	-4.93	4.16
Postgraduate	8.11	25.76	17.59	17.65	9.48
Some College	19.52	33.95	38.10	14.43	18.59
Female	50.32	46.95	50.34	-3.37	0.02
Male	49.68	53.05	49.66	3.37	-0.02
Not VCA County	93.19	68.41	88.85	-24.78	-4.35
VCA County	6.81	31.59	11.15	24.78	4.35

To address the potential biases that may result from the demographic differences between the population and survey sample, I calculate survey weights. The first step of this process accounts for the fact individuals had an unequal probability of selection into the sample based on where they lived. In particular, respondents in VCA counties were 4.615 times more likely than respondents in non-VCA counties. Based on this, I first generate a set of design weights that are equal to 0.217 for respondents in VCA counties and 1 for respondents in non-VCA counties.

Next, I calculate post-stratification weights to account for differences between the demographic characteristics of the California population and the Civiqs survey panel. Traditional cell weighting requires that there are observations in the survey sample for each possible combination of the different variables that the researcher wants to post-stratify on. For the variables that I want to post-stratify on - age, race, education, and gender - this is not the case, so I would not to eliminate demographic variables to use cell weighting. Therefore, I use raking to calculate the post-stratification weights, which is an algorithm that gives me the ability to flexibly use all of the available information on the demographic characteristics of the survey respondents.

To implement the raking procedure I use the guidelines outlined in DeBell and Krosnick (2009). Raking is done with the variables indicating respondents' race, gender, age, and level of education. The raked post-stratification weights are then combined with the design weights that take into account that individuals in VCA counties were more likely to appear in the sample. Following DeBell and Krosnick (2009), the weights

are trimmed at five times the mean value of the weights, because larger values are more likely to inflate the variance of estimates.

Table E.8: Reduced Form: The Effects of the VCA on Voter Behavior (Weighted Data)

	<i>Dependent variable:</i>		
	Time Spent Researching	Information Sources	Discussion
	(1)	(2)	(3)
VCA County	0.041 (0.034)	0.013 (0.052)	-0.005 (0.045)
Time Fixed Effects	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes
Extended Controls	Yes	Yes	Yes
Observations	3,346	3,346	2,390

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors clustered by county.

Table E.9: Reduced Form: The Effects of the VCA on Voter Knowledge (Weighted Data)

	<i>Dependent variable:</i>				
	Knowledge Index	Party ID Correct	Ideology Correct	Ballot Measure Correct	Self-Assessment
	(1)	(2)	(3)	(4)	(5)
VCA County	0.003 (0.029)	0.016 (0.044)	0.013 (0.043)	0.065* (0.036)	-0.007 (0.036)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes
Extended Controls	Yes	Yes	Yes	Yes	Yes
Observations	3,346	3,346	3,346	3,346	3,346

Note: *p<0.1; **p<0.05; ***p<0.01
Robust standard errors clustered by county.

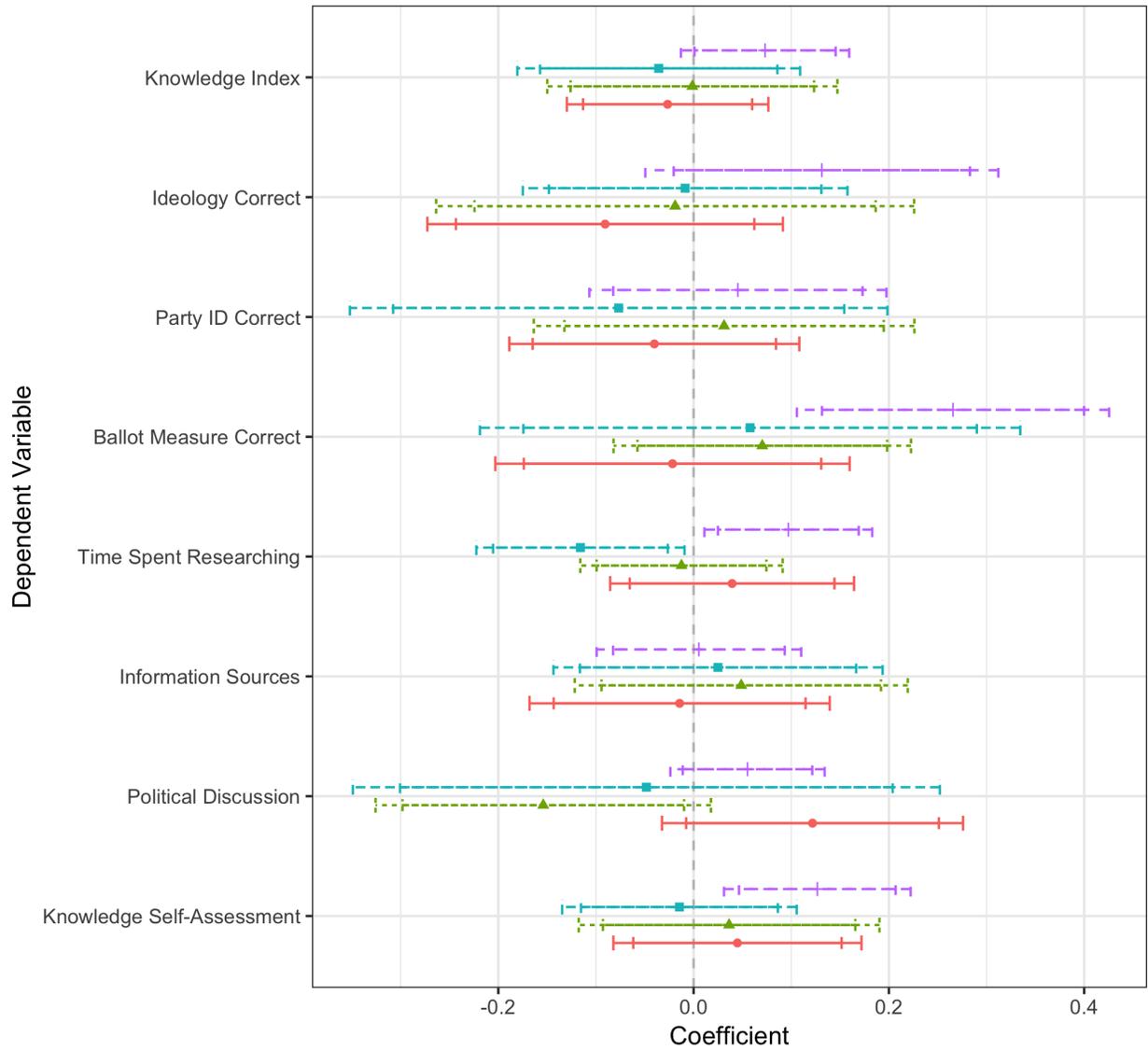
The distributions of the demographic variables in the final weighted data are also displayed in Table E.7. As can be seen, the weights significantly improve the representativeness of the sample from Civiqs. The results of the reduced form models with the weighted data are displayed in Tables E.8 and E.9. For the most robust main results, ballot measure correct and time spent researching, the point estimates with the weighted data are substantively very similar to that for the unweighted data though they are statistically more uncertain. The estimated effect of the VCA on time spent researching is no longer significant with the weighted data though the effect is marginally significant when ballot measure correct is the dependent variable.

E.5 Who is Affected by Vote-by-Mail?

I propose that the effects of vote-by-mail on voter knowledge and behavior will be conditioned by a voter's level of political interest. In particular, I expect that the effects of the VCA on voter knowledge will be largest among individuals who are least interested in the 2018 general election while the effects will be smallest for voters who are highly enthusiastic about politics. This is because voters who care a lot about the election will gather information even if they are not mailed a ballot. Indeed, previous research finds that voters in California with high levels of political interest are the most likely to use voter information guides Mummolo and Peterson (2017). Meanwhile, individuals who are not very interested in politics will be affected by the extra nudge that receiving a ballot in the mail gives them.

In order to test this hypothesis, I take advantage of the fact that a subset of respondents in the sample are able to be matched to the TargetSmart voter database. TargetSmart is a company that aggregates state voter files and generates predictions about a variety of characteristics of registered voters. I categorize respondents of my survey based on their enthusiasm for the 2018 general election, which is derived from one of TargetSmart's proprietary models. The measure is based on turnout in previous elections and survey data measuring individuals' interest in the 2018 midterm election. I split respondents into quartiles based on their enthusiasm scores and estimate the effects of the VCA on the outcome variables separately. The lowest quartile represents individuals that were not very interested in the 2018 general election. The two middle quartiles include voters that have medium levels of interest. Finally, individuals who were very enthusiastic about the election are in the top quartile. It is important to note that for these analyses, the data is restricted to the 1,386 respondents that I am able to observe in the TargetSmart data, so the sample size is much smaller than that for the main set of results discussed in the previous section.

The results of the reduced form models with the data split into quartiles based on respondents' enthusiasm about the 2018 general election are displayed in Figure E.1. The Figure shows the regression coefficient on the causal variable of interest, a dummy variable equal to one for individuals who live in a VCA county in the second wave of the survey. For each dependent variable, the first estimate that is displayed is for respondents in the lowest quintile and the second estimate is for the second lowest quintile, which are the low enthusiasm and low-medium enthusiasm groups. Meanwhile, the results for respondents in the medium-high and high enthusiasm groups are the third and fourth estimates. In addition to the point estimates, Figure E.1 reveals 90% and 95% confidence intervals.



Interest — High — Medium-High — Low-Medium — Low Interest in Election — — Low — — Low-Medium

Figure E.1: Reduced form estimates of the effects of the VCA on voter knowledge and behavior across voters' interest in 2018 election. I report 90% and 95% confidence intervals calculated with robust standard errors clustered at the county-level. All models include the extended set of individual-level control variables.

The results are supportive of the predictions of the theory about who should be most affected by vote-by-mail. Across all of the dependent variables, except information sources used and political discussion, the estimated causal effects are largest among those in the lowest quintile of the election interest measure. That being said, the effect of the VCA on the dependent variables indicating they identified the ideological position and party ID of candidates correctly do not achieve traditional levels of statistical significance, even after splitting respondents into quintiles.

Let us now examine the impact of the VCA on the knowledge index. Although the average effect estimated from the reduced form and TS2SLS models in Table 5 was substantively very small, among respondents who are least enthusiastic about the 2018 general election the estimated effect of the VCA on the knowledge index is large (.073) and marginally significant ($p < .1$). Moreover, the impact of vote-by-mail on answering ballot measure questions correctly, the amount of time spent conducting research, and how knowledgeable respondents feel appear to be driven exclusively by changes in the knowledge and behavior of individuals who are in the bottom quintile of the variable measuring respondents' interest in the election. For these dependent variables, the point estimates are largest for the models with those in the low enthusiasm group. Although these effects are only marginally statistically significant ($p < .1$) when time spent researching is the dependent variable, for the other two outcomes the effect is significant at the .05 level.

This set of results provides additional evidence in favor of my argument that vote-by-mail causes voters to gather more information about politics than they would otherwise. The group of individuals that I expect to be most affected by this change, those who are least interested in politics, do indeed drive the impact of using a mail-in ballot on voter knowledge and behavior.