Dual Effects of Vote-by-Mail Elections on Voter Turnout*

Yuki Atsusaka† Robert M. Stein‡
atsusaka@rice.edu stein@rice.edu

Andrew Menger§
andrewmmenger@gmail.com

August 7, 2020

Include Bob's citation

Abstract

Growing interest in Vote-by-Mail (VBM) elections has made empirical evaluations of the mode of voting crucial. However, prior studies lack a theoretical explanation of *why* VBM could increase voter turnout, making the evaluation of the election reform challenging. We propose an explanation based on two potential mechanisms through which VBM may increase voter turnout (information and convenience effects) and test its implications by 1.7 million administrative records from Colorado and North Carolina and New Mexico as control states. A difference-in-differences analysis with exact matching shows that the VBM implementation significantly boosted turnout among frequent voters and old voters, suggesting a large convenience effect, whereas it had a very small effect among infrequent voters, suggesting a negligible information effect. Our results imply that VBM primarily increases turnout among already mobilized voters, while suggesting the importance of theories in the evaluation of election reform.

Keywords: Vote-by-mail; election reform; convenience voting; semiparametric approach

---

*This study is supported by the New Venture Fund. For helpful comments, we thank Vin Arceneaux, Barry Burden, Shiro Kuriwaki, Soichiro Yamauchi, and participants of the panels at MPSA and SPSA.

†Ph.D. Candidate, Department of Political Science, Rice University.
‡Lena Gohlman Fox Professor, Department of Political Science, Rice University.
§Affiliation.

Cluster the SE at the state level

Word Count: 3047

Introduction

Vote-by-Mail (VBM) elections, where every registered voter is mailed a ballot up to three weeks before each Election Day, have been touted as an antidote to various problems in American politics (Vote at Home, 2020). Recently, a body of research has examined VBM, increasingly enriching our knowledge about the magnitude of VBM effects on turnout and their partisan, participatory, and representational implications in Oregon (Southwell and Burchett, 2000; Southwell, 2009; Gronke and Miller, 2012), Washington (Gerber, Huber and Hill, 2013), California (Bergman and Yates, 2011; Kousser and Mullin, 2007), Colorado (PantheonAnalysis, 2017), Utah (PantheonAnalysis, 2018), and multiple states (Gronke, Galanes-Rosenbaum and Miller, 2007; Richey, 2008; Larocca and Klemanski, 2011), and the role of VBM has been further highlighted amid the COVID-19 pandemic (Barber and Holbein, 2020; Thompson et al., 2020; Bonica et al., 2020) (Online Appendix A.2 provides a review of previous findings). Despite such rapidly growing academic and public attention to the mode of voting, prior studies have not yet provided a theoretical explanation of *why* VBM could increase voter turnout (if any), which makes the evaluation of the election reform challenging. This is highly problematic in the contexts of election reform and political science research more broadly because the lack of *theoretical* explanations makes it difficult for researchers to assess whether VBM actually increases turnout among voters at whom the election reform intends to target and probe what to do if it does not.

To resolve this issue and advance the research of VBM and political participation more broadly, this paper proposes two potential mechanisms through which VBM may increase voter turnout. Specifically, we theorize that the total effect of VBM is derived from a *dual effect* of mobilizing uninformed voters by notifying about upcoming elections *and* making voting more convenient for costly voters. Moreover, given the prior knowledge that the information effect via mailing is small (Arceneaux and Nickerson, 2009), the key observable implication is that VBM increases turnout primarily among voters who are already well informed but find it costly to return their ballots to the polling stations on Election Day, such as frequent voters and older voters. In contrast, we expect that the total VBM effect does not vary much by other subpopulations (e.g., by race, party, and gender) assuming that each group has a similar level of informed registered voters.

To test these implications, we use 1.7 million administrative records ("snapshot" official voter history and registration files) from Colorado and North Carolina. A difference-in-differences analysis with exact matching shows that the VBM implementation in Colorado increased its voter turnout by 5.7% points on average. Moreover, we find that the VBM implementation significantly boosted turnout among frequent voters and old voters, suggesting a large convenience effect, whereas it had a very small effect among infrequent voters, suggesting a negligible information effect. Our results imply that VBM primarily increases turnout among already mobilized voters, while suggesting the importance of theories in the evaluation of election reform.

Dual Effects of Notification and Convenience

VBM has become an increasingly popular mode of voting in multiple American states (National Conference of State Legislatures, 2020) (Online Appendix A.1 offers a brief history of VBM). A number of studies has assessed the effects of mail-assisted voting on voter turnout, finding mostly positive effects (e.g., Gerber, Huber and Hill, 2013; Larocca and Klemanski, 2011; Bonica et al., 2020; Thompson et al., 2020; Barber and Holbein, 2020). To understand positive (and absence of such) effects of VBM — either in the entire electorate or in important subpopulations — it is crucial to discern why and in what ways the adoption of VBM could increase voter turnout, although previous research has not offered a comprehensive theory to explain and predict VBM effects (or lack thereof) on turnout. Our research fills this critical void in the literature.

We identify two ways VBM elections enhance voter participation: *notification* about an upcoming election and the *convenience* with which to cast a ballot. The effect of these attributes on voting are conditional on the voters history of registration and prior voting. We then formulate that the total effect of VBM is an additive effect of the information and convenience. To be explicit, our theoretical explanation is as follows:

$$\text{Total VBM Effect} = \text{Notification Effect} + \text{Convenience Effect} \tag{1}$$

The literature on voter mobilization (Gerber and Green 2000; Green and Gerber 2004; Nickerson 2005; 2007; Michelson 2006) **(need article names for these)** links voter turnout to highly personalized messaging from candidates and parties. Dale and Strauss (2009) note, however, that empirical findings on personal and impersonal campaign messaging are not consistent on this point (Michelson 2006; Nickerson 2008; Green and Gerber 2004: 37)**(need article names for these)**. Dale and Strauss' Noticeable Reminder theory argues that impersonalized messaging can be efficacious for turning out registered voters. Registered voters have already signaled their willingness to participate in the political process. What these voters require to vote is "a reminder to make time in their busy schedules to go to the polls" (Dale and Strauss, 2009, 787). Connection to the political system has been cemented through voter registration, especially if registration has occurred closer to the pending election. The act of registering or re-registering to vote made these voters susceptible to reminders to vote without a face-to-face contact or other personal content to the message. Messaging must however be noticeable, unavoidable and proximate to Election Day.

VBM elections appear to increase voter turnout by providing registered voters with a noticeable reminder, that is *unavoidable and proximate* to when they are likely to vote (i.e., complete their mail ballot. These conditions can vary among registered voters in VBM states and provide a means of testing the notification effects of VBM elections on voter participation. Specifically, we expect that recently registered voters to be more susceptible to the turnout effects of receiving and unsolicited mail ballot than voters who have been registered to vote for a longer time.

Receiving a ballot in the mail is obviously less costly and more convenient than traveling, on or before Election Day, to a polling location to cast a ballot. The convenience associated with VBM mitigates the high costs of locating and getting to a polling place on or before Election Day (Brady and McNulty, 2011; Fitzgerald, 2005; Haspel and Knotts, 2005). VBM elections afford voters the opportunity to complete their ballot over several days (or even weeks) and at home or the location of their choosing.[1] These features of VBM elections afford voters significant convenience and reduced costs, resulting in a higher number of cast ballots and completed ballots (Menger, Stein

[1]These features were not available with conventional precinct Election Day voting.

and Vonnahme, 2018).

We would expect this aspect of VBM to be more pivotal in retaining voters who have turned out in the past, and not in turning out many new or infrequent voters. Dale and Strauss (2009) note that persons who have recently registered to vote and those who have also voted in more recent elections have more accurate knowledge of the costs and benefits of voting. Frequent voters, at least those who voted in more recent elections are more likely to be aware of and pickup the convenience of voting by mail.

Having theorizing the above two effects, we also acknowledge that prior studies report that a notification effect from mailing is usually significantly small (less than 1 percentage points), if not absence (Green and Gerber, 2019) **need to edit this citation**. This makes it possible for us to modify Equation (1) into the following form:

$$\text{Total VBM Effect} \approx \text{Convenience Effect} \tag{2}$$

A key observable implication from Equation (2) is that *if* we would observe a positive (total) effect of VBM on turnout among registered voters, it should be among voters who are more susceptible to the convenience effect; namely those who are already well mobilized or/and who have more physical barriers (e.g., time and transportation) in voting. In particular, we expect that the total VBM effect is larger among frequent voters and older voters than among infrequent voters and non-old voters. Given the small expected information effect, we also expect that the total VBM effect does not vary much by other subpopulations, assuming that they have similar levels of voters who would benefit from the reduced cost. Online Appendix A.3 provides additional discussions on the two mechanisms.

Data and Identification Strategy

To test the observable implications, we first define our population of interest as *a set of voters who had been registered in the location before and after the adoption of VBM*. We then define our quantities of interest as a set of *conditional average treatment effects on the treated* (CATTs):

$$\tau_{CATT} = \mathbb{E}_{\mathbf{X}}\left[\mathbb{E}[Y_i^{d=1} - Y_i^{d=0} | D_i = 1, Z_i = z, \mathbf{X_i}]\right], \tag{3}$$

where $Y_i^{d=1}$ and $Y_i^{d=0}$ are potential outcomes (turnouts) for voter i, $D_i = 1$ the treatment status, $Z_i = z$ an effect modifier denoting voter i belongs to a subpopulation z, and $\mathbf{X_i}$ covariates. We assume weak ignorability ($Y^{d=0} \perp\!\!\!\perp D_i | \mathbf{X_i}$) to hold. Substantively, this means that we study VBM effects among voters to whom the election reform is targeted at and we are interested in how much such effects vary by subpopulations (see Online Appendix C.1 for details).

To estimate the CATTs, we use 1.7 million administrative records from the 2012 and 2016 presidential elections in Colorado and North Carolina. Colorado has adopted VBM elections in 2013, whereas North Carolina has not and thus is used as a counterfactual (Keele and Minozzi, 2010). More specifically, we collected and merged *snapshot* voter history files and registration records, coded individual race, imputed missing data, and examined descriptive statistics (see Online Appendix B). This leaves individual-level data in two states before and after the intervention (VBM adoption) occurs.

For our identification strategy, we adopt a (parametric linear) difference-in-differences (DID) model. The key idea is to simulate voter turnout in Colorado had the VBM policy not been adopted and use such counterfactual turnout to estimate the causal effect of the policy adoption. We do so by using the information from the pre-intervention period in Colorado and information from North Carolina. The validity of this approach depends largely on the parallel trends assumption, which states that, in our case, the unobserved confounders that create a systematic difference in turnouts in the two states should be constant over time. For this reason, our identification is valid *even if Colorado and North Carolina voters have very different political cultures* that cannot be adjusted by available covariates as long as such difference stays the same across the two elections.

Although we provide suggestive evidence that the assumption holds in our setting (Online Appendix C.4), this assumption is inherently nonrefutable (cannot be empirically verified with any data) and thus our inference could be jeopardized by the potential violation of the assumption. To mitigate this concern, we employ a semi-parametric approach in which we carry out exact matching prior to the application of the DID model so that we can restrict the set of control units to those with similar sociopolitical attributes to treated units. Specifically, we perform exact matching to preprocess the original data and then implement a DID model by a saturated weighted least squares regression, where weights (for control units) are obtained from exact matching (see also Online Appendix C.4). This semi-parametric approach has been considered in ONeill et al. (2016) and shown to reduce biases that could appear when the parallel trends assumption is violated (Ryan et al., 2019). We further use the idea of *pattern specificity* (Rosenbaum, 2005) to compare our treated units to different control units from New Mexico. The key idea is to make sure that a common pattern emerges regardless of the choice of control units.

Empirical Findings

First, we find that the total effect of VBM on turnout among all Colorado voters (i.e., ATT) is 0.057 (s.e.=0.003). In Online Appendix A.2, we quantified our prior belief on the total VBM effect as $\mathcal{N}(0.041, 0.0009)$. Given our new finding, we update this prior belief by to form a posterior belief on the VBM effect according to Bayes' rule. Figure 1 visualizes our posterior belief along with our prior belief (with findings of previous studies) and our finding. Our posterior mean (and mode) is 0.057 with an 89% credible interval (0.052, 0.062). These results suggest that Colorado voters would have had 5.7% points *lower* turnout in 2016 if (counterfactually) the state *had not adopted* VBM elections in 2013.

Next, Figure 2 displays estimated CATTs with different effect modifiers. Consistent with our expectations, we find that the estimated effects are significantly larger among frequent voters and older voters than among infrequent voters and non-old voters (We coded each voter as a frequent voter if she had voted in the 2010 midterm election and as an infrequent voter otherwise.). Moreover, the graph shows that the effect size does not vary much for other subpopulations. Importantly, we did not find that VBM increases turnout only among Democratic or non-Democratic voters. The

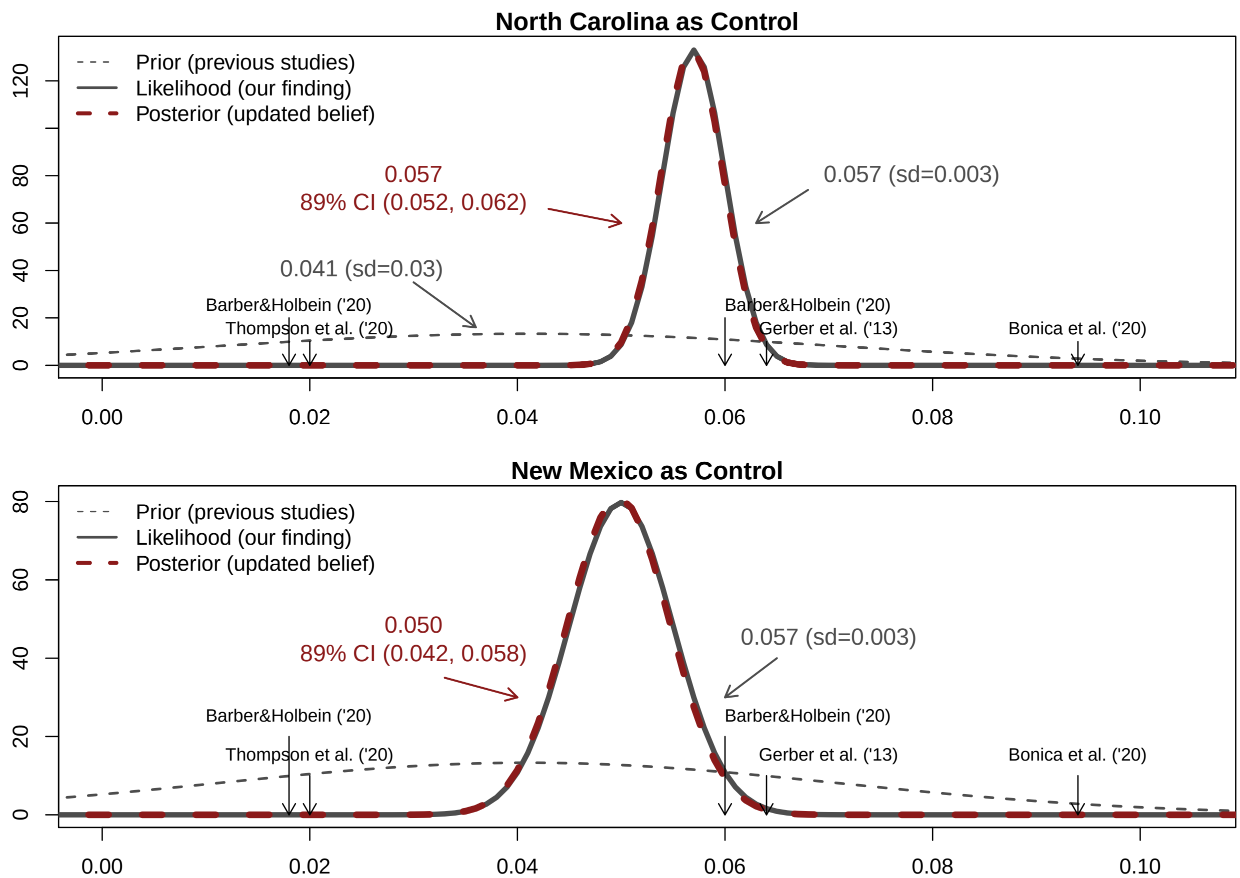


Figure 1: **Posterior Belief on the Average VBM Effect on Turnout.**
Note: Our posterior belief on the total VBM effect when using North Carolina (upper panel) and New Mexico (lower panle) as control states given previous research and our finding.

results imply that most of the (total) VBM effects are stemmed from what we formulate as the convenience effect rather than the information effect. The finding for frequent voters is particularly consequential because it implies that VBM increases turnout among already mobilized voters who are expected to know about upcoming elections.

Robustness Checks (4-5 Sentences)

In our Online Appendix, we demonstrate that our substantive conclusions hold when .... (Table X), ... (Figure Y...)...

Online Appendix D offers additional findings, robustness checks, and circumstantial evidence for the parallel trends assumption, confirming that our substantive conclusion remains the same

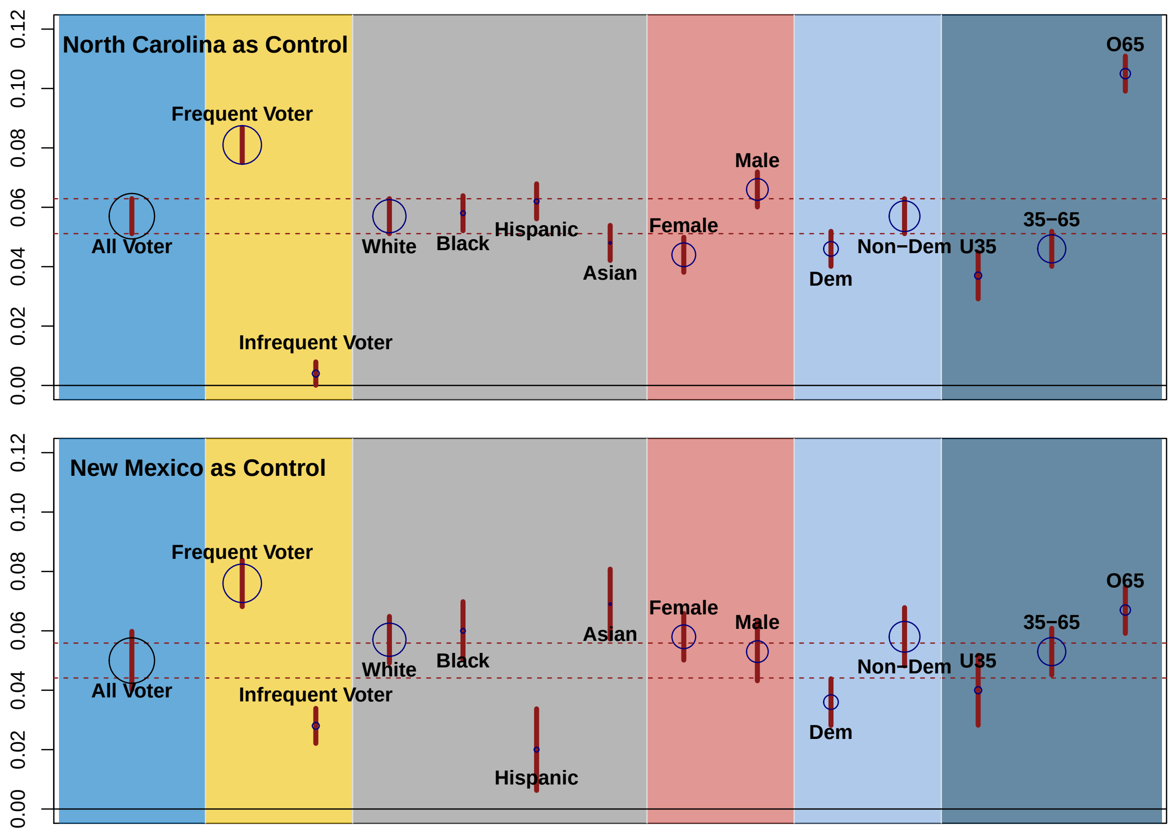


Figure 2: **Estimated Average Effects of VBM on Turnout with Different Effect Modifiers.** *Note*: This figure visualizes the estimated effect on various subpopulations of interest when using North Carolina (upper panel) and New Mexico (lower panel) as control states. For each group, the size of its open circle is proportional to the size of the group.

and is not susceptible to our particular decisions on estimation, inference, and research design. We also show that the VBM did not change the composition of voters who actually turned out from 2012 to 2016 (Online Appendix D.4). Taken together, we find supportive evidence for our theoretical claim: VBM increases turnout by primarily reducing the cost of voting (rather than informing voters of and mobilizing them for upcoming elections).

Implications for Future Research (1-2 Sentences)

This part should be revised, but for now it is included as a space holder: We have identified two mechanisms for how VBM elections effect voter turnout; convenience and information. We find strong circumstantial evidence that the turnout effects from VBM elections are associated with the convenience rather than the mobilizing traits of this mode of voting. Our observation that the convenience rather than the information effects of VBM elections has a more demonstrative effect on voter turnout receives further support from the most recent and methodologically rigorous research on the turnout effects of VBM elections. Barber and Holbein (2020), Thompson et al. (2020), Gerber, Huber and Hill (2013) and Bonica et al. (2020) did not find any partisan effects from VBM elections. Similarly, these authors find that older voters are most advantaged by VBM elections, as are younger voter voters. Only Bonica et al. (2020) and Gerber, Huber and Hill (2013) find evidence, albeit modest, that traditionally underrepresented votes, non-whites received a boost in turnout from VBM elections.

References

Abadie, Alberto. 2005. "Semiparametric difference-in-differences estimators." *The Review of Economic Studies* 72(1):1–19.

Arceneaux, Kevin and David W Nickerson. 2009. "Who is mobilized to vote? A re-analysis of 11 field experiments." *American Journal of Political Science* 53(1):1–16.

Barber, Michael and John. Holbein. 2020. "Voting During the COVID-19 Pandemic: Vote-by-Mail Does Not Advantage Democrats." *Working Paper* .

Bergman, Elizabeth and Philip A Yates. 2011. "Changing election methods: How does mandated vote-by-mail affect individual registrants?" *Election Law Journal* 10(2):115–127.

Bonica, Adam, Jacob M. Grumbach, Charlotte Hill and Hakeem Jefferson. 2020. "All-Mail Voting in Colorado Increases Turnout and Reduces Turnout Inequality." *Working Paper* .

Brady, Henry E and John E McNulty. 2011. "Turning out to vote: The costs of finding and getting to the polling place." *American Political Science Review* 105(1):115–134.

Dale, Allison and Aaron Strauss. 2009. "Don't forget to vote: Text message reminders as a mobilization tool." *American Journal of Political Science* 53(4):787–804.

Fitzgerald, Mary. 2005. "Greater convenience but not greater turnout: The impact of alternative voting methods on electoral participation in the United States." *American Politics Research* 33(6):842–867.

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(1):91–116.

Green, Donald P and Alan S Gerber. 2019. *Get out the vote: How to increase voter turnout*. Brookings Institution Press.

Gronke, Paul, Eva Galanes-Rosenbaum and Peter A Miller. 2007. "Early voting and turnout." *PS: Political Science & Politics* 40(4):639–645.

Gronke, Paul and Peter Miller. 2012. "Voting by mail and turnout in Oregon: Revisiting Southwell and Burchett." *American Politics Research* 40(6):976–997.

Haspel, Moshe and H Gibbs Knotts. 2005. "Location, location, location: Precinct placement and the costs of voting." *Journal of Politics* 67(2):560–573.

Keele, Luke and Rocío Titiunik. 2018. "Geographic Natural Experiments with Interference: The Effect of All-Mail Voting on Turnout in Colorado." *CESifo Economic Studies* 64(2):127–149.

Keele, Luke and William Minozzi. 2010. How much is Minnesota like Wisconsin? States as counterfactuals. In *States as Counterfactuals (August 23, 2010). APSA 2010 Annual Meeting Paper*.

Kousser, Thad and Megan Mullin. 2007. "Does voting by mail increase participation? Using matching to analyze a natural experiment." *Political Analysis* 15(4):428–445.

Larocca, Roger and John S Klemanski. 2011. "US state election reform and turnout in presidential elections." *State Politics & Policy Quarterly* 11(1):76–101.

Menger, Andrew, Robert M Stein and Greg Vonnahme. 2018. "Reducing the Undervote With Vote by Mail." *American Politics Research* 46(6):1039–1064.

National Conference of State Legislatures. 2020. *Vote by Mail*.
URL: *https://www.ncsl.org/research/elections-and-campaigns/all-mail-elections.aspx*

ONeill, Stephen, Noémi Kreif, Richard Grieve, Matthew Sutton and Jasjeet S Sekhon. 2016. "Estimating causal effects: considering three alternatives to difference-in-differences estimation." *Health Services and Outcomes Research Methodology* 16(1-2):1–21.

PantheonAnalysis. 2017. "Utah 2016: Evience for the positive turnout effects of Vote at Home (also known as vote by mail) in participation counties.".

PantheonAnalysis. 2018. "Colorado 2014. Comparisions of Predicted and Actual Turnout.".

Richey, Sean. 2008. "Voting by mail: Turnout and institutional reform in Oregon." *Social Science Quarterly* 89(4):902–915.

Rosenbaum, Paul R. 2005. "Observational study." *Encyclopedia of statistics in behavioral science* .

Ryan, Andrew M, Evangelos Kontopantelis, Ariel Linden and James F Burgess Jr. 2019. "Now trending: Coping with non-parallel trends in difference-in-differences analysis." *Statistical methods in medical research* 28(12):3697–3711.

Southwell, Priscilla L. 2009. "Analysis of the turnout effects of vote by mail elections, 1980–2007." *The Social Science Journal* 46(1):211–217.

Southwell, Priscilla L and Justin I Burchett. 2000. "The effect of all-mail elections on voter turnout." *American Politics Quarterly* 28(1):72–79.

Thompson, Daniel M., Jennifer Wu, Jesse Yoder and Andrew B. Hall. 2020. "All-Mail Voting in Colorado Increases Turnout and Reduces Turnout Inequality." *Working Paper* .

Vote at Home. 2020. *Proven Track Record*.
URL: *https://www.voteathome.org/project/proven-track-record/*

Online Appendix

For "Dual Effects of Vote-by-Mail Elections on Voter Turnout"

Table of Contents

A VBM Elections and Previous Research

A.1 Brief History of VBM Elections

First adopted by Oregon in 1996, Washington and Colorado have also implemented VBM elections statewide for all elections in 2005 and 2013, respectively. In 2014 Utah initiated VBM elections on an optional county-wide basis, and most recently California has introduced optional countywide VBM elections beginning in 2020. Hawaii adopted VBM voting for the 2020 Presidential election. Since 2000 California counties have had the discretion to conduct VBM elections in precincts where there are 250 or fewer registered voters, presumably to save money on the costs of poll workers, voting stations and polling locations. Unlike absentee voters, registered voters in VBM-only precincts are automatically mailed a ballot whether they request one or not with a postage-paid return envelope. These voters do not have the option of voting in person. The only other development is the number of states that have adopted no-excuse absentee voting for the November election. But these adoptions are not permanent and only in response to the COVID-19.

A.2 Previous Findings

To date, a body of research has examined the effect of adopting VBM elections on voter turnout. Table A.1 lists published studies on this topic with the place of interest, the type of elections, data source, and main findings on estimated effects (positive, negative, or null). Overall, previous research has found positive effects of VBM elections on turnout, although the estimated magnitudes vary across studies.

While most research shows a significant and positive effect of VBM on voter turnout, these findings are not fully consistent and some researchers report either no effect on turnout or even a negative effect (Gronke and Miller, 2012; Southwell, 2009; Bergman and Yates, 2011; Kousser and Mullin, 2007; Keele and Titiunik, 2018) The findings of negative effects are restricted to California counties discretionary use of VBM elections, but some studies of VBM in Oregon have reported the absence of a positive effect from the policy adoption.

Importantly, more recent studies (including unpublished works) have studied the VBM effect based on stronger causal identifications than did earlier studies. For example, using a staggered implementation of VBM as a quasi-natural experiment, Gerber, Huber and Hill (2013) find that VBM increases turnout by two to four percentage points. Based on a similar experimental leverage, Thompson et al. (2020) report that VBM elections increase turnout by two to three percentage points, with no evidence of a partisan advantage. More recently, Bonica et al. (2020) show report turnout increases in excess of nine percentage points in Colorado and by larger margins for subpopulations (i.e., younger voters and voters of color. Barber and Holbein (2020) also report significant positive turnout effects from VBM elections.

We incorporate findings from these recent studies as our *prior knowledge* on the sign and magnitude of the total VBM effect. To quantify such prior knowledge, we define that our prior belief on the total effect of VBM elections is a random variable that follows a normal distribution

Authors	Place	Election Type	Data	Effect
Southwell and Burchett (2000)	Oregon	Federal	TS, state level	$+$
Gronke et al. (2007)	All states	Presidential	Panel, state level	$+$
Richey (2008)	All states	Federal	Panel, state level	$+$
Gerber, Huber and Hill (2013)	WA (counties)	Federal	Panel, individual	$+$
Larocca and Klemanski (2011)	All states	Federal	CS, survey	$+$
Gronke and Miller (2012)	Oregon	Fed & Special	TS, state level	null & $+$
Southwell (2009)	Oregon	All	TS, state level	null & $+$
Bergman and Yates (2011)	CA (precincts)	Local	CS, individual	$-$
Kousser and Mullin (2007)	CA (precincts)	Federal	CS, individual	$-$
Sled (2008)	8 states	All	Panel, local level	$+$
PantheonAnalysis (2017)	Colorado	Federal	CS, individual	$+$
PantheonAnalysis (2018)	Utah	Federal	CS, individual	$+$

Table A.1: Previous reserch on the effects of VBM on turnout
Note: WA = Washington, CA = California, TS= Time Series, CS = Cross Sectional, Gronke et al (2007) = Gronke, Galanes-Rosenbaum and Miller (2007).

with mean 0.041 and variance 0.0009 (standard deviation 0.03) (i.e., $\mathcal{N}(0.041, 0.0009)$). We decide the mean of this *prior distribution* to be 0.041 by taking an arithmetic mean of the lowest and highest point estimates provided by Barber and Holbein (2020) and Gerber, Huber and Hill (2013), respectively: $\frac{1}{2}(0.018 + 0.064) = 0.041$. Here, we did not use the point estimate provided by Bonica et al. (2020) (0.094) as the highest estimate because their estimate is slightly far from other studies, whereas the lowest and highest estimates we use are also supported by close findings by Thompson et al. (2020) (0.021) and Barber and Holbein (2020) (0.06), respectively. Nevertheless, to acknowledge that the true effect size could be close to the one provided by Bonica et al. (2020), we choose the standard deviation to be 0.03 so that the entire prior distribution can have non-zero density on the finding by Bonica et al. (2020).

A.3 Additional Discussion on the Two Mechanisms

B Additional Information on Data

B.1 Data Collection

We collected historical voter files from Colorado and North Carolina for the 2010, 2012, and 2016 elections. These files represent "snapshots" taken at the time of each election, and are made available online by the Colorado Secretary of State and the North Carolina State Board of Elections, respectively. These snapshot in time data mean that we are able to examine the behavior of all individuals registered to vote at the time of each election. These data stand in contrast to a "contemporary" voter file – that is, the voter file that a state elections office could generate on any given day – that will contain vote history data over time, but only for those voters who would be eligible to vote if an election were to be held on that day. It would not include voters who were registered to vote in each of the last three presidential elections, but who have died, moved out of state, or have become inactive after the 2016 election.

We began by keeping only those voters who were registered to vote on Election Day in all 2012 and 2016. The dataset includes a variety of information about each voter, including their age at the time of the election, their gender, and their party registration as well a prior voting history for the 2014 midterm Congressional election. We first code for whether each person voted in the election, As required by federal law, The North Carolina voter files includes the voters self-identified race and ethnicity. This information is not required of Colorado voters and was imputed.

B.2 Coding Individual Race

To evaluate the conditional average treatment effects on the treated by racial groups, we also need to code for individual race and ethnicity (hereafter "race"). However, neither voter history files nor registration data include information about individual race. Given this limitation, we predict individual's race by drawing on a Bayesian approach proposed by Imai and Khanna (2016).

Specifically, for each registered voter, we predicted her race by choosing the race that has the highest posterior probability that she belongs to the group conditional upon her background characteristics including surname, residential address, age, gender, and registered party. For example, when our posterior probabilities for a voter look like:

Pr(Black $= 0.75 \mid$ Smith, BlockGroup1001, age30, female, and democrat)
Pr(White $= 0.15 \mid$ Smith, BlockGroup1001, age30, female, and democrat)
Pr(Asian $= 0.05 \mid$ Smith, BlockGroup1001, age30, female, and democrat)
Pr(Hispanic $= 0.05 \mid$ Smith, BlockGroup1001, age30, female, and democrat)

we code her as a black voter.

To apply this approach, we first geocoded the set of registered voters by using `censusgeocode` module in Python and obtained which Block Group and Census Track each voter belongs to. After coding this information, we computed the above posterior probabilities by using `wru` in R developd by (Khanna and Imai, 2019). Finally, for each voter, we coded her race by choosing the group that gives the highest posterior probability.

B.3 Identification for Missing Data

Like many other administrative records, our data contains missing values on multiple covariates. We assume that all variable, except for turnout in 2010, follows missing completely at random (MCAR). With the MCAR assumption, we implemented listwise deletion for (i.e., dropped) voters who contain any missing value on the covariates, except for turnout in 2010. For some unidentified reason, 58.08 % of this variable has missing values for Colorado voters.

For the variable denoting whether each unit voted in 2010 (hereafter voted2010), we impute missing values in three different ways. The first two strategies are based on the idea of partial identification. Specifically, we bound the estimated effects by using the lowest possible value and the highest possible value for voted2010. Let V_i be voted2010. Because V_i is a binary variable with a statistical support: Supp$[V_i] = \{0, 1\}$, this can be done by imputing 0 (the lowest possible value) and 1 (the highest possible value) in all missing values, respectively. In principle, we can then bound the lowest and highest possible values for the total VBM effects by applying our (causal) identification strategy to the two imputed data sets. Hence, we can expect that the "true" effects would be located somewhere between the lower and upper bounds.

The third strategy is based on the missing at random (MAR) assumption and uses a binary choice model to impute the missing values. Informally, we estimate a logistic regression with V_i as the dependent variable and other covariates as predictors (while deleting units with missing values). We then generate predicted probabilities for V_i and code 0 if the predicted probability is less than 0.5 and 1 if it is greater than or equal to 0.5. The expectation is that this enables us to point identify the missing value for each voted2010 (if missing) and that the estimated effects based on this imputed data would be between the above bounded effect estimates. Below, Table D.2 shows that this is actually the case and our main results are based on the third approach.

More formally, let $M_i \in \{0, 1\}$ be a binary variable denoting whether a value is missing for V_i. We then assume that the following two conditions are satisfied: $V_i \perp\!\!\!\perp M_i | \mathbf{X_i}$ (independence of voted2010 and the missingness conditional on covariates) and $\Pr(M_i = 1|\mathbf{X_i}) > 0$ (nonzero probability of voting in 2010 conditional on covariates). With these assumptions, we can show that $\Pr(V_i = 1|M_i = 0, \mathbf{X_i}) = \Pr(V_i = 1|\mathbf{X_i}) = \Pr(V_i = 1|M_i = 1, \mathbf{X_i})$. This implies that we can use the (population) predicted probabilities for V_i given covariates in non-missing data to impute the (population) predicted probabilities in missing data. Since we do not know the population probabilities for non-missing data, we estimate them using a logistic regression with all other covariates assuming that it is a good approximation of the underlying conditional expectation function. According to a plug-in principle, we then use $\widehat{\Pr}(V_i = 1|M_i = 0, \mathbf{X_i})$ (predicted probabilities based on the logistic regression) to impute $\widehat{\Pr}(V_i = 1|M_i = 1, \mathbf{X_i})$. We then code if 0 if the predicted probabilities are less than 0.5 and 1 if greater than or equal to 0.5.

B.4 Descriptive Statistics

Here, we examine descriptive statistics on the turnout rates for all voters and by relevant subpopulations in Colorado and North Carolina in presidential elections before and after the adoption of VBM elections in Colorado (in 2012 and 2016). This allows us to examine changes in turnout rates by different demographic groups without any causal identification. Table B.1 reports three quantities for various subpopulations: (1) difference in turnout rates between the states in 2012; (2) difference in turnout rates between the states in 2016; and (3) difference in the two differences.

		Δ_{12}(CO$_{12}$-NC$_{12}$)	Δ_{16}(CO$_{16}$-NC$_{16}$)	$\boldsymbol{\Delta}$ ($\Delta_{16} - \Delta_{12}$)
All Voters		0.136	0.195	0.059
	Frequent voters	-0.075	0.007	0.082
	Infrequent voters	0.212	0.25	0.038
Race	White	0.132	0.19	0.058
	Black	0.141	0.209	0.068
	Hispanic	0.215	0.263	0.048
	Asian	0.207	0.247	0.04
Party	Democrat	0.155	0.227	0.072
	non-Democrat	0.129	0.174	0.045
Gender	Female	0.134	0.188	0.054
	Male	0.139	0.203	0.064
Age	< 35	0.11	0.176	0.066
	35-65	0.106	0.138	0.032
	65+	0.191	0.266	0.075

Table B.1: **Differences in Turnouts in Data (Before Matching)**

Several points are worth nothing. First, except for frequent voters in 2012, registered voters in Colorado have higher turnouts than those in North Carolina. Second, Column 3 suggests that the difference-in-differences in the raw data (data set before preprocessing via exact matching) produces a positive value in all (sub)population (standard deviations are suppressed here).

C Additional Discussion on Identification

C.1 Populations and Quantities of Interest

Our primary population of interest is *a set of voters who had been registered in Colorado between the 2012 and 2016 elections*. In addition, we consider a set of politically salient subpopulations in the primary population as our secondary populations of interest. This means that we do not consider voters who registered in Colorado *only* at the time of the 2012 *or* 2016 presidential election.

Defining the populations of interest this way is critical when one wants to accurately identify and estimate the information and convenience effects of VBM elections. First, we can isolate the information and convenience effects from the registration effect (the adoption of VBM encourages more people to register and turn out). Second, we can be more certain that the composition of the primary population and subpopulations will (almost) stay the same.

Given our populations of interest, we now define our quantities of interest. Let Y_i be a binary random variable for the outcome, denoting whether a registered voter i turned out to vote in an election. Let D_i denote a binary random variable for the treatment status, representing if the same voter had an option to use a mail ballot. Based on the potential outcomes framework, let us define $Y_i^{d=1}$ as a potential outcome (turnout) for the voter had she been assigned to a jurisdiction with a mail ballot, and $Y_i^{d=0}$ as a potential outcome (turnout) for the same voter had she been assigned to a jurisdiction without the VBM adoption. Here, we assume that the consistency assumption holds as $Y_i = D_i Y_i^{d=1} + (1 - D_i) Y_i^{d=0}$. Assuming an additive effect measure, we define our first quantity of interest (with respect to the primary population or "all voters") as the average treatment effect on the treated (ATT):

$$\tau_{ATT} = \mathbb{E}[Y_i^{d=1} - Y_i^{d=0} | D_i = 1]. \tag{1}$$

Substantively, this means that we study what turnout rate would have looked like in the 2016 presidential election in Colorado if VBM elections had not been adopted in 2013 (in contrast to the reality in which voters were actually exposed to VBM elections).

Now, we consider an *effect modifier* Z_i as a categorical variable taking a discrete value z, which denotes that voter i belongs to group z. Specifically, we are interested in several group structures including the frequency of voting, race and ethnicity, gender, partisanship, and age. These groups are called effect modifiers because causal effects of interest (i.e., ATT in our case) are expected to be modified by (or conditioned upon) such groups. Based on this, we define our next quantities of interest as the conditional average treatedment effect on the treated (CATT):

$$\tau_{CATT} = \mathbb{E}[Y_i^{d=1} - Y_i^{d=0} | D_i = 1, Z_i = z]. \tag{2}$$

In practice, since we use observed outcomes of voters who had the option of use mail ballots and those who had not (and not a variant of randomized experiments), we are concerned with the

CATT conditional on a set of covariates.

$$\tau^{X}_{CATT} = \mathbb{E}_X\left[\mathbb{E}[Y_i^{d=1} - Y_i^{d=0} | D_i = 1, Z_i = z, \mathbf{X} = \mathbf{x}]\right], \tag{3}$$

where $\mathbf{X}$ is a vector of p covariates and $\mathbf{x}$ denotes a vector of specific covariate values, defined over the p-dimensional covariate space $\mathcal{X}$. We assume that the conditional ignorability holds such that within a multidimensional strata of background characteristics the potential outcomes of voters in the treatment condition and control condition do not depend on the treatment assignment.

C.2 Advantages of Using North Carolina as a Counterfactual

The key to our identification strategy is to assume that North Carolina voters are good counterfactuals for Colorado voters (Keele and Minozzi, 2013). We believe that North Carolina is an appropriate control state for several conditions. Most importantly, North Carolina has never used VBM elections, but offers all the other modes of voting as Colorado does. During the period under study, North Carolina voters were able to vote absentee by mail with no excuse, in-person before Election Day and in-person on Election Day. These options for casting a ballot were also available to Colorado voters before and after the adoption of VBM elections in 2013.

Moreover, both states have a competitive partisan environment. In 2016, the two party vote share differential was 5% in Colorado and 4% in North Carolina. Both states have a racially and ethnically diverse population. 36% of North Carolinas population is non-white white and 31% of Colorados population is non-white. The majority-minority population in Colorado is Hispanic (21%) and an identical proportion of North Carolinas population is African-American (21.5%). The age distribution of both states is comparable with 16% and 14% of Colorados and North Carolinas population over the age 65 respectively. A fifth of both states population (22%) are under the age of 18. These data indicate that North Carolina is a good comparison state with which to evaluate the adoption of VBM elections in Colorado. These comparisons are displayed in Table C.1.

C.3 Preprocessing Data

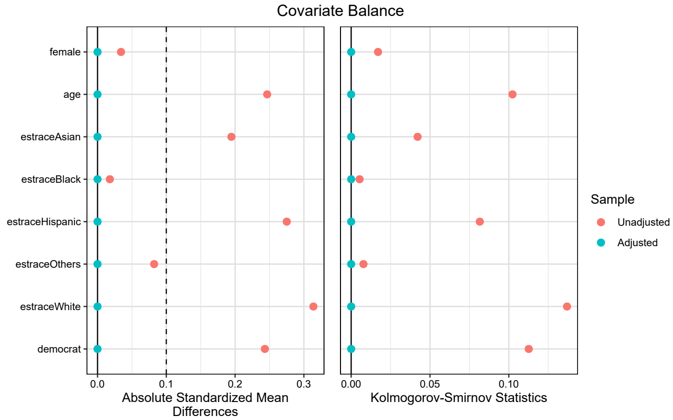
For credible observational studies, it is critical to reduce the imbalance in the covariate distributions between treated and control units in data so that our estimation and inference will be less model dependent (Ho et al., 2007). We thus preprocess our data by exact matching and then apply a difference-in-differences design to the matched data. To apply this semi-parametric approach (e.g., Abadie, 2005), we ideally want to perform exact matching on all available covariates including turnout in 2010, age, gender, race and ethnicity, and party affiliation, except for the covariate we used for an effect modifier using all units. Due to the extraordinary large size of observations ($\approx$ 1.7 million), however, we perform a simple random sampling on the original data (sampling 3% of the entire data) and apply matching to the sampled data.

As a quick visual inspection, Figure C.1 compares the standardized bias and Kolmogorov-

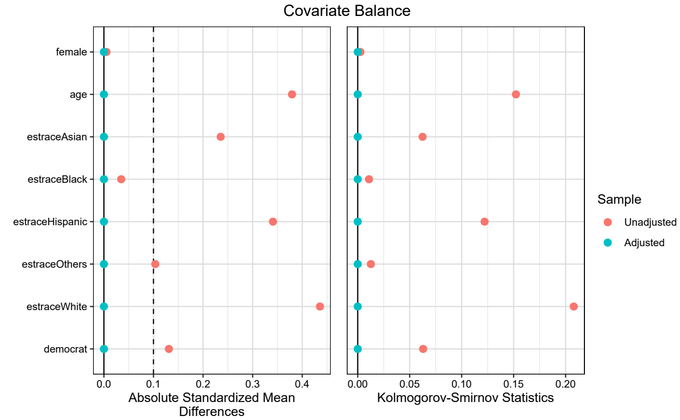
	Colorado	North Carolina
Partisan competition (2016)	5% Dem	3.6% Rep
U.S. House seats	7 (4-D, 3-R)	13 (9-R,3-D, 1-I)
Non-White	31%	36%
Under 18	22%	23%
Over 65	16%	14%
In-person before Election Day	YES	YES
Early voting	YES	YES
Same day registration	YES	YES
No excuse mail voting	YES	YES

Table C.1: **Comparison between Treated and Control States**
Note: This table shows the comparison between Colorado and North Carolina on several state-level variables of interest. The same day registration is available in Colorado since 2013 and in North Carolina since 1994.

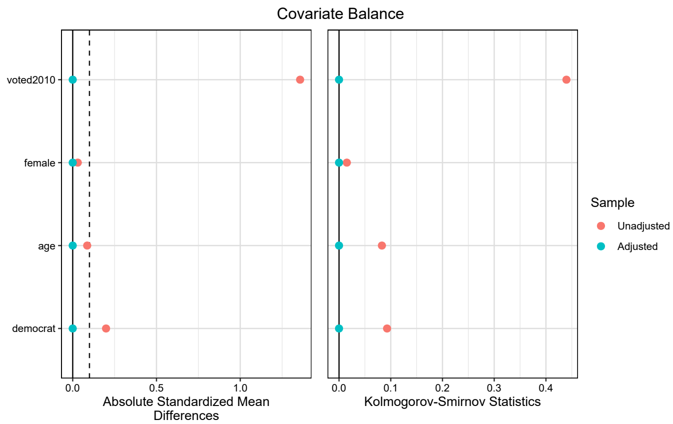
Smirnov statistics for each covariate before and after matching for each population of interest. The results demonstrate that after preprocessing our data achieves a higher balance on covariates than raw data.



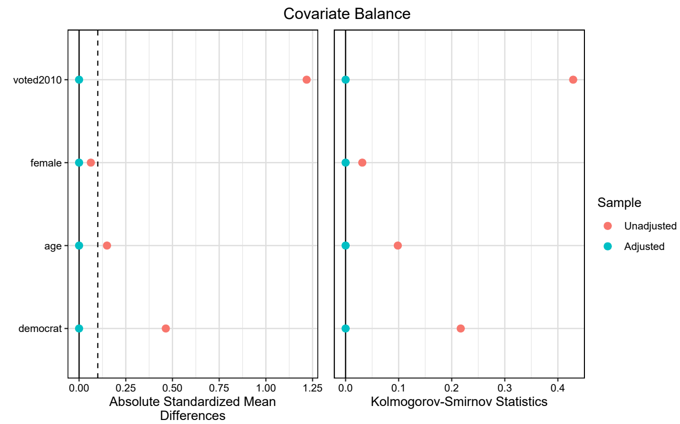
(a) Frequent voters



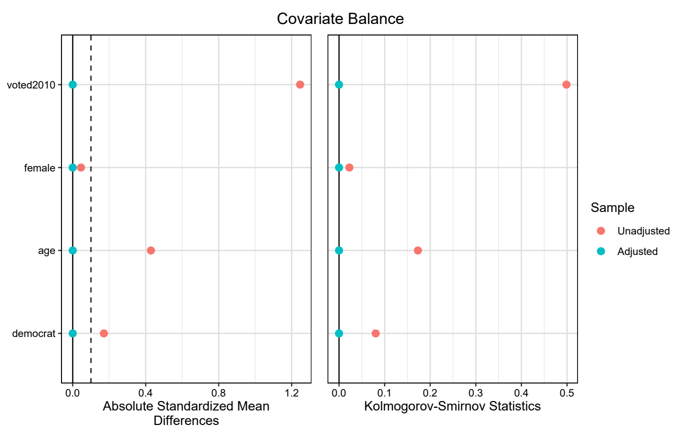
(b) Infrequent voters



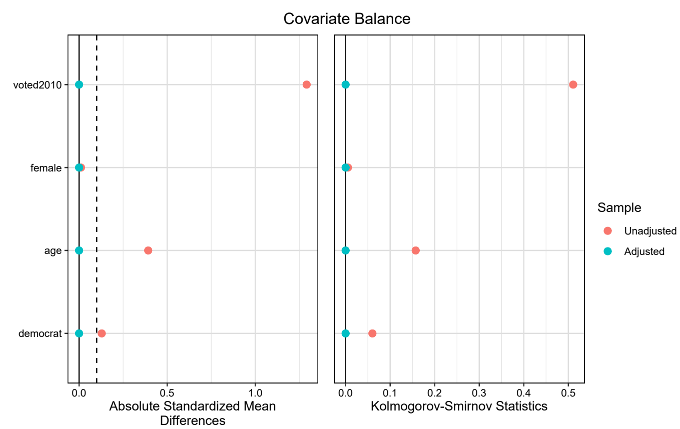
(c) White



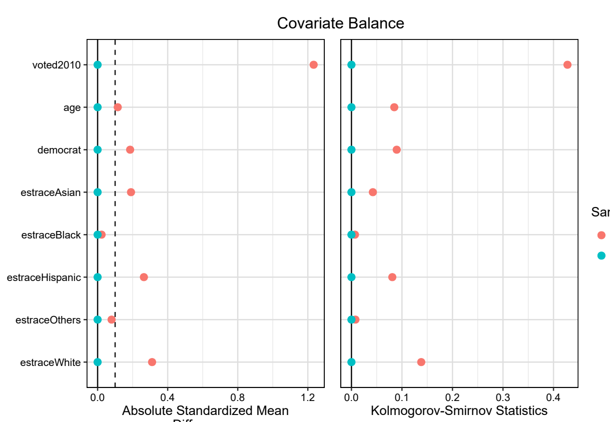
(d) Black



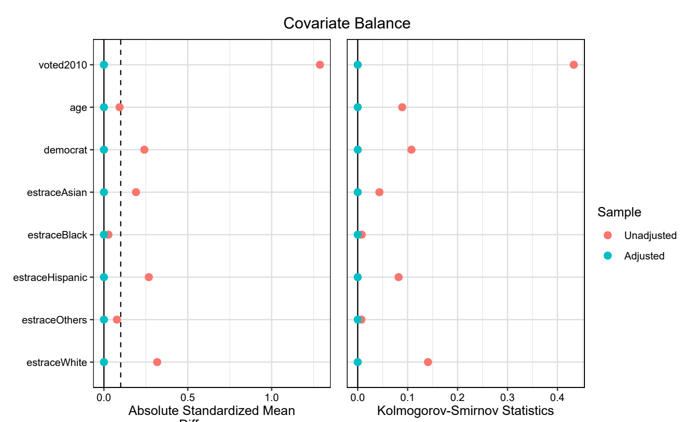
(e) Hispanic



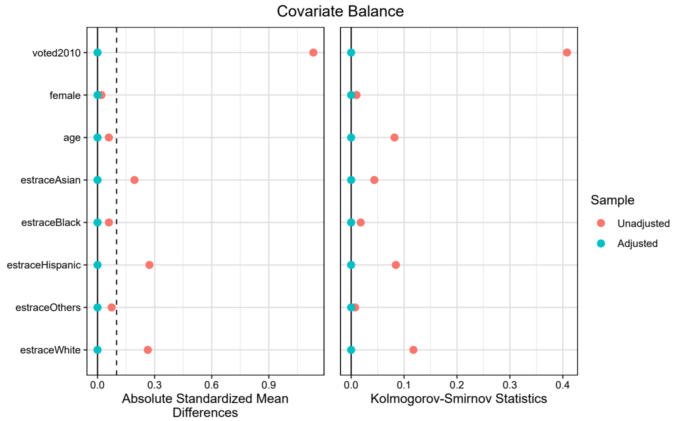
(f) Asian



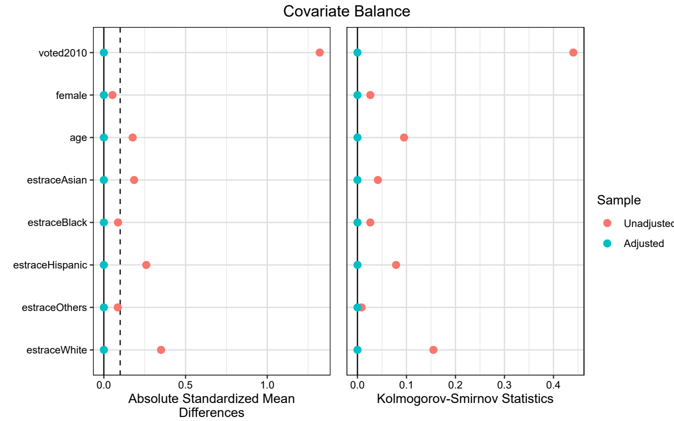
(g) Female



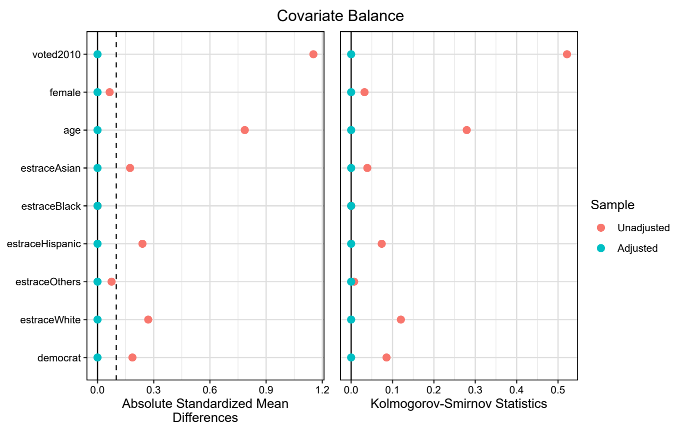
(h) Male



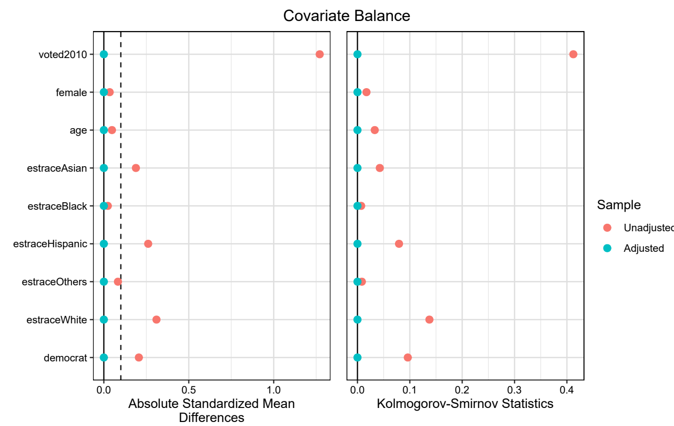
(a) Democrat



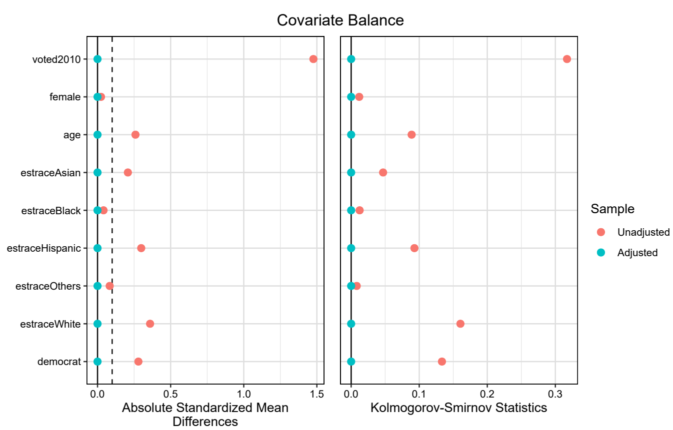
(b) non-Democrat



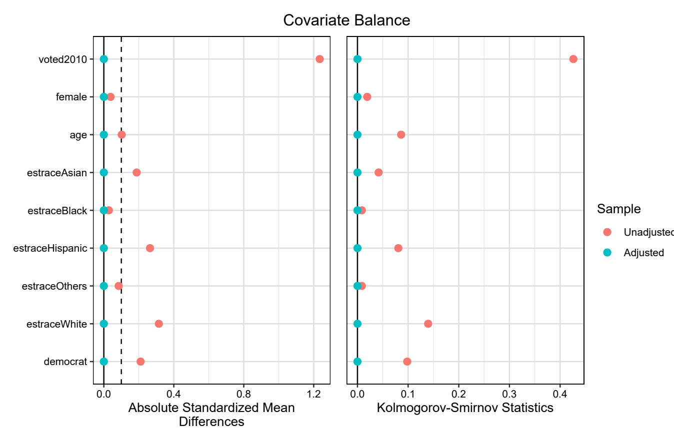
(c) Under 35



(d) 35 to 65



(e) Over 65



(f) All voters

Figure C.1: **Covariate Balance After Preprocessing**
Note: This figure visualizes the covariate balance before and after proprocessing data. We preprocess our data with exact matching (one-to-one matching without replacement) on four covariates. It shows that preprocessing increased the covariate balance, ensuring that treated units and control units look similar on observed features.

C.4 Discussion on Assumptions for Identification

To draw valid inferences under the difference-in-differences design, we must assume that the **turnout trends in the absence of the intervention in both states are the same** (i.e., the *parallel trends assumption*) (Wing, Simon and Bello-Gomez, 2018; Bertrand, Duflo and Mullainathan, 2004; Angrist and Pischke, 2008, 227-233). In the two-states and two-years design including ours, there is no way to empirically verify the parallel trends assumption (such by replacing the true intervention with a false intervention). In Online Appendix D.3, nevertheless, we offer circumstantial evidence that the turnout trends could be the same in the two states absence the intervention using aggregate election data.

Moreover, we assume that there is only one type of treatment effect and each unit's outcome is only a function of her own treatment status and not others (i.e., the *stable unit treatment value assumption (SUTVA)*). As discussed below, our research design might violate the SUTVA by allowing the interference *among treated units*. Indeed, by allowing voters to fill in their ballots *at their convenient locations* and *with anybody*, VBM might induce a form of interference between voters especially in the intervened location (e.g., state, county, precinct). For example, consider two voters who are sharing the same residence (e.g., married couples, partners, roommates, etc). If one of them received a mail ballot (or mail ballots) and talked to the other person about the upcoming election (or even about the mail ballot), it creates an interference between the two where one's outcome is now a function of her own treatment status and the other person's treatment condition. Estimation of the VBM effects with the presence of interference requires advanced techniques (Sobel, 2006). Here, we assume that no interference between units exists, but it must be scrutinized in future analysis. In contrast, we suspect that the SUTVA may not be violated due to the interference *between* treated and control units as discussed by Keele and Titiunik (2018).

D Additional Findings

D.1 Supporting Evidence for Main Results

To confirm that our substantive conclusion is robust to several decisions that we made to derive our main results, we estimate our quantities of interest by not performing preprocessing (exact matching) and using different strategies to impute voted2010 variable (discussed in Online Appendix B.3). The results are displayed in Table D.2, which confirms that our substantive conclusion remains the same.

Our main results are based on Column (4). Columns (1)-(3) are results based on a saturated (difference-in-differences) OLS without preprocessing, while Columns (4)-(6) are results based on a saturated weighted least squares (WLS) with preprocessing. Columns (1) and (4) are results from data where missing values for voted2010 were imputed by logistic regressions, whereas Columns (2) and (5) and Columns (3) and (6) are results based on partial identifications in which the lowest and highest possible values (i.e., 0 or 1) are used for imputation, respectively.

It is worth noting that the estimated effects for all categories in Column (4) are located between the estimated effects in Column (5) and Column (6), which offer lower and upper bounds of the estimated effects via partial identification (for missing data). More importantly, our substantive conclusions (i.e., effects are larger among frequent voters than infrequent voters, effects are larger among old voters than other voters, effects do not vary much by other groups) are not susceptible to the different “specifications” of our models.

	Saturated (Difference-in-Differences) OLS/WLS **(North Carolina)**					
	(1)	(2)	(3)	(4)	(5)	(6)
All voters	0.048	0.059	0.059	**0.057**	0.041	0.069
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)
Frequent	0.082	0.061	0.093	**0.081**	0.054	0.081
	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)
Infrequent	0.041	0.050	-0.015	**0.004**	0.035	-0.035
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
White	0.049	0.059	0.059	**0.057**	0.042	0.066
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Black	0.057	0.067	0.067	**0.058**	0.048	0.068
	(0.001)	(0.002)	(0.001)	(0.003)	(0.003)	(0.003)
Hispanic	0.038	0.048	0.048	**0.062**	0.050	0.074
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Asian	0.030	0.039	0.039	**0.048**	0.034	0.065
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Female	0.043	0.054	0.054	**0.044**	0.038	0.054
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Male	0.054	0.064	0.064	**0.066**	0.045	0.076
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Democrat	0.056	0.069	0.069	**0.046**	0.041	0.057
	(0.001)	(0.001)	(0.001)	(0.003)	(0.004)	(0.003)
non-Democrat	0.039	0.048	0.048	**0.057**	0.035	0.066
	(0.001)	(0.001)	(0.001)	(0.003)	(0.004)	(0.003)
Under 35	0.034	0.066	0.061	**0.037**	0.044	0.082
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)
35 to 65	0.026	0.027	0.026	**0.046**	0.027	0.047
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Over 65	0.079	0.080	0.079	**0.105**	0.076	0.105
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Covariates	✓	✓	✓	✓	✓	✓
Exact matching				✓	✓	✓
Imputing 2010 turnout	Logit	Lowest	Highest	Logit	Lowest	Highest

Table D.1: **Estimated Effects of VBM Adoption on Turnout**
Note: This table reports estimated effects of the VBM adoption on voter turnout with different specifications. Our main results are based on Column (4) and shown in bold.

	Saturated (Difference-in-Differences) OLS/WLS **(New Mexico)**					
	(1)	(2)	(3)	(4)	(5)	(6)
All voters				**0.050**	0.045	
	(0.001)	(0.001)	(0.001)	(0.005)	(0.003)	(0.003)
Frequent				**0.076**	0.048	
	(0.001)	(0.001)	(0.001)	(0.004)	(0.002)	(0.003)
Infrequent				**0.028**	0.043	
	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)
White				**0.057**	0.041	
	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.003)
Black				**0.060**	0.056	
	(0.001)	(0.002)	(0.001)	(0.006)	(0.003)	(0.003)
Hispanic				**0.020**	0.068	
	(0.002)	(0.002)	(0.002)	(0.007)	(0.003)	(0.003)
Asian				**0.069**	0.059	
	(0.003)	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)
Female				**0.058**	0.046	
	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.003)
Male				**0.053**	0.046	
	(0.001)	(0.001)	(0.001)	(0.005)	(0.003)	(0.003)
Democrat				**0.036**	0.028	
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.003)
non-Democrat				**0.058**	0.046	
	(0.001)	(0.001)	(0.001)	(0.005)	(0.004)	(0.003)
Under 35				**0.040**	0.060	
	(0.001)	(0.001)	(0.001)	(0.006)	(0.004)	(0.004)
35 to 65				**0.053**	0.034	
	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.003)
Over 65				**0.067**	0.053	
	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.003)
Covariates	✓	✓	✓	✓	✓	✓
Exact matching				✓	✓	✓
Imputing 2010 turnout	Logit	Lowest	Highest	Logit	Lowest	Highest

Table D.2: **Estimated Effects of VBM Adoption on Turnout (New Mexico)**
Note: This table reports estimated effects of the VBM adoption on voter turnout with different specifications. Our main results are based on Column (4) and shown in bold.

D.2 Placebo Tests

To further confirm the internal validity of our analysis, we perform simple robustness checks. Specifically, we perform a series of placebo tests by replacing the original outcome (i.e., turnout) with several *false outcomes*. In particular, we use a dummy variable for being Democrat, female, and white as our placebo outcome for each test, and we apply the tests for data sets with and without preprocessing. If our identification strategy "works" as we intended, we should expect to find null results (or should not find effects) for these placebo tests. The left panel of Figure D.1 visualizes the results. It indicates that our placebo tests do not find any statistically and/or substantially significance effect on our false outcomes, suggesting that our DID estimation is a valid procedure.

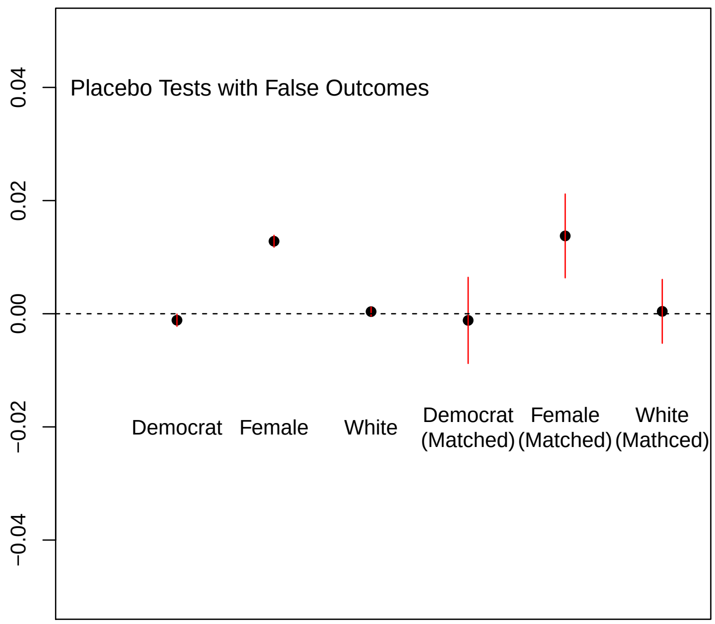


Figure D.1: **Placebo Tests with False Outcomes and Treatments**
Note: This figure presents the results of placebo tests with false outcomes and treatments, respectively.

D.3 Verifying the Parallel Trends Assumption

As noted above, the validity of our estimates hinges upon the parallel trends assumption on voter turnout in the two states without the presence of the intervention. However, in our data, no empirical method can verify the assumption, leaving us no choice but to merely believe that the assumption holds. To (at least) provide circumstantial evidence, nontheless, we visualize voter turnout in both states between 2000 and 2018 using the voting-eligible population (VEP) turnout data from McDonald's United States Elections Project (McDonald, 2008). It should be noted that the VEP is different from our primary population of interest (i.e., a set of voters who had been registered between the 2012 and 2016 elections). With this in mind, we nevertheless check if "overall turnout trends" would look alike in Colorado and North Carolina over time before the VBM adoption in 2013.

Figure D.3 shows that the time trends seem to be fairy similar before the adoption of VBM in Colorado in 2003. From the visual inspection, we could not make a credible claim about the validity of the assumption, but at least have a suggestive evidence that the parallel tends assumption is more or less satisfied, which grants a confidence in our causal identification. Nevertheless, this result can only be a circumstantial evidence since our main populations of interest are limited within specific groups both in terms of over time registration status and politically salient attributes.

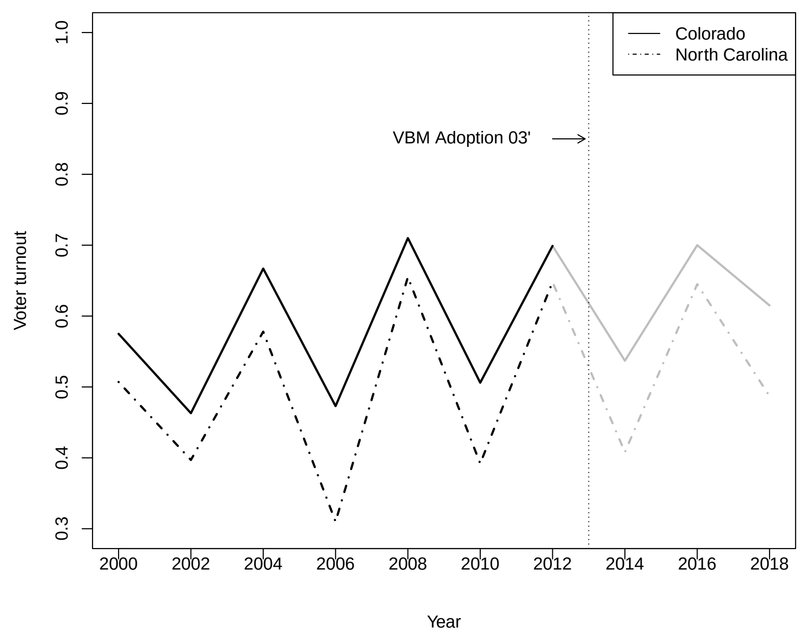


Figure D.2: **Time Trends in Voter Turnout in Colorado and North Carolina**
Note: This plot portrays the overtime trend in voter turnout in Colorado and North Carolina based on the highest office Voting-eligible population (VEP) turnout collected by McDonald (2008).

D.4 Does VBM Make the Electoral More Representational?

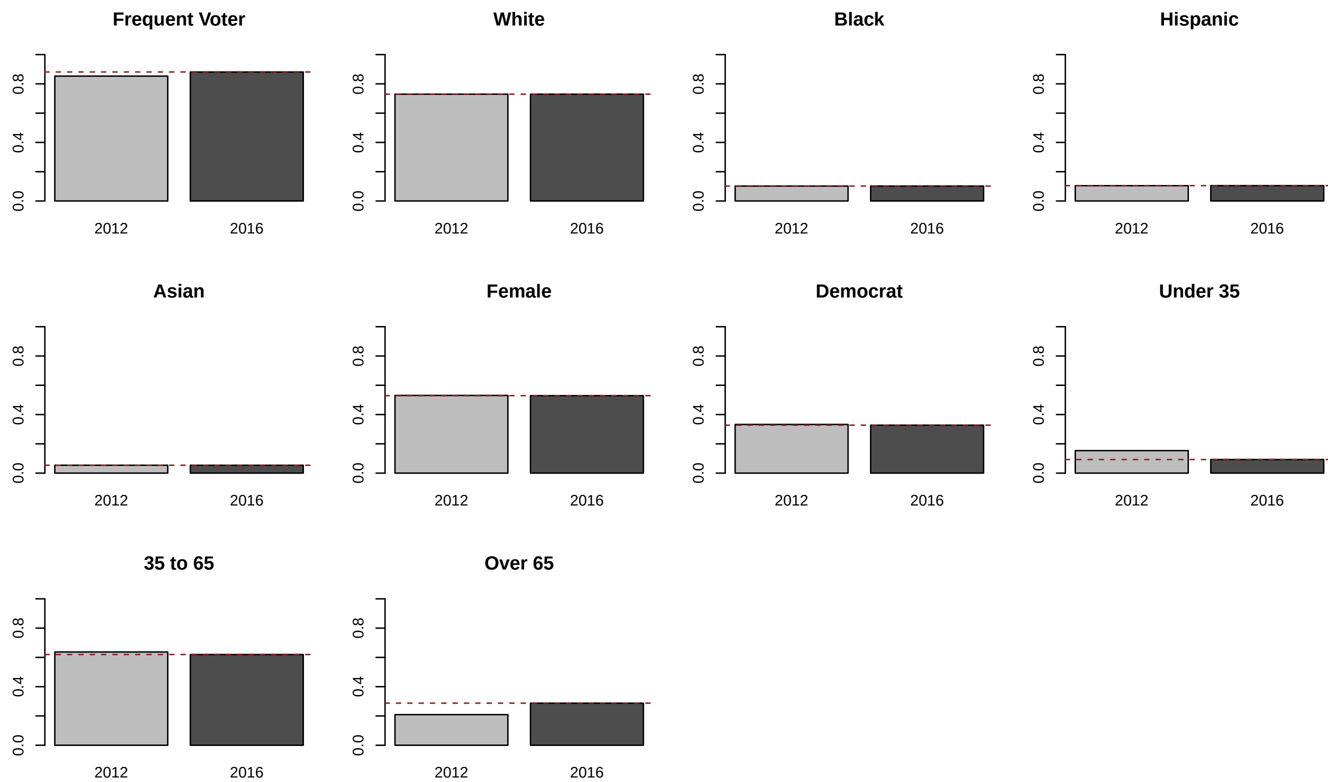


Figure D.3: **Change in the Composition of Voters Who Turned Out**
Note: This figure visualizes the change in proportion of each group among those who have actually voted in 2012 and 2016.

References for Online Appendix

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Barber, Michael and John. Holbein. 2020. "Voting During the COVID-19 Pandemic: Vote-by-Mail Does Not Advantage Democrats." *Working Paper* .

Bergman, Elizabeth and Philip A Yates. 2011. "Changing election methods: How does mandated vote-by-mail affect individual registrants?" *Election Law Journal* 10(2):115–127.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly journal of economics* 119(1):249–275.

Bonica, Adam, Jacob M. Grumbach, Charlotte Hill and Hakeem Jefferson. 2020. "All-Mail Voting in Colorado Increases Turnout and Reduces Turnout Inequality." *Working Paper* .

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(1):91–116.

Gronke, Paul and Peter Miller. 2012. "Voting by mail and turnout in Oregon: Revisiting Southwell and Burchett." *American Politics Research* 40(6):976–997.

Ho, Daniel E, Kosuke Imai, Gary King and Elizabeth A Stuart. 2007. "Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference." *Political analysis* 15(3):199–236.

Imai, Kosuke and Kabir Khanna. 2016. "Improving ecological inference by predicting individual ethnicity from voter registration records." *Political Analysis* 24(2):263–272.

Keele, Luke and Rocío Titiunik. 2018. "Geographic Natural Experiments with Interference: The Effect of All-Mail Voting on Turnout in Colorado." *CESifo Economic Studies* 64(2):127–149.

Keele, Luke and William Minozzi. 2013. "How much is Minnesota like Wisconsin? Assumptions and counterfactuals in causal inference with observational data." *Political Analysis* pp. 193–216.

Khanna, Kabir and Kosuke Imai. 2019. *wru: Who are You? Bayesian Prediction of Racial Category Using Surname and Geolocation*. R package version 0.1-9.
URL: *https://CRAN.R-project.org/package=wru*

Kousser, Thad and Megan Mullin. 2007. "Does voting by mail increase participation? Using matching to analyze a natural experiment." *Political Analysis* 15(4):428–445.

Larocca, Roger and John S Klemanski. 2011. "US state election reform and turnout in presidential elections." *State Politics & Policy Quarterly* 11(1):76–101.

McDonald, Michael. 2008. "United States election project." *United States Elections Project* .

PantheonAnalysis. 2017. "Utah 2016: Evience for the positive turnout effects of Vote at Home (also known as vote by mail) in participation counties.".

PantheonAnalysis. 2018. "Colorado 2014. Comparisions of Predicted and Actual Turnout.".

Richey, Sean. 2008. "Voting by mail: Turnout and institutional reform in Oregon." *Social Science Quarterly* 89(4):902–915.

Sobel, Michael E. 2006. "What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference." *Journal of the American Statistical Association* 101(476):1398–1407.

Southwell, Priscilla L. 2009. "Analysis of the turnout effects of vote by mail elections, 1980–2007." *The Social Science Journal* 46(1):211–217.

Southwell, Priscilla L and Justin I Burchett. 2000. "The effect of all-mail elections on voter turnout." *American Politics Quarterly* 28(1):72–79.

Thompson, Daniel M., Jennifer Wu, Jesse Yoder and Andrew B. Hall. 2020. "All-Mail Voting in Colorado Increases Turnout and Reduces Turnout Inequality." *Working Paper* .

Wing, Coady, Kosali Simon and Ricardo A Bello-Gomez. 2018. "Designing difference in difference studies: best practices for public health policy research." *Annual review of public health* 39.