Cross-State Strategic Voting

Gordon B. Dahl   Joseph Engelberg   Runjing Lu   William Mullins*

June 2023

Abstract

We estimate that 3.1% of the US voter population is registered to vote in two states, opening up the possibility for them to choose where to vote. Double registration is 3.5 times more likely in the wealthiest 1% of zip codes vs. the bottom quartile, giving wealthy Americans more voting power. Double-registrants respond to both incentives and costs, disproportionately choosing swing states (higher incentive) and states which automatically send out mail-in ballots (lower cost). We call this behavior cross-state strategic voting (CSSV) and estimate there were 217,000 such votes involving swing states in the 2020 presidential election. While there are more Democrat double-registrants, Republicans are more responsive to swing-state incentives. The net effect did not alter the 2020 election, although it could change the outcome in closer elections (e.g., Florida in 2000) or if one party increased CSSV relative to the other.

Keywords: Election Incentives and Costs, Strategic Voting, Voting Fairness

JEL Classification Number: D72

*Gordon B. Dahl: UC San Diego, NBER, Norwegian School of Economics, CESifo, CEPR, IZA (gdahl@ucsd.edu); Joseph Engelberg: UC San Diego (jengelberg@ucsd.edu); Runjing Lu: University of Alberta (runjing1@ualberta.ca); William Mullins: UC San Diego (wmullins@ucsd.edu) We thank Carlos Avenancio-Leon, Eli Berman, Julie Cullen, Ramona Dagostino, Nir Jaimovich, Craig McIntosh, Karthik Muralidharan, Paul Niehaus, Antoinette Schoar, and Tom Vogl, as well as seminar participants at several universities and conferences.

1. Introduction

Trust in elections is critical to the legitimacy of a democracy, but many Americans have little confidence in how elections are administered (e.g., Marist Poll, 2021, MIT Election Lab, 2021). Court cases and the literature have focused on *within*-state election issues, such as certification, polling locations and voter ID laws. Far less is known about *cross*-state issues because states run their elections independently, with no formal mechanism to coordinate voter registrations or votes across states.

The lack of central coordination introduces the possibility that individuals could vote in states they no longer live in, creating concern about fairness and election integrity. If some individuals can choose where to vote and others cannot, this violates the principle of one person, one vote which says "one person's voting power ought to be roughly equivalent to another person's within the same state" (Cornell Law Dictionary, 2022).[1] Due to the Electoral College, a vote in a swing state has greater impact, giving those who can choose which state to vote in disproportionate influence. Concerns about these cross-state issues have led to recent legislation. For example, Senate Bill 1260, signed into Arizona law in June 2021, specifies that "A person is guilty of a class 5 felony who ... [k]nowingly provides a mechanism for voting to another person who is registered in another state."

These concerns and legislative actions exist despite a lack of evidence on how many people are registered to vote in two different states, their characteristics, and their voting behavior. We provide this evidence using comprehensive data on all 203 million U.S. voter registrations. This paper is the first in the literature to (i) estimate the number of double-registrants and their characteristics, and (ii) document how their voting responds to incentives and costs.

We begin by estimating the number of voters who are double-registered as of October 2020. The same name appearing on the voting rolls of two different states could either be the same person or different people who happen to share a name. If it is the same person, the probability of the same birthday is 1; if they are different people, this probability will be much smaller. We use the empirical frequency that same-named people share the same

[1]"In Reynolds v. Sims (1964), the Supreme Court held that "The Equal Protection Clause requires substantially equal legislative representation for all citizens in a State regardless of where they reside."

birthday and the Law of Total Probability to estimate that 3.1% of voters are registered in two states. In other words, pairs of registrants in different states with the same name have the same birthday more often than would be expected by chance, and we use that excess probability to estimate that 6.1 million individuals are double-registrants.

We next consider the rate of double registration by measures of wealth and income, using the same Law of Total Probability approach. Based on zip code-level housing values, the bottom quartile has a 2.1% double-registration rate. This rate jumps in the upper tail of the distribution. For individuals at or above the 90th, 95th, and 99th percentiles, the rates are 4.5%, 5.5%, and 7.4%, respectively. In other words, compared to the bottom quartile, the rate of double-registration is twice as high among the top 10% and 3.5 times among the top 1%. Using zip code-level income yields a similar pattern. Historically, voting power depended on property ownership and income in the US; legislation and court decisions have attempted to sever this link. However, our evidence points to richer voters having disproportionately more voting power via double-registration. This fact, along with growing concerns about wealth and income inequality (e.g., Piketty and Saez, 2003, Smith et al., 2023), exacerbate any fairness concerns that come with double-registration.

The Law of Total Probability allows us to estimate the size and distribution of the double-registered population, but does not identify which particular individuals are double-registrants. To create such a sample, we exploit cell phone data, which we have for a subpopulation of registrants. The characteristics of this group closely match those of the full population (Table 1). We show that two people with the same name, gender, birthdate information, and cell phone number registered in two different states are almost certainly the same person. Using this novel approach, we identify a sample of 590,991 double-registrants.

Having identified a sample of double-registrants, we next turn to their voting behavior in the 2020 presidential election. Ceteris paribus, these individuals may prefer to vote in the state with a strategic advantage. For example, a voter registered in both Mississippi and Georgia may choose to vote in Georgia because it is a swing state and Mississippi is not. Similarly, a voter registered in both Oregon and Idaho may prefer to vote in Oregon because it automatically sent them a mail-in ballot, unlike Idaho. We call this behavior *Cross-State*

Strategic Voting (CSSV).[2]

Regardless of voter preferences, most state election laws specify that citizens should vote in the state of their legal domicile, and they can have only one (Greabe, 2012). Characteristics which help courts determine an individual's domicile include where they have a driver's license, where they pay state taxes, and where they spend the majority of the year. For example, the guidance given to service members, who relocate frequently, by the government is: "Your voting residence is within your state of legal residence or domicile ... [it] is used for state income tax purposes, and determines eligibility to vote for federal and state elections ... You have the option to establish residency or domicile each time you are transferred to a new location. Once you change your residence or domicile, you may not revert to a previous residence without re-establishing a new physical presence according to residency laws of that state" (Federal Voting Assistance Program, 2022).

If individuals are choosing where to vote based on costs and incentives (CSSV), there are two possible ways this could happen. First, someone may vote in a state which is not their legal domicile. This could be purposeful or due to ignorance of state voting laws.[3] Alternatively, voters may be willing to incur some costs to establish – or reestablish – domicile in a state they prefer to vote in. For example, a double-registrant who moved from Arizona to California might spend resources or time in Arizona to reestablish domicile there in order to vote in a swing state. Absent these strategic behaviors, we would not expect where double-registrants vote to coincide with whether states are swing or whether states make it easier to vote.

Nevertheless, we find strong evidence of CSSV. For each double-registrant, we call the state with the earlier date of registration the "first" state and the state with the later registration the "second" state. If an individual's first state is a swing state but the second is not, their likelihood of voting in the first state is 4.6 percentage points (pp) higher than if neither state is swing. This amounts to a 31% increase relative to a mean of 15%. In other

[2]Strategic voting (e.g., Duverger, 1954, Farquharson, 1969, Myerson and Weber, 1993) typically refers to voting as if one's vote is pivotal, which can lead to voting for a less-preferred candidate if the preferred candidate is unlikely to win. In contrast, voters engaging in CSSV are strategic in their choice of state to vote in.

[3]Voting in a state which is not one's legal domicile is illegal and has resulted in a handful of prosecutions (New York Times, 2022).

words, when the opportunity to participate in a consequential election only exists in the first state, voters are more likely to vote there, despite having registered in another state more recently. Likewise, if only the second state is swing, they are 2.9 pp less likely to vote in the first state.

Our main measure of voting costs is whether a state automatically mails out ballots. We find that double-registrants disproportionately choose to vote in the state which makes voting less costly. If only the first state auto-mails ballots, the likelihood of voting there increases by 63% of the mean. However, if only the second state does, the likelihood of voting in the first state decreases by 37%. We find smaller effects for early voting, another measure of voting costs.

Choosing which state to vote in presupposes that a person will vote in the first place. Thus, one potential concern is that incentive and cost considerations influence the likelihood a person will vote at all rather than the particular state they will vote in. We address this concern by considering a subset of double-registrants who are at least 99% likely to vote based on observables, such as their voting history. This "identification at infinity" approach continues to show strong evidence of CSSV.

Our main analysis uses a binary measure for whether a state is swing. However, not all swing states are created equal: Iowa was much less of a swing state than Georgia in 2020. Moreover, the difference in closeness between a double-registrant's pair of states will also vary. For example, Georgia could be paired with a similarly close swing state like Arizona or with a less close swing state like Iowa. We find that, holding the closeness of the first state fixed, as the second state's election becomes less close, double-registrants are more likely to vote in their first state.

A natural question is whether the same incentives and costs lead double-registrants to cast a ballot in both states. CSSV predicts that double voting should be most prevalent when both states are swing, which is exactly what we find. We also document that auto-mailed ballots increase the likelihood of double voting. While not the focus of this paper, we find double voting to be extremely rare in the 2020 election (at most 0.034% of voter registrations), roughly in line with the estimate in Goel et al. (2020) for the 2012 election.

Finally, we consider how CSSV affected the 2020 election. We estimate a large number of CSSV-induced votes: 317,000 at the national level, of which 217,000 involve a swing state. The distribution of these votes to Trump and Biden is driven by three key findings. First, there are *twice* as many Democratic double-registrants as Republicans. Second, double-registered Republicans are more responsive to incentive-based CSSV: if one of their states is swing, they are more likely to vote there compared to Democrats. Third, mail-in ballot states are disproportionately blue states (and not swing states) and Democrats are more sensitive to cost-based CSSV than Republicans, keeping more Democrat votes in mail-in ballot states and away from swing states. While the first effect favors Biden, the second two benefit Trump, so that on net CSSV did not affect the 2020 election. However, CSSV could be important in closer elections such as the 2000 election, where the gap was 537 votes in Florida and 366 in New Mexico.

Our paper relates to three strands of the literature. The first tests for the presence of classical strategic voting, where individuals behave as if their vote is pivotal, including Fujiwara (2011), Kawai and Watanabe (2013), Spenkuch (2015), Anagol and Fujiwara (2016), Pons and Tricaud (2018), Spenkuch (2018), and Clark et al. (2022). Our contribution is to identify a new and potentially consequential form of strategic voting: *Cross-State Strategic Voting*. In other words, while the existing literature demonstrates strategic choice of *who* to vote for, we show individuals strategically choose *where* to vote.

Second, we contribute to the literature on potential violations of election law (e.g., Gronke et al., 2008, Minnite, 2010, Cottrell et al., 2018, Eggers et al., 2021, Cantoni and Pons, 2021, Ferlenga and Knight, 2022). Given new legislation targeting double-registrants, we contribute by estimating the size and characteristics of this population and examining their behavior.

Third, our paper relates to work examining determinants of voting, including features of elections designed to lower voting costs, such as early voting, voting technology, and mail-in ballots (e.g., Braconnier et al., 2017, Bursztyn et al., 2017, Yoder et al., 2021, Cantoni, 2020, Kaplan and Yuan, 2020, Bechtel et al., 2018, Spenkuch and Toniatti, 2018, Fujiwara et al., 2016, DellaVigna et al., 2016, Fujiwara, 2015, Ortoleva and Snowberg, 2015, Gerber et al., 2013, Gentzkow et al., 2011). We find consistent evidence that costs matter using a new empirical design and a unique sample of voters. We also document an unintended effect of reducing such costs: states which make voting easier pull votes from swing states, most of which are not low-cost.

The remainder of the paper proceeds as follows. The next section describes our voter registration data. Section 3 estimates the prevalence of double-registrants. Section 4 presents evidence of CSSV, while Section 5 estimates its effect on the 2020 election. Section 6 concludes.

2. Data

Our data on voter registrations across the United States is from L2, a well-known, nonpartisan data vendor used by political campaigns and the academic literature (e.g., Allcott et al., 2020, Billings et al., 2021, Spenkuch et al., 2023). The L2 dataset contains information on registered voters in all 50 states (and DC), including their name, address, birthdate, and voting history. In addition, L2 merges the state-level voter files with other administrative and commercial datasets, adding additional voter characteristics such as cell phone number.

The data also contain information on individuals' political partisanship. For 34 states (and DC), L2 assigns political affiliation using self-reported voter registration. For the remaining states, L2 infers party using a variety of data sources, including voter participation in primaries, demographics, exit polling, and commercial lifestyle data.[4]

L2 data is regularly tested by political campaigns in the field. Academic papers have also verified the accuracy of the partisanship measures in voter files. For example, Bernstein et al. (2022) validates the accuracy of L2 partisanship by comparing state files to L2 data; Brown and Enos (2021) runs a survey to verify L2 partisanship; Pew (2018) compares commercial voter file data to Pew national survey microdata. In addition, when we sum L2's individual-level voting records by state they closely match certified state totals. In all but six states the vote totals in the L2 files exceed those in the certified totals (by 0.8% on average), consistent with (i) "residual voting" (casting a ballot to vote for some offices but not for president Stewart III et al. 2020), and (ii) some ballots not being counted for procedural reasons. In five of the remaining six states the difference is small (-0.4% on average), which could be

[4]Party identification is inferred in Alabama, Georgia, Hawaii, Illinois, Indiana, Michigan, Minnesota, Missouri, Montana, North Dakota, Ohio, South Carolina, Texas, Vermont, Virginia, Washington.

due to minor differences in when vote totals are certified versus when L2 receives the data.[5]

2.1 Summary statistics

Table 1 presents summary statistics. The first column considers the full dataset of L2 voters as of October 2020; the second column is for voters for which we have the data (e.g., cell phone number) necessary to identify them as double-registrants; the final column reports statistics for our sample of identified double-registrants. Specifically, these are individuals with the same cell phone number, first name, last name, gender, month of birth, and year of birth but are registered in two different states. As we will show in Section 4.1, the likelihood of two people from different states having these same matched characteristics but *not* being the same person is less than 0.001%.

There were 203 million voter registrations as of October 2020 (31% Republican and 41% Democrat); 74% voted in the 2020 presidential election. The characteristics of voters with the required variables for our matching procedure (in column 2) are remarkably similar to the full dataset in terms of voting history, partisan fractions, and demographics. In contrast, double-registered voters (column 3) are quite different. This is unsurprising, because double-registrants have necessarily moved, and so tend to have characteristics associated with geographic mobility: higher wealth (27% in column 3 vs. 20% in column 2), higher income (30% vs. 21%), younger (59% vs. 40% for the fraction born after 1975), and less likely to be a homeowner (50% vs. 67%). Double-registrants are also less likely to be Republicans (21% vs. 33%). In fact, Democrats outnumber Republicans 2:1 among the double-registrant sample in column 3 (46% vs. 21%), consistent with the fact that Republicans are less likely to move.[6]

Appendix Figure A3 plots a histogram for the elapsed time between when an individual registers to vote in their first and second states. The figure shows that most double-registrants are not recent movers: two-thirds have a gap between registrations of more than

[5]The lone outlier is Utah, where the certified vote total is 16% larger than the number of cast ballots in the L2 data. Upon examining the original voter file that L2 received from the state of Utah, we find the same difference. In other words, the discrepancy is attributable to the state of Utah, not L2. Our results are unaffected by dropping Utah.

[6]Using Cooperative Election Study survey data for those registered to vote in 2020, we find that Republicans are 2.6% less likely to have moved over the past year compared to Democrats.

5 years.

3. Prevalence of double-registrants

In this section, we estimate how many of the 203 million registrations correspond to the same individual in two states. For example, if a person is initially registered in Arizona and then moves to and registers in California, but remains on the voter roll in Arizona, they will be counted twice in the 203 million total.[7] We also estimate how this prevalence varies by wealth, income, and party.

3.1 Estimation

We estimate the number of double-registrants in our sample by measuring how many people have the same name and birthdate in two different states, in excess of the number expected by chance. We define Name Uniqueness (NU) as the total number of times a (first name, last name) combination occurs in our voter registration data. We remove the 54,067,832 registered voters whose names occur only once in the sample (NU = 1) because these names have no corresponding match in any other state.

Among the remaining voters, we find pairs with the same first name, last name and birth year in two states, and consider the subset for which we have birth day and month for both voters in the pair.[8]

The estimate proceeds as follows. By chance, pairs of *different* people with the same name and birth year will have the *same* birth day and month in 1/353 or 0.2833% of cases, assuming birthdays are uniformly distributed through the year.[9] If the empirically observed percentage of cases with the same birth month and day is higher than 0.2833%, then we can

[7]L2 performs a data cleaning process before making their data available to academics and campaigns. To the extent that L2 removes some double-registrants in this process, our estimate of the number of double-registrants will be an underestimate.

[8]27 states are missing day of birth information at a non-trivial rate, and states often record the day of birth as 1 when information is missing. We estimate the percentage of double-registrants among pairs of observations for which day is available and is not the first of the month. We apply this estimate to the full population of matched pairs of registrations, based on the assumption that the two missing conditions are uncorrelated with double-registrant status.

[9]Using the empirical distribution of birthdays would not materially alter the results of this calculation. Relatedly, we first match on name and year because specifying a counterfactual probability based on day, month, and year is less straightforward.

infer that some of these pairs are in fact the same person. For a pair of records, we define D_1 and D_2 as the birth day and month of records 1 and 2, respectively. We can then apply the Law of Total Probability:

$$\begin{aligned} P(D_1 = D_2) = P(D_1 = D_2 \mid \text{same person}) \times P(\text{same person}) \\ + P(D_1 = D_2 \mid \text{not same person}) \times P(\text{not same person}) \end{aligned} \tag{1}$$

Absent recording errors, $P(D_1 = D_2 \mid \text{same person}) = 1$. Assuming births are evenly distributed across days of the year, $P(D_1 = D_2 \mid \text{not same person}) = 0.2833\%$. For each level of NU we can estimate $P(D_1 = D_2)$. Using this information and rearranging terms, for each NU we obtain:

$$P(\text{same person}) = \frac{P(D_1 = D_2)\ -\ 0.2833\%}{1\ -\ 0.2833\%} \tag{2}$$

For example, when NU = 2, we find that $P(D_1 = D_2) = 87.511\%$, but we would expect this to occur by chance in only 0.2833% of pairs if they are not the same person. Using equation 2 we can infer that $\frac{87.511\% - 0.2833\%}{1 - 0.2833\%} = 87.476\%$ of these pairs are the same person, which translates to 1,294,288 voters. Since our sample has 16,923,844 voters with NU=2, we estimate that 1,294,288/16,923,844 = 7.65% of them are double-registrants.

We calculate this for each value of NU from 2 to 150 and plot them in Appendix Figure A1. As the value of NU increases, the probability that a pair is the same person registered in two states falls, converging to approximately 2.55%. The intuition for this decline is that a pair of registrations with the same very unusual name (e.g., NU = 2) has a considerable chance of being the same person, but this likelihood is much lower (and converges to 2.55%) for more common name pairs (e.g., NU=50 or higher).[10]

Summing over all levels of NU generates an estimate of 5.1 million voters that were double registered for the 2020 general election. This implies that of the 203 million voter registrations there are 192.8 million single-registrants and 5.1 million double-registrants, so that double-registrants make up 2.6% of all voters.

[10] We performed this calculation for all values up to NU=150. Because the probability remains approximately 2.55% after NU=100 we use this value from then on.

This analysis has assumed all double-registrants have identical recorded first and last names in both state registrations. However, this is not always the case for a variety of reasons, such as name changes following marriage, name recording errors, use of nicknames (e.g., Thomas vs. Tom) and so on. We correct for this by estimating a scaling factor capturing the ratio of double-registrants with identically recorded first and last names to the full set of double-registered individuals (see Appendix 1 for details). Applying the adjustment factor of 1.1954, we estimate there were 6.1 million double-registrants, or 3.1% of voters as of October 2020.

3.2 Heterogeneity by wealth, income, and party

Figure 3 displays the rate of double registration using measures of property wealth and income, using the same Law of Total Probability approach. Based on zip code-level housing prices using Zillow's home value index (upper panel), the bottom quartile has a double-registration rate of 2.1%. The middle 50 percent (25-75th percentile) rises modestly to 3.0%. In contrast, individuals living in wealthier zip codes are more likely to double register: the rates are 4.5%, 5.5%, and 7.4% at or above the 90th, 95th, and 99th percentiles, respectively. In other words, individuals living in the wealthiest neighborhoods are far more likely to have the ability to engage in cross-state strategic voting relative to the majority of Americans.

In the lower panel we use zip code-level adjusted gross income from IRS tax data.[11] The bottom quartile has a double-registration rate of 2.7%. The middle half is similar. However, the highest-income zip codes have much higher rates: 4.7% (≥90th), 5.3% (≥95th), and 6.4% (≥99th). This pattern mirrors the gradient in the upper panel, albeit less pronounced.

Historically, voting power in the US was directly tied to either owning property or paying taxes. A series of legislative reforms and court decisions served to grant voting power more equally. However, the evidence here points to the right tail of the wealth and income distributions having disproportionately more voting power via double-registration.

Next, we consider heterogeneity by party affiliation. We find that Republicans have a double registration rate of 1.9%. In contrast, Democrats have a substantially higher rate of

[11] The variable for individual-level income in L2 is missing 60% of the time. Moreover, individual-level income is not missing at random, with a higher double-registration rate among those with missing income. Therefore, we use a zip code-level proxy.

3.2% (Independents have the highest double-registration rate of 4%). Combined with the fact that there are more Democrats than Republicans in the voter population, this means that Democrat double-registrants outnumber Republicans two to one. Therefore, Democrats have more opportunity to exert voting power via double-registration.

4. Cross-state strategic voting

4.1 Identifying a sample of individual double-registrants

The preceding section allows us to estimate the size of the double-registered population, but does not identify which individuals are double-registrants. To construct a sample of identified double-registrants, we take advantage of cell phone data. As we showed in Table 1, the characteristics of the population for which we have cell phone data closely matches those of voters without it. We match voters from different states on first name, last name, gender, year and month of birth, and cell phone number. Note that cell phone number is a non-redundant piece of information for matching because numbers are merged to an individual voter registration using name and *address* by the data provider (and hence within each state voter file) before we perform cross-state matching.

The analysis proceeds as follows: we first find all pairs of voters from different states that have the same first name, last name, gender, year of birth, and matching first 9 digits of their cell phone number. The tenth digit of each voter's cell phone number and month of birth are used as "check digits" to ensure a correct match; we use two check digits instead of one to minimize the effects of recording errors on our calculations.[12] At the pair level, we have the following conditional probabilities:

P(Either check digit matches | same person) $\approx$ 1

P(Either check digit matches | not same person) = 1 - (9/10 $\times$ 11/12) = .175

Empirically, we find that among the 629,035 matches, 99.9703% of them have at least one check digit that matches (either the same cell phone 10th digit, same month of birth, or both).

[12] We do not use day of birth to define the sample as this would exclude individuals in states that do not consistently record day of birth, as noted in Section 3.

Using the above conditional probabilities and the Law of Total Probability, we find that 99.964% of our matches are indeed the same person. To define our sample, we require that *both* check digits match (reducing the sample to 611,129 matches). Applying Bayes rule yields a final estimated probability of over 99.999% that the matches in our sample correspond to the same person. For our final regression sample we further drop the approximately 20,000 voters registered in more than two states.[13] This leaves us with a sample of 590,991 double-registrants.

4.2 Main results

We now explore whether a double-registrant chooses to vote in a state depends on that state's election characteristics. Specifically, elections in "swing states" are more consequential for national-level outcomes, which may incentivize voters to participate more than elections that are near-certain. Indeed, despite the near-zero probability that an individual vote will be pivotal in any election, individuals seem more likely to vote in closer elections (e.g., Blais, 2006, Alvarez et al., 2006, Bursztyn et al., 2017). In addition, some states make it easier to vote, for example by automatically mailing ballots to all registered voters, which lowers the time and effort cost of voting. We call voting that responds to these incentives and costs when choosing where to vote Cross-State Strategic Voting (CSSV).

To put our analysis in context, a majority of states have laws that restrict voting to one's state of domicile. Domicile is a legal term: "the place where a person has fixed his habitation and has a permanent residence, without any present intention of removing there from" (Black's Law Dictionary). Individuals may have multiple residences but only one legal domicile at a time. Typically, the government determines a person's domicile based on characteristics such as where they have a driver's license, where they pay state taxes, and where they spend the majority of the year. For example, service members, who relocate often, are advised that "Your voting residence is within your state of legal residence or domicile... [it] is used for state income tax purposes, and determines eligibility to vote for federal and state elections" (Federal Voting Assistance Program, 2022).

[13] We also exclude registered voters whose birthdays are on January 1st because there is an abnormal mass of voters, suggesting that some states assign this date by default when only year of birth is known. We also drop 62 voters recorded as registering in two states on the same day.

If double-registrants are choosing where to vote based on incentives and costs (CSSV), there are two possible ways this could happen. First, they could vote in a state which is not their legal domicile, either purposefully or due to ignorance of state laws. Alternatively, they could incur some costs to establish or reestablish domicile in a state. Returning to our service member example, the government instructs: "You have the option to establish residency or domicile each time you are transferred to a new location. Once you change your residence or domicile, you may not revert to a previous residence without re-establishing a new physical presence according to residency laws of that state" (Federal Voting Assistance Program, 2022). Absent any strategic behaviors among double-registrants, we would not expect where they vote to coincide with whether states are swing or whether states make it easier to vote.

Among our sample of 590,991 double-registrants we know the date of voter registration in each of their two states; we call the state with the earlier registration the "first" state and the state with the later registration the "second" state. A voter's second state is their most likely state of domicile, as this is the one they most recently registered to vote in. Consistent with this, only 15% of double-registrants voted in their first state in the 2020 presidential election.

In Figure 1, we show both the source (second states) and destination (first states) of these vote flows. States that are neither swing nor auto-mail lose the largest share of their votes and attract the fewest. In other words, within each destination group, the white bars (*from: neither*) are always taller than the black (*from: swing*) and grey (*from: auto-mail*). For example, in the middle destination group, neither states lose 19% of votes to swing states, whereas swing states and auto-mail states only lose 16% and 13%, respectively. Looking across destination groups, the left-most group of bars (*to: neither*) are shorter than those of the other destination groups. For instance, neither states attract only 14% from other neither states, whereas swing and auto-mail states attract 19% and 27%, respectively.

We estimate these effects more systematically in the following regression framework:

$$\begin{aligned} \textit{Vote first}_i = \gamma_0 &+ \gamma_1 \textit{Only first swing}_i + \gamma_2 \textit{Both swing}_i + \gamma_3 \textit{Only second swing}_i \\ &+ \gamma_4 \textit{Only first low-cost}_i + \gamma_5 \textit{Both low-cost}_i + \gamma_6 \textit{Only second low-cost}_i + \epsilon_i \end{aligned} \tag{3}$$

Vote first$_i$ is an indicator equal to one if individual i, who is double-registered in a pair of states, voted in their first state in the 2020 presidential election, and zero otherwise. The omitted categories are pairs of states that are both not swing, and pairs that are both not low-cost. We identify swing states as those for which the less favored party has at least a 20 percent chance of winning according to the betting market PredictIt (PredictIt, 2020) one month before the election (October 3, 2020). This regression exploits the variation across 2,332 ordered state pairs (out of a theoretical maximum of $51!/(51\text{-}2)! = 2{,}550$).

We find strong evidence of CSSV in Table 2. The positive and statistically significant coefficient on *only first swing* in column 1 indicates that when the first state is swing but the second is not, double-registrants are 4.6 pp more likely to vote in their first state, which translates to a 31% increase relative to the mean. Moreover, when double-registrants' second state is swing and their first is not, they are 20% *less* likely to vote in their first state relative to the mean. Overall, votes among double-registered individuals flow to the states that are more consequential.

We also find evidence that the cost of voting affects where double-registrants choose to vote. Automatically receiving a mail-in ballot reduces the effort required to vote and so typically increases voting participation (e.g., Gerber et al., 2013, Hodler et al., 2015).[14] We find that if only the first state automatically mails out ballots, the likelihood of voting there increases by 72% relative to the mean (column 1). If both states have auto-mailed ballots, this effect shrinks to 47%. However, when only the second state has auto-mailed ballots, double-registrants become less likely to vote in their first state (-40%). Note that our estimates for CSSV in response to incentives (swing states) and costs (auto-mailed ballots) are largely independent, as only one state – Nevada – had both characteristics in 2020.

Early voting which allows people to vote before election day constitute another dimension of voting costs. We find similar, albeit more muted, cost effects for these policies in column 2, while the main incentive and cost effects of swing states and auto-mailed ballot policies are essentially unchanged.

[14] While the USPS will not forward ballots, a mail-in ballot is still more flexible than early in-person voting or in-person voting on election day.

4.3 Robustness

Our outcome variable potentially embeds two components: (1) deciding whether to vote and (2) conditional on voting, choosing which state to vote in. Only the second of these two would be considered CSSV. To isolate this second component, in Table 2 columns 3 and 4 we consider registrants who are predicted to vote in the 2020 presidential election with a very high probability ex-ante. In column 3 we select all individuals who voted in the 2020 primary election. This group has a 97.4% chance of voting in the general election. In column 4, we select individuals who have a high predicted probability of voting estimated from a probit model with demographics interacted with political donor status and voting history (see Appendix Figure A2). We use individuals with a predicted probability of voting at or above 99%.

This "identification at infinity" approach (d'Haultfoeuille and Maurel, 2013) in columns 3 and 4 reveals that our results in column 2 are robust. If anything, the estimated coefficients are larger among these near-certain voters. This could be due to two reasons. First, 25% of double-registrants do not vote at all, which will attenuate the coefficients in column 2. Second, near-certain voters might be more politically engaged and therefore more responsive to strategic incentives.[15]

4.4 Relative closeness of elections

Thus far we have used a binary measure for whether a state is swing and compared swing to non-swing states. However, not all swing states are equally close: Iowa (which went +8% for Trump) is much less of a swing state than Georgia (+0.2% Biden). Moreover, the difference in closeness between a double-registrant's pair of states will also vary. For example, Georgia could be paired with a similarly close swing state like Arizona (+0.3% Biden) or with a less-close swing state like Iowa. If the relative closeness of elections affects CSSV behavior, the incentive to vote in Georgia will be higher when it is paired with Iowa.

Figure 2 and Table 3 explore these effects. Panel (a) of Figure 2 plots the probability of voting in the first state against the first state "win gap," defined as the absolute value of the

[15] 2.7% of double-registrants are FEC donors; this rises to 4.9% among primary voters and to 7.7% among those who have a predicted probability of voting at or above 99%.

Democratic minus Republican win probability in prediction markets. As the first state's win gap becomes larger (i.e., the election becomes less close), individuals are less likely to vote there. Panel (b) changes the horizontal axis to be the difference in win gaps between the second and first state in each pair. While panel (a) has 51 possible values for the win gap (50 states plus DC), panel (b) has 2,332 possible values for the difference in win gaps (the number of ordered state pairs in our sample). There is a clear positive slope in panel (b), indicating that as an individual's incentive to vote in their first state increases (because the difference in win gaps between their second and first state is larger), they are more likely to vote in their first state. Panels (c) and (d) repeat these graphs, but residualize out the early voting and auto-mailed ballot variables of Table 2 from the voting probability. This reduces the dispersion of the observations around the fitted lines without changing the slopes.

Table 3 mirrors Table 2, but we replace the binary measures for swing states with two continuous variables: (i) the first state win gap (*first win gap*) and (ii) the difference in win gaps between the second and first state (*second – first win gap*). We find that a one standard deviation (30 pp) increase in the first state's predicted win gap (*first win gap*) reduces a double-registrant's likelihood of voting there by 0.9 pp (30×-0.031). A one standard deviation increase in *second – first win gap* (46 pp) increases the likelihood of voting in the first state by 2.4 pp. In other words, holding the first state win gap fixed, as the second state's election becomes less close, voters are more likely to vote in their first state. The coefficients on *auto-mailed ballot* and *early voting* are largely unaffected.

4.5 CSSV and double voting

A natural question is whether double-registrants double vote *for president*, since this is illegal under the Voting Rights Act of 1965 (52 U.S.C. §10307(e)).[16] We observe a mean rate of double voting of approximately 1% among double-registrants in our sample, amounting to 0.034% of voter registrations. Hence, we reach the same conclusion as Goel et al. (2020) that double voting is exceedingly rare.

Moreover, even this low rate is likely an over-estimate for two reasons. First, outcome

[16] National Conference of State Legislatures (2021) finds that there have only been a small number of cases of double voting prosecuted under state law and none under federal law.

means close to zero are sensitive to small data recording errors in this setting (Goel et al., 2020). While data recording errors could inflate the *level* of double voting, they should not inflate the coefficient estimates for incentives and costs, which are the main focus of the paper.[17]

Second, even if a voter did cast a ballot in two states, this need not imply double voting for president. Failing to vote for president despite turning in a ballot is not uncommon ("residual voting"). In fact, Stewart III et al. (2020) find that 1.4% of ballots were residual votes in the 2016 election. "Ballot roll-off" is a closely related concept, where someone votes for president but not for other items on the ballot. Because how an individual votes is confidential, in our data we only observe whether someone cast a ballot, but not which items they voted on. Thus, it is possible that a double voter casts a ballot in state A for president (but not for U.S. senator) and in state B for senator (but not for president).

We now investigate whether the same CSSV forces that affect which state to vote in – incentives and costs – influence the likelihood of double voting. We estimate equation 3 but change the outcome to be an indicator for voting in two states in the 2020 presidential election. Table 4 reports the results and finds evidence of CSSV. When only the first state is swing, the probability of double voting increases by 24% of the mean in column 2, and when both states are swing this rises to 43%. Double-registrants respond even more strongly to costs. Double voting increases by 104% if both states send out auto-mailed ballots. The effect is similar for *only first auto-mailed* (106%) and smaller for *only second auto-mailed* (19%), likely because cost considerations matter less in the state where double-registrants live.[18]

Columns 3 and 4 condition on near-certain voters, as in Table 2. The influence of incentives increases, with the coefficient in column 4 on *both swing* rising to 64% of the subsample mean. For cost-related variables, we see a decrease in effect sizes.

[17]Meyer and Mittag (2017) shows that a binary dependent variable with conditionally random measurement error attenuates coefficients in most situations.

[18]Appendix Table A4 shows that these estimates are robust to the same set of robustness checks as in Appendix Table A3.

5. Did CSSV affect the outcome of the 2020 election?

The prior section documents strong evidence of CSSV behavior in the 2020 election. To quantify the number of votes involved, columns 1 and 2 of Table 5 provide back-of-the-envelope estimates of the total number of CSSV votes flowing in and out of each of the 13 swing states. We do this by estimating the number of double-registrants using the same strategy as in Section 3, calculating the number of votes that flow in and out of each pair of states using estimated CSSV coefficients, and summing the inflows and outflows for each swing state. To put these numbers in perspective, column 3 lists the actual vote margins in the 2020 election. In some states, the inflows and outflows are of the same order of magnitude as the actual victory margins.

However, whether this behavior affected the outcome of the election depends on whether one party engaged in CSSV more than the other in swing states. The overall effect depends on (i) the share of Democrats vs. Republicans among double-registrants and their distribution across states and (ii) the intensity of CSSV behavior among Democrats vs. Republicans. Ultimately, what matters is the party composition of both inflows and outflows in swing states. For example, if the inflows were all votes for Biden and the outflows were all votes for Trump, then the elections in Arizona and Georgia would have flipped to Trump in the absence of CSSV. However, this is an extreme example, and in reality each flow will include votes for both candidates. Indeed, the calculated inflows and outflows in columns 1 and 2 use party-specific estimates for the number of double-registrants and for the CSSV coefficients.

In Section 2 we showed that there are approximately twice as many Democratic as Republican double-registrants. Appendix Table A5 runs the regression from Table 2 column 2 separately for Democrats, Republicans, and Independents. While voters from all parties display CSSV behavior, an F-test confirms that the coefficients are different (p value<0.01). The coefficients on *only first swing* and *only second swing* are larger in absolute value among Republicans than Democrats, while the converse is true for the auto-mail coefficients.[19] In other words, Republicans are more responsive to incentives, while Democrats are more re-

[19] While the *both swing* (*both auto-mail*) coefficient is bigger (smaller) for Democrats, there are few voters whose states are both swing, which means they have little effect on the 2020 election.

sponsive to costs.

The remaining columns in Table 5 estimate the number of votes gained or lost by each party's candidate and the net effect. To translate party affiliation into votes for each candidate, we use exit polls (Schaffner et al., 2021). Aggregating these incentive effects across all state pairs, we report the net number of votes flowing into each swing state for the candidates in columns 4-5 and the difference between them (Dem–Rep) in column 6. For example, in Georgia, the Democratic candidate (Biden) gained 2,947 votes due to incentive-based CSSV, while the Republican candidate (Trump) gained 2,150, for a net difference of +798. Biden gained more incentive-based CSSV votes despite Republicans' greater responsiveness to incentives because there are more Democratic double-registrants. In contrast, the net difference for cost-based CSSV is -1,389. Since Georgia was not an auto-mail state, Democrats' higher responsiveness to this cost measure combined with their greater numbers disproportionately pulled Biden votes out of Georgia and into auto-mail states. Combining the effects, CSSV added only 591 votes in favor of Trump in Georgia on net, despite there being 15,836 (=10,209+5,627) total CSSV-induced votes.

These two patterns observed in Georgia are present more generally across the other swing states, none of which had auto-mailed ballots except Nevada. First, incentive-based CSSV generally favored Biden because the larger number of Democratic double-registrants more than offset the higher responsiveness of Republicans to incentives. Second, cost-based CSSV generally favored Trump by disproportionately pulling votes away from Biden in swing states both because more Democrats reside in auto-mail states and because Democrats are more responsive to costs. In summary, there were a large number of CSSV votes in 2020 – 217,000 of which involve at least one swing state – but the *net* effect was not pivotal, because Republican and Democratic CSSV-induced votes largely cancelled out. However, CSSV could be important in closer elections such as in 2000, where the gap was 537 votes in Florida and 366 in New Mexico.

6. Conclusion

U.S. states administer federal elections independently, giving rise to the possibility of double-registered voters. We estimate there are 6.1 million of these individuals and that they respond to incentives and costs, voting more often in swing states (higher incentive) and states that automatically send out mail-in ballots (lower cost).

Double-registration is relevant to long-standing discussions around election policies and campaign behavior. Reforms targeting double-registrants may reduce their number, but at a potential cost to political enfranchisement and state autonomy.[20] As another example, automatic mailing of ballots is often proposed as a policy to increase voter participation. Our findings point to an unintended effect: making voting easier in one state pulls votes away from other states via CSSV. This was evident in 2020, as auto-mail states – which are disproportionately Democratic – drew votes away from Biden in swing states. Absent reforms, however, campaigns could choose to target double-registrants (or potential double-registrants) in swing states to increase the impact of CSSV.

Regardless of how CSSV interacts with election policy or campaign strategy, the fact that some individuals are choosing where to vote violates the principle of one person, one vote by providing some Americans more voting power than others. Moreover, the voters who receive this additional power are not random: double-registrants are 3.5 times as prevalent in the wealthiest 1% of zip codes compared to the bottom quartile. This fact is not only antithetical to the progress made by courts and legislators to decouple voting power from wealth but also concerning because redistributive policies are often on the ballot.

[20] Even if states could correctly identify double-registrants via cross-state information sharing, disenfranchisement is a risk because identifying an individual's domicile is difficult. Some states participate voluntarily in the Electronic Registration Information Center (ERIC) which collects their voter registrations and, upon request, sends reports to member states.

References

Allcott, H., Braghieri, L., Eichmeyer, S., and Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, 110(3):629–76.

Alvarez, R. M., Boehmke, F. J., and Nagler, J. (2006). Strategic voting in British elections. *Electoral Studies*, 25(1):1–19.

Anagol, S. and Fujiwara, T. (2016). The runner-up effect. *Journal of Political Economy*, 124(4):927–991.

Bechtel, M. M., Hangartner, D., and Schmid, L. (2018). Compulsory voting, habit formation, and political participation. *Review of Economics and Statistics*, 100(3):467–476.

Bernstein, A., Billings, S. B., Gustafson, M., and Lewis, R. (2022). Partisan residential sorting on climate change risk. *Journal of Financial Economics*.

Billings, S. B., Chyn, E., and Haggag, K. (2021). The long-run effects of school racial diversity on political identity. *American Economic Review: Insights*, 3(3):267–84.

Blais, A. (2006). What affects voter turnout? *Annual Review of Political Science*, 9:111.

Braconnier, C., Dormagen, J.-Y., and Pons, V. (2017). Voter registration costs and disenfranchisement: experimental evidence from France. *American Political Science Review*, 111(3):584–604.

Brown, J. R. and Enos, R. D. (2021). The measurement of partisan sorting for 180 million voters. *Nature Human Behaviour*, pages 1–11.

Bursztyn, L., Cantoni, D., Funk, P., Schönenberger, F., and Yuchtman, N. (2017). Identifying the effect of election closeness on voter turnout: Evidence from Swiss referenda.

Cantoni, E. (2020). A precinct too far: Turnout and voting costs. *American Economic Journal: Applied Economics*, 12(1):61–85.

Cantoni, E. and Pons, V. (2021). Strict ID laws don't stop voters: Evidence from a US nationwide panel, 2008–2018. *Quarterly Journal of Economics*, 136(4):2615–2660.

Clark, T. S., Montagnes, B. P., and Spenkuch, J. L. (2022). Politics from the bench? Ideology and strategic voting in the US Supreme Court. *Journal of Public Economics*, 214:104726.

Cornell Law Dictionary (2022). Flush with cash, Biden eclipses trump in war for the airwaves. https://www.law.cornell.edu/wex/one-person_one-vote_rule#:~:text=One%2Dperson%2C%20one%2Dvote%20refers%20to%20the%20rule%20that,person%27s%20within%20the%20same%20state. (accessed May 8, 2023).

Cottrell, D., Herron, M. C., and Westwood, S. J. (2018). An exploration of donald trump's allegations of massive voter fraud in the 2016 general election. *Electoral Studies*, 51:123–142.

DellaVigna, S., List, J. A., Malmendier, U., and Rao, G. (2016). Voting to tell others. *The Review of Economic Studies*, 84(1):143–181.

Duverger, M. (1954). *Political Parties*. Wiley, New York.

d'Haultfoeuille, X. and Maurel, A. (2013). Another look at the identification at infinity of sample selection models. *Econometric Theory*, 29(1):213–224.

Eggers, A. C., Garro, H., and Grimmer, J. (2021). No evidence for systematic voter fraud: A guide to statistical claims about the 2020 election. *Proceedings of the National Academy of Sciences*, 118(45):e2103619118.

Farquharson, R. (1969). *Theory of Voting*. Yale Univ. Press, New Haven.

Federal Voting Assistance Program (2022). How to Determine Your Voting Residency. https://www.fvap.gov/info/laws/voting-residence (accessed Dec 15, 2022).

Ferlenga, F. and Knight, B. G. (2022). Vote early and vote often? Detecting electoral fraud from the timing of 19th century elections. *NBER working paper*.

FiveThirtyEight (2020). Statewide opinion polling for the 2020 United States presidential election. https://en.wikipedia.org/wiki/Statewide_opinion_polling_for_the_2020_United_States_presidential_election (accessed April 22, 2022).

Fujiwara, T. (2011). A regression discontinuity test of strategic voting and Duverger's law. *Quarterly Journal of Political Science*, 6(3–4):197–233.

Fujiwara, T. (2015). Voting technology, political responsiveness, and infant health: Evidence from Brazil. *Econometrica*, 83(2):423–464.

Fujiwara, T., Meng, K., and Vogl, T. (2016). Habit formation in voting: Evidence from rainy elections. *American Economic Journal: Applied Economics*, 8(4):160–88.

Gentzkow, M., Shapiro, J. M., and Sinkinson, M. (2011). The effect of newspaper entry and exit on electoral politics. *American Economic Review*, 101(7):2980–3018.

Gerber, A. S., Huber, G. A., and Hill, S. J. (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.

Goel, S., Meredith, M., Morse, M., Rothschild, D., and Shirani-Mehr, H. (2020). One person, one vote: Estimating the prevalence of double voting in US presidential elections. *American Political Science Review*, 114(2):456–469.

Greabe, J. M. (2012). A federal baseline for the right to vote. *Colum. L. Rev. Sidebar*, 112:62.

Gronke, P., Galanes-Rosenbaum, E., Miller, P. A., and Toffey, D. (2008). Convenience voting. *Annual Review of Political Science*, 11:437–455.

Hodler, R., Luechinger, S., and Stutzer, A. (2015). The effects of voting costs on the democratic process and public finances. *American Economic Journal: Economic Policy*, 7(1):141–71.

Kaplan, E. and Yuan, H. (2020). Early voting laws, voter turnout, and partisan vote composition: Evidence from Ohio. *American Economic Journal: Applied Economics*, 12(1):32–60.

Kawai, K. and Watanabe, Y. (2013). Inferring strategic voting. *American Economic Review*, 103(2):624–62.

Marist Poll (October 2021). NPR/PBS NewsHour/Marist Poll. https://maristpoll.marist.edu/wp-content/uploads/2021/10/NPR_PBS-NewsHour_Marist-Poll_USA-NOS-and-Tables_B_202110251104.pdf (accessed May 3, 2022).

Meyer, B. D. and Mittag, N. (2017). Misclassification in binary choice models. *Journal of Econometrics*, 200(2):295–311.

Minnite, L. C. (2010). *The Myth of Voter Fraud*. Cornell University Press.

MIT Election Data and Science Lab (2020). County Presidential Election Returns 2000-2020.

MIT Election Lab (2021). Voter Confidence. https://electionlab.mit.edu/research/voter-confidence (accessed Sept. 19, 2022).

Myerson, R. B. and Weber, R. J. (1993). A theory of voting equilibria. *American Political Science Review*, 87(1):102–114.

National Conference of State Legislatures (2020a). Absentee and mail voting policies in effect for the 2020 election. https://web.archive.org/web/20201031183259/https://www.ncsl.org/research/elections-and-campaigns/absentee-and-mail-voting-policies-in-effect-for-the-2020-election.aspx (October 29, 2020, accessed via Wayback Machine).

National Conference of State Legislatures (2020b). State laws governing early voting. https://web.archive.org/web/20201031202754/https://www.ncsl.org/research/elections-and-campaigns/early-voting-in-state-elections.aspx (October 22, 2020, accessed via Wayback Machine).

National Conference of State Legislatures (2021). Double voting. https://www.ncsl.org/research/elections-and-campaigns/double-voting.aspx (accessed September 9, 2022).

New York Times (2020). Flush with cash, Biden eclipses trump in war for the airwaves. https://www.nytimes.com/interactive/2020/10/17/us/politics/trump-biden-campaign-ad-spending.html (accessed April 22, 2022).

New York Times (2022). In voter fraud, penalties often depend on who's voting. https://www.nytimes.com/2022/09/07/us/voter-fraud-penalties.html (accessed September 9, 2022).

Ortoleva, P. and Snowberg, E. (2015). Overconfidence in political behavior. *American Economic Review*, 105(2):504–35.

Pew (2018). Commercial voter files and the study of US politics. https://www.pewresearch.org/methods/2018/02/15/commercial-voter-files-and-the-study-of-u-s-politics/ (accessed June 15, 2022).

Piketty, T. and Saez, E. (2003). Income inequality in the United States, 1913–1998. *Quarterly Journal of Economics*, 118(1):1–41.

Pons, V. and Tricaud, C. (2018). Expressive voting and its cost: Evidence from runoffs with two or three candidates. *Econometrica*, 86(5):1621–1649.

PredictIt (2020). State-Level U.S. Presidential Election Prediction Markets. https://www.predictit.org/markets/17/US-Elections (accessed April 4, 2022).

Schaffner, B., Ansolabehere, S., and Luks, S. (2021). Cooperative Election Study Common Content, 2020.

Smith, M., Zidar, O., and Zwick, E. (2023). Top wealth in America: New estimates under heterogeneous returns. *Quarterly Journal of Economics*, 138(1):515–573.

Spenkuch, J. L. (2015). Please don't vote for me: Voting in a natural experiment with perverse incentives. *Economic Journal*, 125(585):1025–1052.

Spenkuch, J. L. (2018). Expressive vs. strategic voters: An empirical assessment. *Journal of Public Economics*, 165:73–81.

Spenkuch, J. L., Teso, E., and Xu, G. (2023). Ideology and performance in public organizations. *Econometrica (forthcoming).*

Spenkuch, J. L. and Toniatti, D. (2018). Political advertising and election results. *Quarterly Journal of Economics*, 133(4):1981–2036.

Stewart III, C., Alvarez, R. M., Pettigrew, S. S., and Wimpy, C. (2020). Abstention, protest, and residual votes in the 2016 election. *Social Science Quarterly*, 101(2):925–939.

Yoder, J., Handan-Nader, C., Myers, A., Nowacki, T., Thompson, D. M., Wu, J. A., Yorgason, C., and Hall, A. B. (2021). How did absentee voting affect the 2020 US election? *Science Advances*, 7(52):eabk1755.

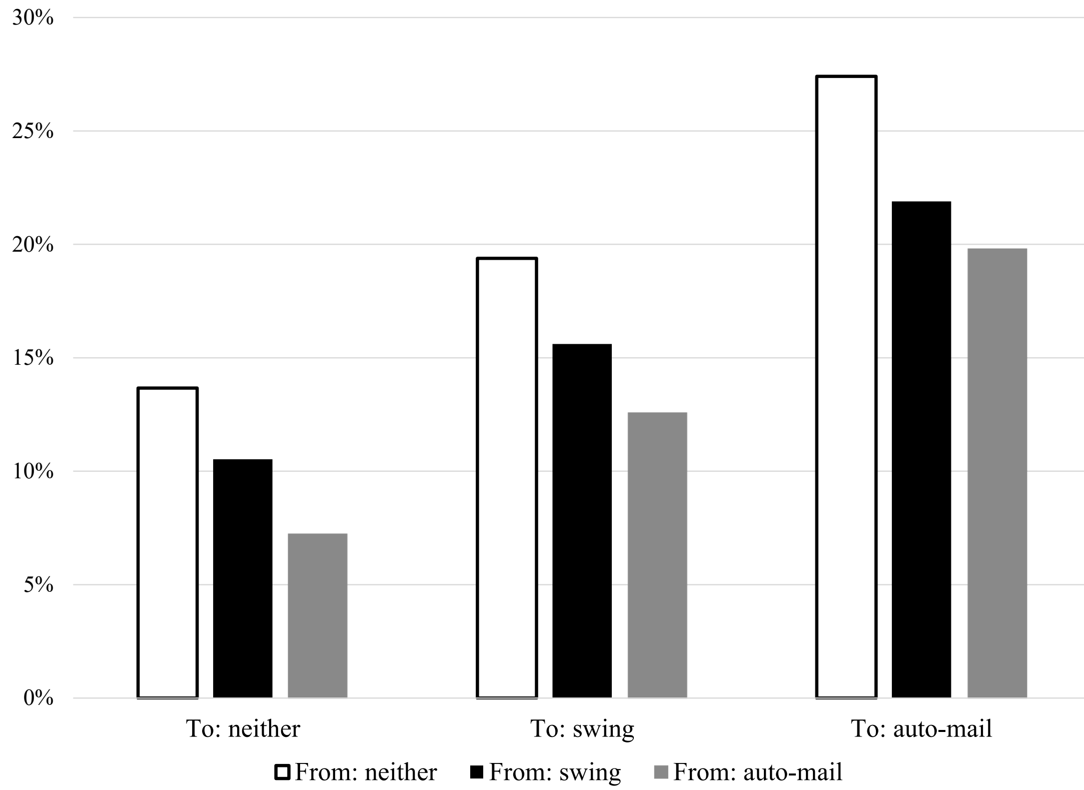
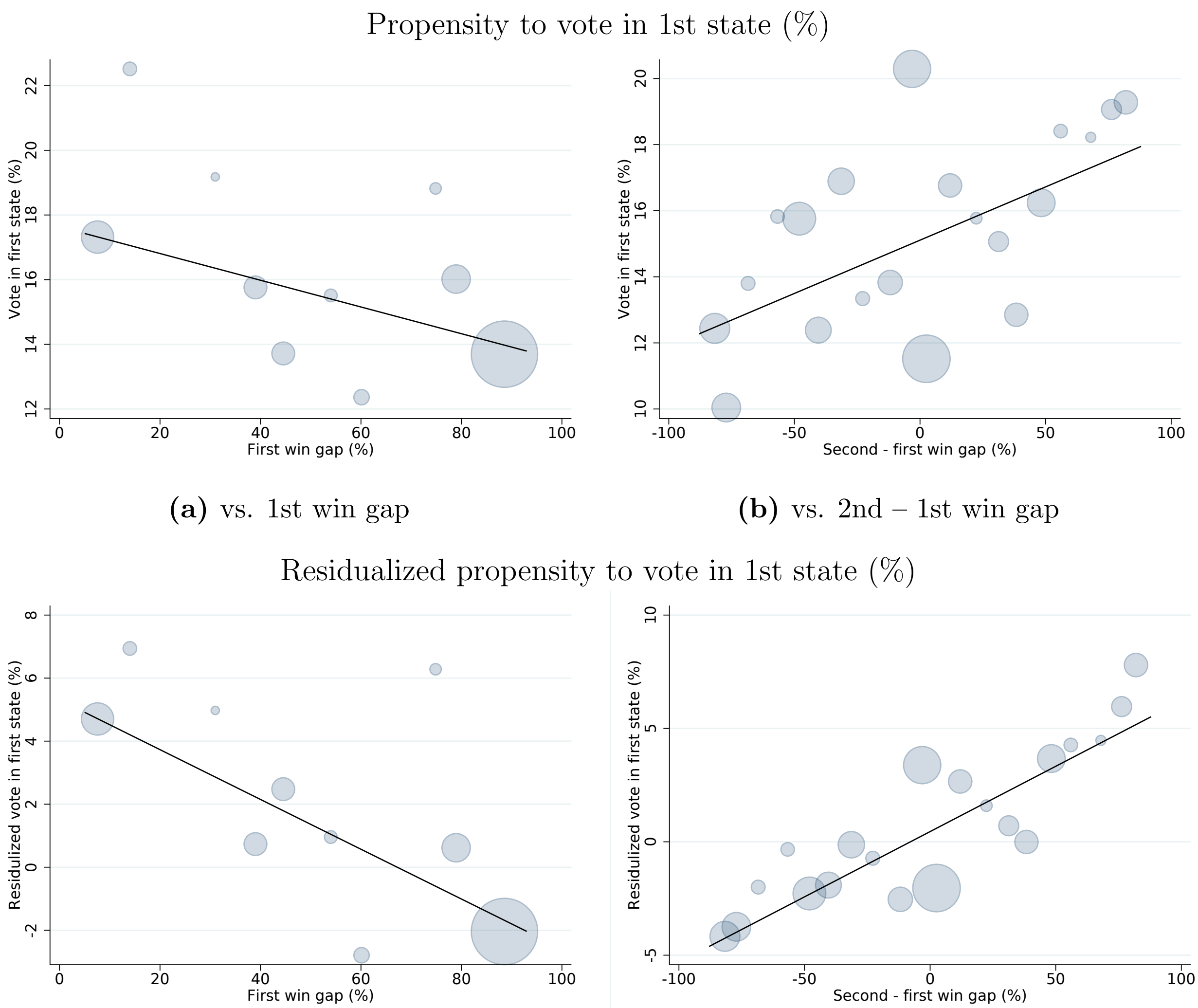


Figure 1: Vote Flows from Second to First State of Voter Registration

Note: Vote flows are defined as the number of double-registrants who vote in their first state divided by the number of double-registered voters in our sample for each first-state × second-state characteristic combination. Characteristic combinations are the interaction of *swing*, *auto-mail*, and *neither* (i.e., not swing or auto-mail) between the first and second state in a pair. *From:* refers to voters' second state of registration; *To:* refers to their first state. *Swing* is an indicator for the 13 states where the probability of winning for any candidate on October 3, 2020 did not exceed 80 percent (PredictIt, 2020), and *auto-mail* denotes states that automatically sent mail-in ballots to registered voters in 2020 (National Conference of State Legislatures, 2020b). Nevada is both swing and auto-mail; to create mutually exclusive categories we classify it as swing in this figure only.



(a) vs. 1st win gap

(b) vs. 2nd – 1st win gap

(c) vs. 1st win gap

(d) vs. 2nd – 1st win gap

Figure 2: Voting in First State of Registration and Election Closeness

Note: These are binscatters of a double-registrant's probability of voting in their first state plotted against (i) the win gap in their first state, i.e., the predicted closeness of an election (panels a and c), and (ii) the difference in the win gap between their second and first states, i.e., the relative closeness of the elections in a pair (panels b and d). Win gap is defined as the absolute value of the Democratic minus Republican win probability in prediction markets on Oct 3, 2020 (PredictIt, 2020). Panels a and b plot the raw vote probabilities on the y-axis, while panels c and d residualize out the auto-mailed ballot and early voting variables in Table 3. Dot sizes are proportional to the number of observations.

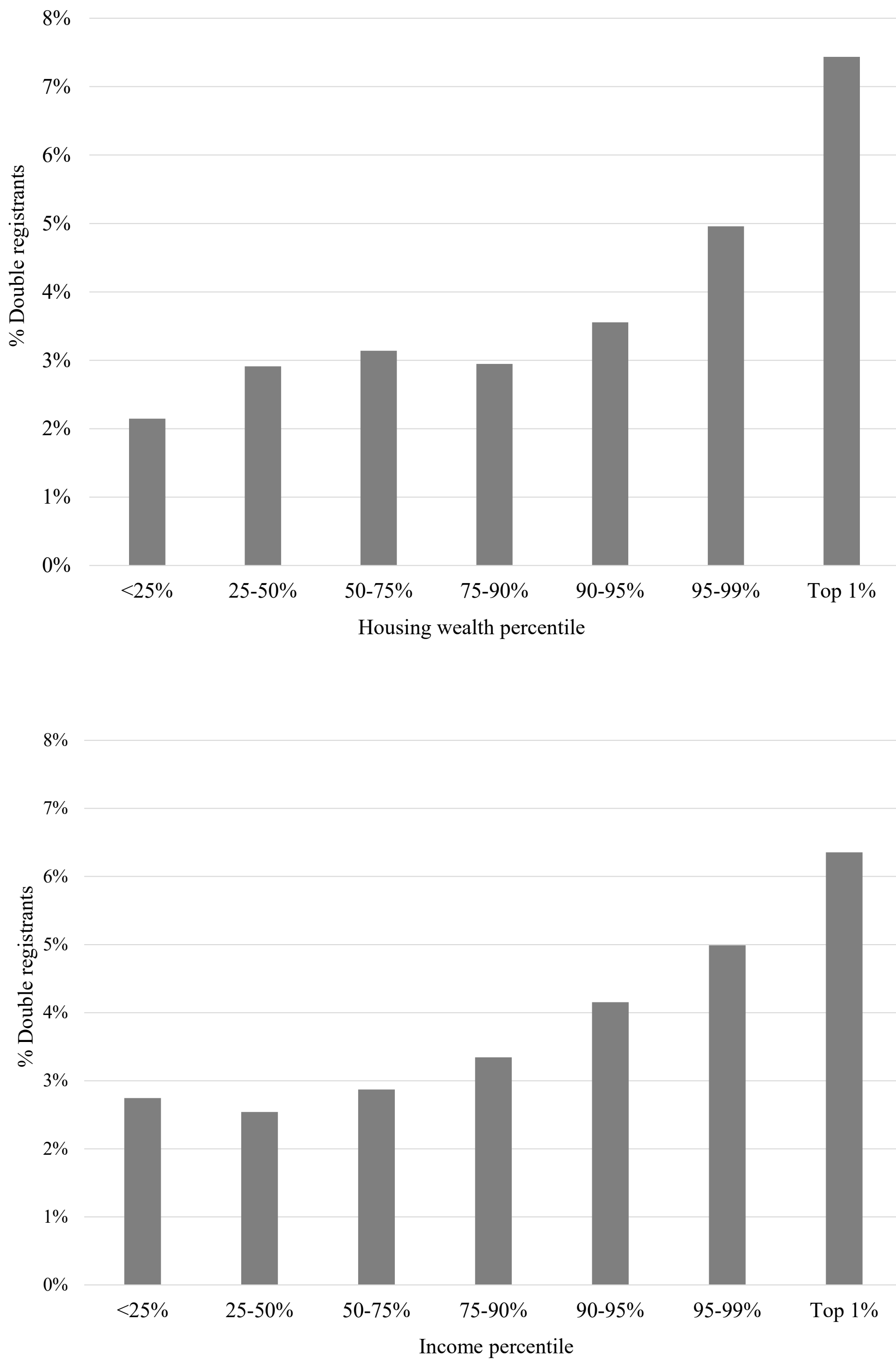


Figure 3: Double registration rates by housing wealth and income

Note: These bar charts report the share of double registrants by housing wealth and income bins. Housing wealth is proxied for using Zillow zip code-level data to assign each voter to a housing value bin. The bottom chart measures income using IRS adjusted gross income by zip code. The first three bars in each chart represent the first three quartiles, and the subsequent bars break out the top quartile.

Table 1: Summary Statistics for Registered Individuals

	All registrants	Registrants with cell phone & other info	Known double registrants
2020 presidential election (%)			
Vote	74.4	78.1	81.3
Vote in first state	-	-	14.9
Voting history (%)			
Vote in general elections	64.0	65.6	66.6
Vote in primary elections	23.0	23.7	22.3
Vote in minor elections	8.7	8.8	6.5
Party & donations (%)			
Republican	31.1	32.7	21.4
Democrat	41.0	41.6	46.3
Independent	28.0	25.6	32.2
FEC donor	1.5	1.8	2.7
Demographics (%)			
Male	47.0	46.4	52.2
White	70.9	70.4	70.8
Hispanic	13.6	14.4	13.9
Black	11.8	11.7	10.6
Asian	3.7	3.5	4.7
Low wealth	32.3	30.2	21.4
Middle wealth	49.3	49.6	51.2
High wealth	18.4	20.2	27.4
Low income	22.7	22.1	18.7
Middle income	57.6	57.1	51.5
High income	19.7	20.8	29.8
Married	40.8	43.5	42.1
Any children	35.5	38.4	44.6
Homeowner	63.8	66.8	49.5
Born pre-1955	22.9	22.7	12.9
Born 1955-64	17.7	19.7	13.0
Born 1965-74	16.0	18.2	15.7
Born 1975-84	15.6	17.7	22.2
Born 1985-94	16.8	15.5	31.0
Born post-1994	10.9	6.3	5.3
Observations	202,535,296	90,234,280	590,991

Note: Column 1 reports summary statistics for all registrants as of October 2020; column 2 for registrants with necessary information for matching individuals across states, where the main constraint is the availability of mobile phone numbers; and column 3 for registrants who we identify as double-registrants. *Vote in first state* is an indicator for voting in the first (earlier) state of registration. Voting history refers to the share of even-year general, even-year primary, and minor elections that an individual voted in before 2020. In column (3) party affiliation is based on the second (later) state of registration. *FEC donor* is an indicator for making at least one FEC donation by 2020. *Low wealth*, *middle wealth*, and *high wealth* denote zip codes where Zillow monthly average home value index for single-family houses in 2020 is below $200,000, $200,000 to $500,000, and above $500,000, respectively. *Low income*, *middle income*, and *high income* denote zip codes where average annual adjusted gross income from IRS data is below $50,000, $50,000 to $99,999, and above $100,000, respectively.

Table 2: How Incentives and Costs Affect Voting in First State of Registration

	Dependent var.: Vote in first state of registration			
	(1)	(2)	(3)	(4)
	All double-registrants		Near certain general election voters	
			Voter in primary	Predicted voter Pr≥0.99
Only first swing	4.628***	4.560***	5.645***	7.942***
	(0.807)	(0.758)	(0.919)	(1.351)
Both swing	1.582*	1.299*	0.234	0.739
	(0.956)	(0.724)	(1.056)	(1.461)
Only second swing	-2.947***	-2.917***	-5.749***	-6.397***
	(1.045)	(0.964)	(1.976)	(2.366)
Only first auto-mailed ballot	10.762***	9.251***	17.600***	16.049***
	(1.629)	(1.595)	(2.571)	(3.438)
Both auto-mailed ballot	6.906***	6.124***	4.894*	4.917
	(1.558)	(1.463)	(2.809)	(3.314)
Only second auto-mailed ballot	-5.870***	-5.544***	-11.777***	-10.400***
	(0.730)	(0.652)	(1.247)	(1.581)
Only first early voting		4.105***	5.828***	7.062***
		(0.969)	(1.608)	(1.982)
Both early voting		2.851***	3.422***	1.352
		(0.560)	(1.010)	(1.387)
Only second early voting		-0.108	0.093	-2.573**
		(0.669)	(1.045)	(1.284)
Observations	590,991	590,991	174,595	63,536
R^2	0.018	0.021	0.058	0.055
Outcome mean	14.85	14.85	18.57	18.97

Note: The variables *only first ...*, *both ...*, and *only second ...* indicate whether the first state, both the first and the second state, and only the second state have specific features: swing status, auto-mailed ballots, and early voting. *Swing* is an indicator for the 13 states whose PredictIt winning likelihood for either party on October 3, 2020 did not exceed 80% (PredictIt, 2020); *auto-mailed ballot* denotes states that automatically sent mail-in ballots to registrants in 2020 (National Conference of State Legislatures, 2020b); *early voting* refers to states that allowed voting for at least 19 days before Nov 3, 2020 (the sample median among those with early voting) (National Conference of State Legislatures, 2020a). All independent variables are divided by 100 for presentation purposes. Columns (1)-(2) use all double-registrants in the sample; column (3) double-registrants who voted in a 2020 primary election; column (4) double-registrants whose predicted likelihood of voting in the 2020 presidential election is 0.99 or higher (see note to Figure A2 for details). Standard errors are clustered at the state-pair level.
*** 1%, ** 5%, * 10% significance level

Table 3: RELATIVE CLOSENESS OF ELECTIONS AND CSSV BEHAVIOR

	Dependent var.: Vote in first state of registration
First win gap	-0.031**
	(0.013)
Second – first win gap	0.052***
	(0.012)
Only first auto-mailed ballot	9.668***
	(1.577)
Both auto-mailed ballot	6.209***
	(1.474)
Only second auto-mailed ballot	-5.824***
	(0.652)
Only first early voting	4.042***
	(0.960)
Both early voting	2.875***
	(0.590)
Only second early voting	-0.134
	(0.662)
Outcome mean (%)	14.85
Observations	590,991
R^2	0.022

Note: This table reports the relationship between a double-registered individual's probability of voting in the first state and the closeness of the election in the first state, as well as the difference in the closeness between the first and second states. *First win gap* refers to the difference in the PredictIt winning likelihood between the Democratic and the Republican parties in the first state on Oct 3, 2020, and *second – first win gap* refers to the difference in the win gaps between the first and the second states. Auto-mailed ballot and early voting variables are divided by 100 for presentation purposes. Other variable definitions and sample selection follow those in Table 2 column 2. Standard errors are clustered at the state-pair level.
*** 1%, ** 5%, * 10% significance level

Table 4: How Incentives and Costs Affect Voting in Both States of Registration

	Dependent var.: Vote in both states of registration			
	(1)	(2)	(3)	(4)
	All double-registrants		Near certain general election voters	
			Voter in primary	Predicted voter Pr≥0.99
Only first swing	0.244**	0.238**	0.610***	1.626***
	(0.112)	(0.109)	(0.215)	(0.404)
Both swing	0.441***	0.433***	1.050***	1.893***
	(0.168)	(0.167)	(0.403)	(0.684)
Only second swing	0.080	0.078	0.088	0.112
	(0.123)	(0.120)	(0.266)	(0.452)
Only first auto-mailed ballot	1.157***	1.103***	1.581***	2.501***
	(0.192)	(0.191)	(0.379)	(0.609)
Both auto-mailed ballot	1.102***	1.082***	0.810**	0.584
	(0.215)	(0.216)	(0.317)	(0.540)
Only second auto-mailed ballot	0.178*	0.194*	-0.238	-0.359
	(0.107)	(0.105)	(0.184)	(0.365)
Only first early voting		0.115	0.115	0.113
		(0.128)	(0.271)	(0.471)
Both early voting		0.106	0.054	0.008
		(0.153)	(0.312)	(0.573)
Only second early voting		-0.046	-0.076	-0.370
		(0.122)	(0.260)	(0.469)
Observations	590,991	590,991	174,595	63,536
R^2	0.002	0.002	0.002	0.004
Outcome mean	1.04	1.04	1.92	3.35

Note: All independent variables and sample selections follow those in the corresponding columns in Table 2, including dividing independent variables by 100 for presentation purposes. Standard errors are clustered at the state-pair level.
*** 1%, ** 5%, * 10% significance level

Table 5: Effect of Cross-State Strategic Voting on 2020 Swing States

	CSSV induced votes		Actual 2020 vote margin	Estimated 2020 vote gain for Democratic and Republican presidential candidates						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
				Incentive			*Cost* (auto-mail & early voting)			*Total*
	Inflow	Outflow	Win-Loss	Dem	Rep	Dem-Rep	Dem	Rep	Dem-Rep	Dem-Rep
AZ	5,307	5,506	10,457	1,466	1,610	-144	-1,834	-1,453	-382	-525
GA	10,209	5,627	11,779	2,947	2,150	798	-1,008	382	-1,389	-591
WI	3,091	2,598	20,682	1,500	1,296	205	-1,239	-1,090	-149	56
NV	10,623	5,750	33,596	1,703	1,742	-38	979	202	777	739
NH	1,279	508	59,267	723	490	233	-306	-163	-142	91
NC	7,650	3,584	74,481	2,210	2,178	32	-466	-16	-450	-418
PA	24,711	14,497	80,555	8,293	5,724	2,570	-3,878	-481	-3,397	-828
IA	4,458	1,962	138,611	1,261	1,128	133	-46	25	-71	61
MI	5,569	2,356	154,188	1,503	1,084	419	-57	420	-477	-58
MN	1,344	855	233,012	250	276	-27	-128	69	-197	-224
FL	29,607	28,301	371,686	13,100	13,855	-755	-13,128	-12,910	-217	-972
OH	13,295	5,550	475,669	3,317	2,785	533	823	1,520	-697	-164
TX	20,132	24,610	631,221	8,941	8,869	72	-13,517	-8,698	-4,819	-4,747

Note: This table uses estimates of the number of double-registrants by party for each ordered state pair, estimates of the party-specific CSSV coefficients from Table A5, and exit polls for how party affiliation translates into votes for a candidate (Schaffner et al., 2021). As an example, we estimate that there are 5,415 (Florida, New York) Republican double-registrants, 22,064 (Florida, New York) Democrats, and 10,884 (Florida, New York) Independents using the procedure in Section 3. We then combine these numbers with the CSSV coefficients in Table A5 to calculate the number of votes that flow in and out of this ordered state pair. Since the first state (Florida) is the only swing state in the pair, we calculate 304 ($= 5,415 \times 5.617/100$) votes from Republican double-registrants, 867 votes from Democrats, and 521 votes from Independents flowed out of New York and into Florida. Exit polls in New York indicate that 87.5% of Republicans voted for Trump and 11.5% for Biden, so the 304 incentive-driven votes among (Florida, New York) Republican double-registrants led to a 266 vote gain for Trump and 35 vote gain for Biden in Florida, with the same votes lost in New York. A similar calculation using exit-poll data yields vote flows for Democratic and Independent double-registrants. The table performs these calculations for CSSV-induced votes for all ordered state-pairs and aggregates them. Columns 4 to 6 report the party-specific net inflow minus outflow for incentive-based CSSV; columns 7 to 9 do this for cost-based CSSV. Column 10 sums columns 6 and 9.

Online Appendix

Cross-State Strategic Voting

Gordon B. Dahl, Joseph Engelberg, Runjing Lu, and William Mullins

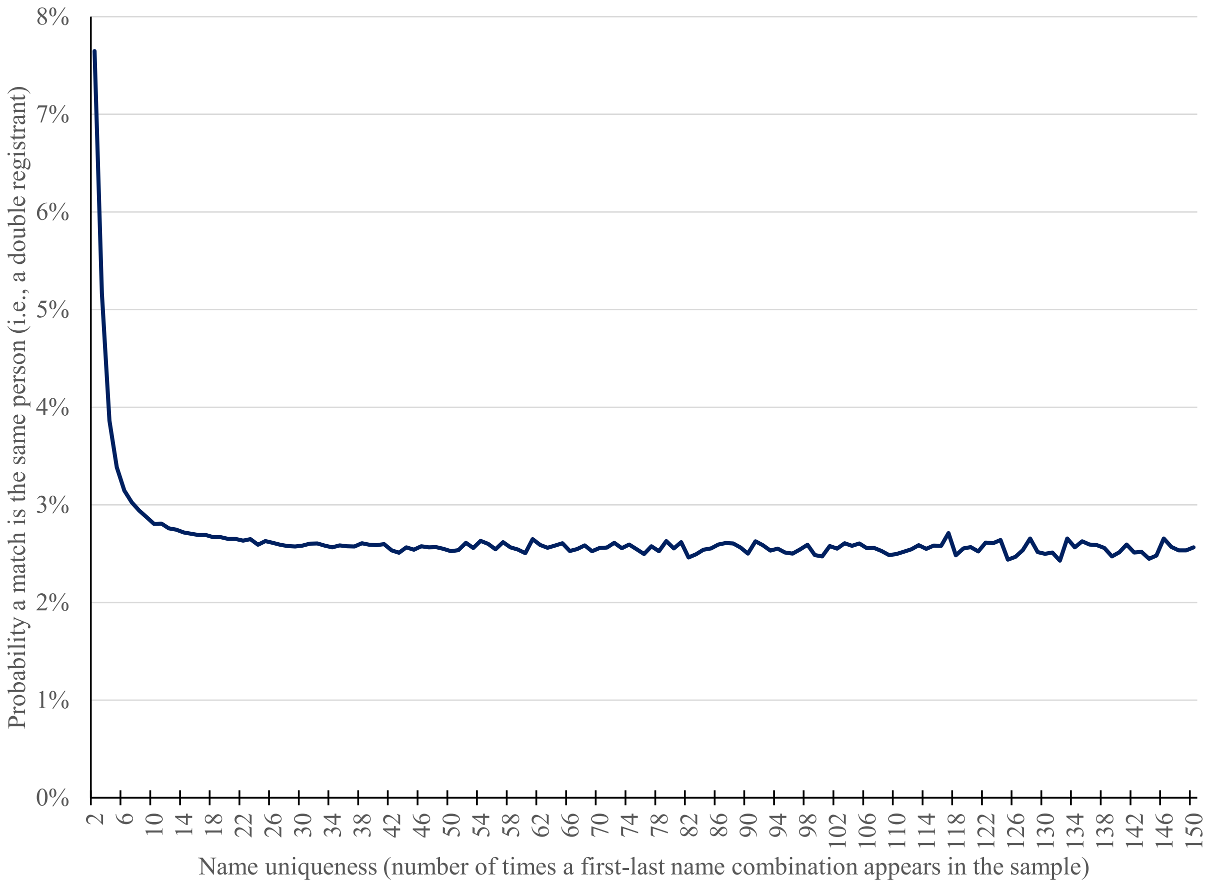


Figure A1: Probability a Name Match Across States is the Same Person by Level of Name Uniqueness

Note: This figure plots the estimated probability that two records in different states with matching first and last names correspond to the same person, estimated separately for each level of name uniqueness. Name uniqueness (NU) is the total number of times a (first name, last name) combination occurs in our sample of 2020 U.S. voter registration records. Section 3 describes the estimation procedure.

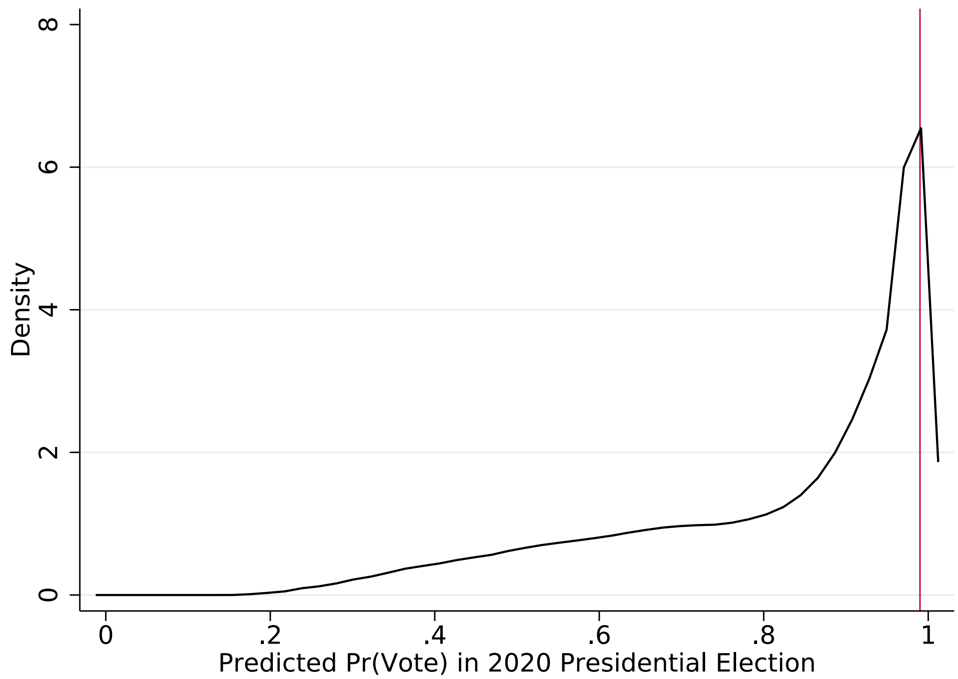


Figure A2: Predicted Probability of Voting in the 2020 Presidential Election

Note: This figure plots the kernel density of double-registrants' predicted likelihood of voting in the 2020 presidential election. The likelihood is estimated from a probit model in our double-registrant sample using as predictors the following variables: occupation, work industry, and demographics (gender, race, birth cohort, marital status, presence of children in the home, homeownership, as well as zip code-level housing wealth and income groups) interacted with political donor status and voting history in even-year general and primary elections (2008-2020). The vertical red line displays the cutoff used in Tables 2 and 4: a predicted voting probability of 0.99.

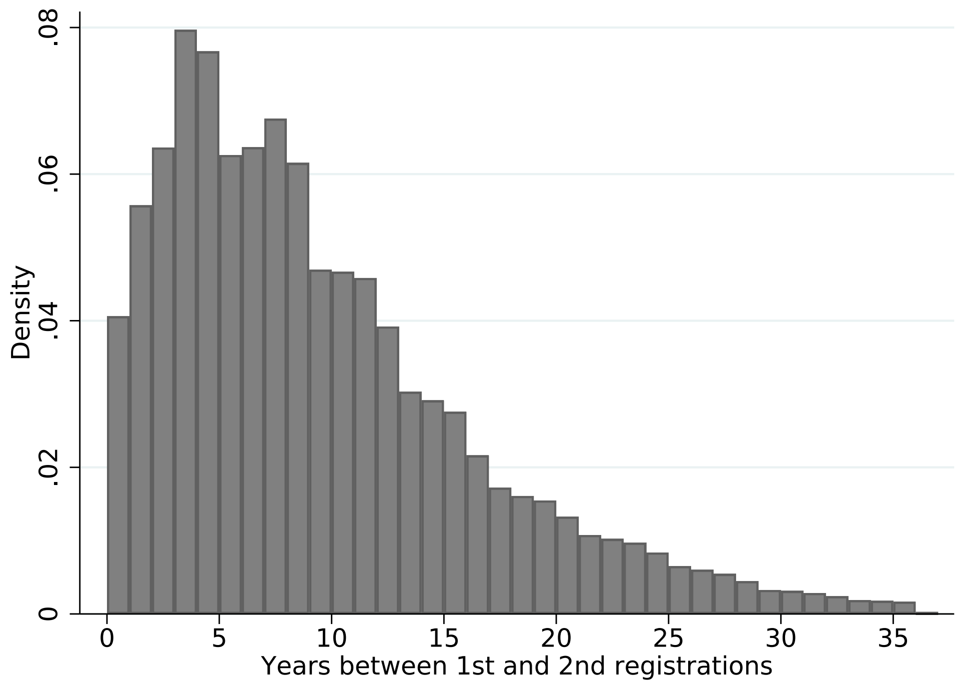


Figure A3: Number of Years between First and Second Registrations

Note: This figure plots the number of years between when an individual registers in their first and second states in our sample.

Table A1: State Election Characteristics

State	Swing	Auto-mailed	Auto absentee	Early voting	Alternate swing definitions		
	PredictIt	ballot	application		NYT	Ex-ante	Ex-post
AK	N	N	N	N	N	Y	N
AL	N	N	N	Y	N	N	N
AR	N	N	N	N	N	N	N
AZ	Y	N	N	Y	Y	Y	Y
CA	N	Y	N	Y	N	N	N
CO	N	Y	N	N	N	N	N
CT	N	N	Y	N	N	N	N
DC	N	Y	N	N	N	N	N
DE	N	N	Y	N	N	N	N
FL	Y	N	N	N	Y	Y	Y
GA	Y	N	N	Y	Y	Y	Y
HI	N	Y	N	N	N	N	N
IA	Y	N	Y	Y	Y	Y	Y
ID	N	N	N	N	N	N	N
IL	N	N	Y	Y	N	N	N
IN	N	N	N	Y	N	N	N
KS	N	N	N	Y	N	N	N
KY	N	N	N	N	N	N	N
LA	N	N	N	N	N	N	N
MA	N	N	Y	N	N	N	N
MD	N	N	Y	N	N	N	N
ME	N	N	N	Y	N	N	Y
MI	Y	N	Y	Y	Y	Y	Y
MN	Y	N	N	Y	Y	Y	Y
MO	N	N	N	N	N	Y	N
MS	N	N	N	N	N	N	N
MT	N	Y	N	Y	N	Y	N
NC	Y	N	N	Y	Y	Y	Y
ND	N	N	N	N	N	N	N
NE	N	N	Y	Y	N	Y	N
NH	Y	N	N	N	N	N	Y
NJ	N	Y	N	Y	N	N	N
NM	N	N	Y	N	N	N	N
NV	Y	Y	N	N	N	Y	Y
NY	N	N	N	N	N	N	N
OH	Y	N	Y	Y	Y	Y	Y
OK	N	N	N	N	N	N	N
OR	N	Y	N	N	N	N	N
PA	Y	N	N	Y	Y	Y	Y
RI	N	N	Y	Y	N	N	N
SC	N	N	N	N	N	Y	N
SD	N	N	N	Y	N	N	N
TN	N	N	N	Y	N	N	N
TX	Y	N	N	N	N	Y	Y
UT	N	Y	N	N	N	Y	N
VA	N	N	N	Y	N	N	N
VT	N	Y	N	Y	N	N	N
WA	N	Y	N	N	N	N	N
WI	Y	N	Y	N	Y	Y	Y
WV	N	N	N	N	N	N	N
WY	N	N	N	Y	N	N	N

Note: This table lists state election characteristics in the 2020 presidential election. See note to Tables 2 and A3 for definitions of characteristics.

Table A2: Active and Inactive Registrations by State: L2 data vs. EAVS
Newly added

	L2 data			EAVS		
State	Active reg.	Inactive reg.	Active (%)	Active reg.	Inactive reg.	Active (%)
AK	532,119	12,746	97.7	595,647	50,446	92.1
AL	2,994,930	365,318	89.1	3,438,213	279,585	92.4
AR	1,175,076	385,799	75.3	1,408,061	423,353	76.8
AZ	3,781,982	307,889	92.5	4,275,729	452,380	90.4
CA	20,877,284	247,325	98.8	-	-	-
CO	3,578,369	88,946	97.6	3,803,762	407,766	90.3
CT	2,124,782	179,036	92.2	2,335,860	188,857	92.5
DC	429,570	24,912	94.5	517,890	107,793	82.7
DE	648,370	40,892	94.1	711,287	28,385	96.1
FL	12,955,508	920,121	93.4	14,517,002	701,422	95.3
GA	6,462,992	369,217	94.6	7,194,889	423,547	94.4
HI	714,106	57,831	92.5	-	-	-
IA	1,966,810	134,028	93.6	2,094,770	148,988	93.3
ID	839,004	59,414	93.4	1,029,763	0	100
IL	7,465,362	630,953	92.2	9,103,542	686,351	92.9
IN	3,652,052	584,177	86.2	4,170,353	521,738	88.8
KS	1,602,484	114,477	93.3	1,764,949	148,624	91.6
KY	2,912,392	249,155	92.1	3,319,307	246,121	93
LA	2,836,316	116,704	96	2,963,901	129,103	95.8
MA	3,922,652	516,452	88.4	4,400,254	412,655	91.4
MD	3,879,585	150,048	96.3	4,142,347	156,595	96.3
ME	934,708	49,270	95	1,135,008	3,568	99.6
MI	6,761,976	532,141	92.7	7,209,300	896,224	88.9
MN	3,436,140	50,922	98.5	3,731,016	0	100
MO	3,590,033	395,751	90.1	3,963,980	374,153	91.3
MS	1,736,102	198,444	89.7	1,982,632	160,517	92.5
MT	571,625	68,707	89.3	675,971	71,468	90.4
NC	5,690,869	771,743	88.1	6,607,121	765,487	89.6
ND	349,540	28,824	92.4	-	-	-
NE	1,050,614	97,025	91.5	1,168,708	98,022	92.2
NH	847,682	52,817	94.1	1,087,145	0	100
NJ	5,454,103	397,474	93.2	5,896,836	413,728	93.4
NM	1,079,136	24,235	97.8	1,255,669	105,202	92.2
NV	1,524,037	229,588	86.9	1,835,401	203,761	90
NY	11,375,576	903,046	92.6	12,362,997	1,191,845	91.2
OH	6,797,392	644,935	91.3	8,073,829	0	100
OK	1,696,503	284,962	85.6	2,021,846	237,261	89.4
OR	2,756,422	323,002	89.5	2,944,588	0	100
PA	7,629,443	688,315	91.7	8,280,348	754,713	91.6
RI	694,819	60,189	92	735,195	73,922	90.8
SC	2,892,210	419,955	87.3	3,535,061	319,148	91.7
SD	468,534	68,826	87.2	578,683	56,573	91
TN	3,710,495	284,693	92.9	4,226,928	209,799	95.2
TX	14,121,565	1,812,361	88.6	15,279,870	1,675,649	90.1
UT	1,238,198	144,374	89.6	1,713,297	148,680	92
VA	5,205,669	288,154	94.8	5,763,187	212,374	96.4
VT	405,685	48,682	89.3	440,920	48,357	90.1
WA	4,462,776	321,208	93.3	4,892,871	362,595	93.1
WI	3,169,168	1,408,051	69.2	3,834,164	0	100
WV	969,037	176,452	84.6	1,062,685	206,339	83.7
WY	226,281	7,629	96.7	303,049	0	100

Note: This table reports the number of active and inactive registrations and the percentage of active registrations in L2 database (columns 1-3) and in EAVS (columns 4-6). ”-” indicates a lack of reported data in EAVS.

Table A3: Voting in First State of Registration – Alternative Definitions of Swing States and Sample

	(1)	(2)	(3)	(4)	(5)
	Dependent var.: Vote in first state of registration				
	Swing NYT	Swing ex-ante	Swing ex-post	Drop Goel et al. (2020) states	Add auto absentee application
Only first swing	5.019***	4.885***	4.539***	5.145***	4.737***
	(0.930)	(0.799)	(0.758)	(0.805)	(0.787)
Both swing	3.518***	1.534**	1.291*	1.309*	1.426*
	(0.679)	(0.770)	(0.723)	(0.741)	(0.754)
Only second swing	-1.844**	-3.493***	-2.898***	-3.158***	-2.842***
	(0.912)	(1.032)	(0.966)	(0.988)	(1.006)
Only first auto-mailed ballot	9.600***	9.353***	9.263***	9.468***	9.660***
	(1.677)	(1.566)	(1.596)	(1.636)	(1.607)
Both auto-mailed ballot	6.909***	5.992***	6.137***	6.282***	6.520***
	(1.458)	(1.495)	(1.463)	(1.517)	(1.476)
Only second auto-mailed ballot	-4.867***	-5.734***	-5.543***	-5.811***	-5.456***
	(0.705)	(0.661)	(0.654)	(0.681)	(0.747)
Only first early voting	3.281***	4.363***	4.075***	4.158***	3.976***
	(1.012)	(0.966)	(0.970)	(0.988)	(0.972)
Both early voting	2.586***	2.852***	2.846***	3.084***	2.647***
	(0.623)	(0.570)	(0.560)	(0.558)	(0.550)
Only second early voting	0.288	-0.348	-0.094	0.213	-0.091
	(0.668)	(0.689)	(0.668)	(0.691)	(0.658)
Only first auto absentee application					1.195
					(0.854)
Both auto absentee application					1.264
					(0.828)
Only second auto absentee application					0.149
					(0.902)
Outcome mean (%)	14.85	14.85	14.85	67.41	67.5
Observations	590,991	590,991	590,991	561,001	590,991
R^2	0.020	0.022	0.021	0.023	0.021

Note: This table reports robustness tests for Table 2. All independent variables are defined in the same way as in Table 2 except for those related to *swing*. Column 1 defines swing states as those listed by the New York Times (2020); column 2 uses statewide opinion polling aggregated by FiveThirtyEight (2020), classifying swing states as those with an expected vote margin within 10 percentage points (pp); column 3 classifies swing states as those whose actual vote margin in the 2020 election was within 10 pp (MIT Election Data and Science Lab, 2020). Column 5 includes additional indicators for the first state, both the first and the second states, and only the second state automatically mailing applications to request an absentee ballot to registered individuals (National Conference of State Legislatures, 2020b), respectively. Column 4 excludes states identified by Goel et al. (2020) as having potentially lower data quality in the 2012 presidential election due to multi-generational households. These are Arkansas, Hawaii, Mississippi, New Hampshire, Wisconsin, Wyoming and D.C. Standard errors are clustered at the state-pair level.
*** 1%, ** 5%, * 10% significance level

Table A4: Double Voting – Alternative Definitions of Swing States and Sample

	(1)	(2)	(3)	(4)	(5)
	Dependent var.: Vote in both states of registration				
	Swing NYT	Swing ex-ante	Swing ex-post	Drop Goel et al. (2020) states	Add auto absentee application
Only first swing	0.222**	0.237*	0.241**	0.255**	0.299***
	(0.099)	(0.125)	(0.108)	(0.104)	(0.106)
Both swing	0.490**	0.346**	0.467***	0.384***	0.517***
	(0.238)	(0.153)	(0.170)	(0.133)	(0.153)
Only second swing	0.048	-0.007	0.083	0.082	0.163
	(0.140)	(0.134)	(0.120)	(0.110)	(0.113)
Only first auto-mailed ballot	1.121***	1.107***	1.110***	1.121***	1.226***
	(0.191)	(0.191)	(0.191)	(0.193)	(0.182)
Both auto-mailed ballot	1.093***	1.053***	1.091***	1.140***	1.275***
	(0.218)	(0.212)	(0.216)	(0.215)	(0.225)
Only second auto-mailed ballot	0.197*	0.160	0.200*	0.168*	0.321***
	(0.114)	(0.106)	(0.105)	(0.095)	(0.098)
Only first early voting	0.066	0.131	0.111	0.194**	0.092
	(0.128)	(0.125)	(0.129)	(0.098)	(0.114)
Both early voting	0.037	0.115	0.105	0.287**	0.023
	(0.154)	(0.151)	(0.153)	(0.132)	(0.144)
Only second early voting	-0.080	-0.046	-0.047	0.081	-0.054
	(0.134)	(0.117)	(0.122)	(0.091)	(0.106)
Only first auto absentee application					0.328***
					(0.099)
Both auto absentee application					0.593***
					(0.222)
Only second auto absentee application					0.405***
					(0.113)
Outcome mean (%)	1.04	1.04	1.04	67.41	67.5
Observations	590,991	590,991	590,991	561,001	590,991
R^2	0.002	0.002	0.002	0.002	0.002

Note: This table reports robustness tests for Table 4. The outcome is an indicator for a double-registrant voting in both states of registration in the 2020 presidential election. Definitions of independent variables and sample selections follow those in the corresponding columns in Table A3. Standard errors are clustered at the state-pair level.
*** 1%, ** 5%, * 10% significance level

Table A5: CSSV by Party Registration

	Dep. var.: Vote in 1st state of registration		
	(1) Democrat	(2) Republican	(3) Independent
Only first swing	3.928***	5.617***	4.790***
	(0.803)	(1.226)	(0.796)
Both swing	1.937**	0.409	1.066
	(0.970)	(1.059)	(0.686)
Only second swing	-2.090*	-4.232***	-2.929***
	(1.243)	(1.016)	(0.901)
Only first auto-mailed ballot	9.482***	7.032***	10.653***
	(1.964)	(1.530)	(1.647)
Both auto-mailed ballot	4.629**	9.068***	7.249***
	(1.979)	(1.421)	(1.482)
Only second auto-mailed ballot	-5.328***	-4.388***	-5.583***
	(0.759)	(0.890)	(0.687)
Only first early voting	3.791***	5.344***	3.630***
	(1.104)	(1.160)	(0.957)
Both early voting	1.553**	3.013***	4.059***
	(0.705)	(0.791)	(0.608)
Only second early voting	-0.365	-1.362	0.710
	(0.751)	(0.868)	(0.763)
Outcome mean (%)	13.88	14.7	16.35
Observations	273,914	126,599	190,478
R^2	0.022	0.022	0.021

Note: This table reports party heterogeneity in the relationship between a double-registrant's propensity to vote strategically and the characteristics of the two states they are registered in. The outcome is an indicator for an individual voting in their first state in the 2020 presidential election. The sample in columns 1, 2, 3 consists of Democrats, Republicans, and Independents, respectively; party affiliation is based on an individual's second state of registration. All specifications and variable definitions mirror those in Table 2 column 2. Standard errors are clustered at the state-pair level
*** 1%, ** 5%, * 10% significance level

Appendix 1. Scaling factor

This appendix describes the scaling factor used in the estimate for the prevalence of double-registrants. The initial analysis in Section 3 assumes all double-registrants have *identical* recorded first and last names in both state registrations. However, this is not always the case for a variety of reasons, such as name changes following marriage, name recording errors, use of nicknames (e.g., Thomas vs. Tom) and so on. We correct for this by estimating a scaling factor capturing the ratio of double-registrants with identical recorded first and last names to the full set of double-registered individuals.

In order to do this we build a reference sample of double-registrants *without* matching on both first and last name, but which correspond to the same person with a high degree of confidence. We begin with all cross-state matches with the same cellphone number, birth date (day, month, and year) and gender. We then apply an additional matching criterion to ensure a high match quality for the reference dataset. Specifically, we require at least one of the following to match: (i) first name, (ii) last name, (iii) middle initial, or that (iv) the full name matches are above a textual similarity threshold.[1] In the reference dataset, 83.65% of pairs have identical first and last names, which means that for every 100 pairs with identically matching first and last names, we estimate that there are an additional 19.5 double-registrants without identical names across both of their records, despite being the same person (i.e., we use a scaling factor of $1/0.83654 = 1.1954$).

[1]Textual similarity is determined by removing spaces between the text of the full name, then calculating the fraction of letter pairs that match, irrespective of position. For example, BOB SMITH and ROB SMITH have a textual similarity score of 6/7 because 6 out of the 7 pairs of adjacent letters match. We require this to be above 0.2.