

NBER WORKING PAPER SERIES

CROSS-STATE STRATEGIC VOTING

Gordon B. Dahl
Joseph Engelberg
Runjing Lu
William Mullins

Working Paper 30972
<http://www.nber.org/papers/w30972>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue
Cambridge, MA 02138
February 2023, revised September 2024

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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2023 by Gordon B. Dahl, Joseph Engelberg, Runjing Lu, and William Mullins. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Cross-State Strategic Voting

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

NBER Working Paper No. 30972

February 2023, revised September 2024

JEL No. D72

ABSTRACT

We estimate that 3.1% of US voters, or 6.1 million individuals, were registered to vote in two states in 2020, opening up the possibility for them to choose where to vote. Double registrants are concentrated in the wealthiest zipcodes and respond to both incentives and costs, disproportionately choosing to vote in swing states (higher incentive) and states which automatically send out mail-in ballots (lower cost). We call this behavior Cross-State Strategic Voting. While others have documented strategic incentives on *who* to vote for, this paper is the first to consider strategic incentives on *where* to vote.

Gordon B. Dahl
Department of Economics
University of California, San Diego
9500 Gilman Drive #0508
La Jolla, CA 92093-0508
and NBER
gdahl@ucsd.edu

Joseph Engelberg
Rady School of Management
University of California at San Diego
Otterson Hall, Room 4S154
9500 Gilman Drive #0553
LaJolla, CA 92093-0553
jengelberg@ucsd.edu

Runjing Lu
University Campus
University of Toronto
Toronto, ON M5S3A3
Canada
runjing.lu@utoronto.ca

William Mullins
Rady School of Management
University of California, San Diego
9500 Gilman Drive, MC 0553
La Jolla, CA 92093-0553
wmullins@ucsd.edu

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.”²

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 define a double-registrant as the same person listed on two different state voter rolls at the same point in time (October 2020).³ 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

¹In *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.”

²A Federal judge issued a preliminary injunction blocking this provision in 2022. The case is currently pending before the 9th U.S. Circuit Court of Appeals.

³A separate, and interesting, question is how many double-registered individuals *know* they are double registered, but this is difficult to measure. Our definition is objective, measurable, and policy-relevant (since unaware double-registrants could be activated).

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 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 2.2% 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 4.7%, 5.8%, and 8.0%, respectively. In other words, compared to the bottom quartile, the rate of double-registration is over twice as high among the top 10% and 3.6 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 583,835 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).⁴ We emphasize that there is nothing illegal about being registered in two states, as individuals are not automatically de-registered when they move.

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.⁵ 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

⁴Strategic 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.

⁵Voting in a state which is not one’s legal domicile is illegal and has resulted in a handful of prosecutions (New York Times, 2022).

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 4.7 percentage points (pp) higher than if neither state is swing. This amounts to a 48% increase relative to a mean of 9.8%. In other 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 1.9 pp less likely to vote in the first state.

This asymmetry in estimated coefficients based on whether only the first state or only the second state is a swing state suggests that state-pair omitted variable bias is unlikely to be driving our results. Specifically, our analysis includes both double registrants who move from one state to another and those who move in the opposite direction. If there were an unobserved characteristic associated with state-pairs that included a swing state, both coefficients would have the same sign.

Our 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 20% of the mean. However, if only the second state does, the likelihood of voting in the first state decreases by 31%.

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.

We also find evidence of incentive- and cost-based CSSV in the 2016 election as well as in a different sample in 2020 constructed using uncommon names and birthdays (rather than cellphones) to identify double registrants. The pervasiveness of CSSV across different elections and algorithms to identify double registrants provides strong evidence that our results are not driven by the idiosyncrasies of the 2020 election or our reliance on cellphone information.

We would expect greater responsiveness to incentives among individuals who spend time in both of their registration states, as they are more likely be aware of their double-registration status and have the ability to act upon it. As a proxy, we use Corelogic’s Tax Assessment database to identify double registrants who own property in both states. We estimate stronger incentive-based CSSV effects among this subsample who maintain a footprint in both states.

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 0.031% of voter registrations), roughly in line with the estimate in Goel et al. (2020) for the 2012 election.

Finally, a back-of-the-envelope calculation indicates that CSSV did not affect the 2020 election; this is because the behaviors of Republican and Democrat voters largely canceled each other out. Nevertheless, CSSV could be important in closer elections, such as the 2000 election, which was won by 537 votes in Florida. In the same election, only 366 votes separated the candidates in New Mexico. The existence of a large population of double registrants opens up the possibility that politicians or campaigns could activate them to vote strategically in swing states.

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, Fukumoto and Horiuchi, 2011, Ichino and Schündeln, 2012, Hidalgo and Nichter, 2016, 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 exploring the 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.

The remainder of the paper proceeds as follows. The next section describes our voter registration data. Section 3 estimates the prevalence of double-registrants and Section 4 presents evidence of CSSV. Section 5 concludes.

2. DATA

Our data on voter registrations across the United States is from L2, a well-known, non-partisan data vendor used by political campaigns and the academic literature (e.g., Allcott et al., 2020, Brown et al., 2021, Billings et al., 2021, Spenkuch et al., 2023, Chyn and Haggag, 2023, Billings et al., 2024, Elder et al., 2024). 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, and cleans the data.

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.⁶

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; see 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 due to minor differences in when vote totals are certified versus when L2 receives the data.⁷

Another way to validate the quality of L2’s registration information is to compare it to official data collected by the Election Administration and Voting Survey (EAVS). Table A1 shows that the number of voters in each state according to the L2 data tends to be around 90% of the corresponding number in EAVS, consistent with L2 cleaning its data.⁸

We determine voter registration dates using the information on state rolls. Since North Dakota does not have voter registration and New Hampshire does not report voters’ registration date in 2020, our data covers 48 states and DC.

⁶Party identification is inferred in Alabama, Georgia, Hawaii, Illinois, Indiana, Michigan, Minnesota, Missouri, Montana, North Dakota, Ohio, South Carolina, Texas, Vermont, Virginia, and Washington.

⁷The 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.

⁸The appendix table also compares active and inactive registrations. The number of active registrations show a similar pattern to the total. The number of inactive registrations are sometimes larger in L2 than in EAVS, likely because some inactive registrants as of October 2020 (when L2 data is recorded) could have voted in the 2020 election, thus becoming active in the EAVS data (collected between November 2020 and July 2021).

2.1 SUMMARY STATISTICS

The main dataset uses voter roll data as of the month immediately preceding the November 2020 general election. Because the National Voter Registration Act of 1993 does not allow states to purge voters within 90 days of federal elections, we can be sure that the double-registered individuals in our sample have the ability to vote in both states they are registered in.⁹

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 are shown in column 2. The availability of a cellphone number is largely responsible for the reduction in the sample size as we move from columns 1 to 2. The set of voters in the two columns are remarkably similar in terms of voting history, partisan fractions, and demographics.¹⁰ 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 (28% 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

⁹There are six states exempt from the Act: Idaho, Minnesota, New Hampshire, North Dakota, Wisconsin, and Wyoming (U.S. Department of Justice).

¹⁰Even though the means are qualitatively close to each other, since the samples sizes are so large, the means are statistically different.

3 (47% vs. 22%), consistent with the fact that Republicans are less likely to move.¹¹

Appendix Figure A1 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 years.

3. PREVALENCE OF DOUBLE-REGISTRANTS

We define a double-registrant as the same person listed on two different state voter rolls at the same point in time (October 2020). In Section 2, we saw 203 million *registrations* across all 50 states and DC, but these may not correspond to 203 million unique *individuals*. In this section, we estimate how many of these 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.¹² We also estimate how this prevalence varies by wealth, income, and party.

Before estimating the number of double registrants, we first describe the various ways one can become a double registrant. Because there is no national voter registry, state laws govern registration, and there is limited cross-state coordination regarding voter registries. A necessary condition for double-registration is moving across state lines. When individuals register to vote in their destination state, they become double-registrants if they have not been de-registered in their origin state. There is no obligation for individuals to de-register, and few do. All but 6 states are subject to the National Voter Registration Act (NVRA) of 1993, which specifies that states must wait for at least two federal election cycles (4 years) before removing a voter from the rolls. In addition, states must comply with certain notification procedures.¹³

¹¹Using Cooperative Election Study survey data for those registered to vote in 2020, we find that Republicans are 2.6 pp less likely to have moved relative to an overall mean of 8.8%.

¹²L2 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.

¹³In terms of a mover's destination state, they may choose to register deliberately or, in some states, they will be automatically registered to vote when they apply for a driver's license, unless they opt out. We find

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.¹⁴

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.¹⁵ If the empirically observed percentage of cases with the same birth month and day is higher than 0.2833%, then we can 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:

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

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

a larger fraction of registrations on non-election days for states with automatic voter registration than states without (97.5% vs. 93.9% using the two years prior to the 2018 election).

¹⁴27 states are missing *day* of birth information at a non-trivial rate, and the data 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.

¹⁵We use 353 instead of 365 as we exclude observations with first of the month birthdays because the data often record day of birth as 1 when day information is missing. Using the empirical distribution of birthdays, instead of assuming uniformity, would not materially alter the results of this calculation. If we instead assume a uniform distribution of birthdates (month, day, and year) over our entire sample, this likewise changes the estimated rate of double registration by less than a tenth of a percentage point.

level of NU we can estimate $P(D_1 = D_2)$. Using this information and rearranging terms, for each NU we obtain:

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

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 $\frac{87.511\% - 0.2833\%}{1 - 0.2833\%} = 87.47\%$ of these pairs are the same person, which translates to 1,294,288 voters.¹⁶ 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 A2. 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).¹⁷

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.¹⁸

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. To calculate this, we build a reference sample of

¹⁶There are 930,548 across-state matches with $NU=2$ and non-missing birthdate information, which also share the same birth year; 87.47575% of 930,548 is 814,004. Since there are 37.10799% across-state matches that have missing birthdate information, we assume missing-at-random and scale up 814,004 to 1,294,288.

¹⁷We 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.

¹⁸This estimate is larger than those for 2012 in Goel et al. (2020) and for 2011 in Pew (2012), potentially because COVID disrupted residential stability in unique ways.

double-registrants *without* matching on both first and last name, but which correspond to the same person with a high degree of confidence.

Specifically, 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. 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.¹⁹ 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. This yields a scaling factor of 1.1954.

Applying this scaling factor, we estimate there were 6.1 million double-registrants, or 3.1% of voters as of October 2020.²⁰

3.2 HETEROGENEITY

Wealth. Figure 1 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 2.2%. The middle 50 percent (25-75th percentile) rises modestly to 3.1%. In contrast, individuals living in wealthier zip codes are more likely to double register: the rates are 4.7%, 5.8%, and 8.0% 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.

Income. In the lower panel we use zip code-level adjusted gross income from IRS tax

¹⁹Textual 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.

²⁰Recall from Section 2 that L2 cleans its data in order to provide high quality data for political campaigns. To the extent that L2’s cleaning removes double registrants from their voter files, this would result in us underestimating the size of the double registrant population.

data.²¹ The bottom quartile has a double-registration rate of 2.8%. The middle half is similar. However, the highest-income zip codes have much higher rates: 4.9% (≥ 90 th), 5.6% (≥ 95 th), and 6.8% (≥ 99 th). 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.

Party. 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 3.4% (Independents have the highest double-registration rate of 4.2%). 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.

Automatic voter registration. Finally, we find similar rates of double registration for states with and without automatic registration: 3.2% vs. 3.0%. This suggests being a double registrant is mainly driven by the *first* state keeping individuals on the voter rolls after a move and less about them automatically becoming registered in their *second* state.

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,

²¹The 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.

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.²² At the pair level, we have the following conditional probabilities:

$$P(\text{Either check digit matches} \mid \text{same person}) \approx 1$$

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

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). 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 voters registered in more than two states or those with missing state registration date.²³ This leaves us with a sample of 583,835 double-registrants.

4.2 CROSS-STATE VOTING INCENTIVES AND COSTS

We now explore which state a double-registrant chooses to vote in depends on that state’s election characteristics. Specifically, elections in “swing states” are more consequential for

²²We 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.

²³We 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 finally drop a handful of voters recorded as registering in two states on the same day.

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).

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 583,835 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 10% of double-registrants in our sample voted in their first state in the 2020 presidential election.

In Figure 2, we show the share of double registrants voting in their first state of registration by first state characteristics (swing, auto-mail, and neither) $\times$ second state characteristics (swing, auto-mail, and neither). The first thing to note is that the white bars are always taller than the corresponding black and grey bars within each group of three bars. This indicates that second states which are neither swing nor auto-mail lose the largest share of votes to first states. For example, when the first state is swing (the middle group), neither states lose around 14% of votes to swing states, whereas swing states and auto-mail states only lose 12% and 11%, respectively.

The second thing to note is that the leftmost group of three bars are always shorter than the other groups. This indicates that first states that are neither swing nor auto-mail attract the fewest votes from second states. For example, first states that are neither swing nor auto-mail (leftmost white bar) attract less than 10% from neither second states, whereas swing and auto-mail first states (the other white bars) attract around 14% and 12%, respectively.

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

$$\begin{aligned} \text{Vote first}_i = & \gamma_0 + \gamma_1 \text{Only first swing}_i + \gamma_2 \text{Both swing}_i + \gamma_3 \text{Only second swing}_i \\ & + \gamma_4 \text{Only first auto-mailed ballot}_i + \gamma_5 \text{Both auto-mailed ballot}_i \\ & + \gamma_6 \text{Only second auto-mailed ballot}_i + \epsilon_i \end{aligned} \quad (3)$$

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 auto-mailed ballot. 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); see Table A2 for state-by-state

classifications.²⁴ This regression exploits the variation across 2,176 ordered state pairs (out of a theoretical maximum of $49!/(49-2)! = 2,352$).

4.3 MAIN RESULTS

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 4.7 pp more likely to vote in their first state, which translates to a 38% 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.²⁵

The fact that we find coefficients of opposite signs based on whether the first state is swing or the second state is swing suggests that state-pair omitted variable bias is unlikely to be driving our results. Specifically, our analysis includes both double registrants who move from one state to another and those who move in the opposite direction. If there were an unobserved characteristic associated with state pairs that included a swing state, both coefficients would have the same sign.²⁶

In interpreting the incentive-based voting result in Table 2, we recognize that outside forces could be encouraging individuals to vote. For example, a campaign could target voters in swing states rather than in safe states, so that individuals with both a swing and a safe state registration will be more likely to receive communications associated with their swing state. It is also possible that state governments of swing states could choose to nudge double registrants into voting in their state via streamlined voting procedures. Regardless, this result demonstrates that votes among double-registered individuals flow to the states that are more consequential.

²⁴We consider alternative definitions of swing state in Appendix Table A3: those listed by the New York Times (2020), aggregated opinion polling by FiveThirtyEight (2020), and actual vote margin (MIT Election Data and Science Lab, 2020). Our main results are robust to swing-state definition.

²⁵The coefficient on *both swing* is positive and marginally insignificant. At first glance, this coefficient may suggest that an omitted variable explains both voting in the first swing state and how competitive it is, regardless of how competitive it is relative to the second swing state. But as we will show in Figure 3, relative competitiveness is strongly predictive of where an individual votes.

²⁶For an omitted variable to be a concern, it would have to be correlated with the direction of moves *within* state pairs, for reasons unrelated to states' swing status.

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).²⁷ We find that if only the first state automatically mails out ballots, the likelihood of voting there increases by 20% relative to the mean (column 1). However, when only the second state has auto-mailed ballots, double-registrants become less likely to vote in their first state (-31%). 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.

In Appendix Table A4, we explore heterogeneity by characteristics that might be associated with more strategic behavior: income and retirees. We find some evidence that retirees engage in more incentive-based CSSV, perhaps because they have the time and are more likely to own a second home to be strategic.

In Appendix Table A6 we also report heterogeneity by party affiliation. Using these estimates, we conduct a back-of-the-envelope calculation to explore how CSSV affected the 2020 election (see Appendix). This calculation requires several assumptions, including extrapolating estimates from our regression sample to the broader voter population and inferring voting behavior from party affiliation. With these caveats in mind, we find that CSSV did not change the outcome of the 2020 election. This is because the CSSV behavior of Republicans and Democrats largely canceled each other out on both the incentive- and cost-based margins in this election.

4.4 NEAR-CERTAIN VOTERS

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 2 and 3 we consider registrants who are predicted to vote in the 2020 presidential election with a very high probability ex-ante. In column 2, we select all individuals who voted in the 2020

²⁷While 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.

primary election. This group has a 97.4% chance of voting in the general election. In column 3, 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 A3). We use individuals with a predicted probability of voting at or above 99%.

This “identification at infinity” approach (d’Haultfoeuille and Maurel, 2013) is the cleanest evidence for CSSV because it nets out the decision of whether to vote which leaves only the choice of where to vote. Reassuringly, the estimated coefficients are larger among these near-certain voters. While the sample sizes are much smaller in columns 2 and 3, and hence the standard errors increase, the coefficients remain statistically significant. For incentive-based CSSV, the coefficient on *Only first swing* is 7.715 in column 3 (relative to 4.719 in column 1) and *Only second swing* is -4.585 (relative to -1.889). For cost-based CSSV, three out of four relevant coefficients in columns 2 and 3 are at least 50% larger compared to those in column 1.

The larger coefficients could be due to two reasons. First, 19% of double-registrants do not vote at all, which will attenuate the coefficients in column 1. Second, near-certain voters might be more politically engaged and therefore more responsive to strategic incentives.²⁸

4.5 ROBUSTNESS

4.5.1 EVIDENCE FROM THE 2016 ELECTION

Our main analysis focuses on 2020 because the data is superior in a couple of dimensions. First, in 2020 our data vendor L2 pulled data from all states in the same month: October of 2020 (just one month before the election). The National Voter Registration Act of 1993 prohibits states from purging voters within 90 days of a general election. This means that we can be certain that individuals in our sample of double registrants have the possibility to vote in both states they are recorded as being registered in. Second, cellphone data is consistently available in the L2 voter database. By 2020 most people have cellphones and keep the same number when they move. This is important because it helps us uniquely

²⁸2.7% of double-registrants are FEC donors; this rises to 4.9% among primary voters and to 7.6% among those who have a predicted probability of voting at or above 99%.

identify double registrants across states.

The data is less ideal for the 2016 election because state files are collected at different points in time, and cellphone data is not consistently available.²⁹ The state files not being collected at the same time could introduce measurement issues, and the absence of cellphone data requires an alternative way of identifying double registrants.

With these caveats in mind, we analyze CSSV for the 2016 election as a robustness check. We exploit uniquely named individuals – people whose names are so unique that their names appear only twice, in two separate states, and also share the same birthdate. For example, if a very unique name only appears on the California and Arizona voter rolls and happens to share the same exact birthdate, we can be nearly certain this is the same person registered in two different states. Using this alternative approach, we can identify 665,641 double registrants.

Results for 2016 using the unique-name sample are reported in Table 3. We find strong evidence for both incentive-based and cost-based CSSV, especially in our cleanest sample of near-certain voters. For example, for voters with a high predicted probability of voting (column 3), the *Only first swing* and *Only second swing* coefficients are 10.763 and -3.976, respectively.³⁰ For our measure of cost-based CSSV, we use early voting in 2016 instead of auto-mailed ballots as we did in 2020 because auto-mailed ballots only became a widespread phenomenon following the COVID pandemic.³¹ This cost of voting measure yields strong evidence of CSSV as indicated by the coefficients of 6.610 on *Only first early voting* and -6.123 on *Only second early voting* in column 3.

²⁹In 2016, the timing of L2’s collection varied across states. For some states, L2 collected data in, for example, March and August and for others May and September. Data from larger and swing states were collected multiple times per year while smaller states were often collected just once a year (and not necessarily for the same month).

³⁰For the 2016 election, we set the voting probability cutoff for near-certain voters to be greater or equal to 95% as otherwise the sample becomes too small. If we apply the same cutoff of 95% instead of 99% to the 2020 election, results are very similar: *Only first swing* is 6.397 (s.e.=1.184), *Only second swing* is -3.198 (0.0918), *Only first auto-mailed ballot* is 1.895 (0.799), and *Only second auto-mailed ballot* is -5.609 (0.875).

³¹Auto-mailed ballots were used in only four states in 2016 (6.3% of the voter population) vs. 11 states in 2020 (21.6% of the voter population).

4.5.2 ALTERNATIVE SAMPLES FOR THE 2020 ELECTION

Unique-name sample. As an additional robustness check for our 2020 election results, we construct a unique-name sample of double registrants following the same approach in the previous section. This sample of 804,610 double registrants has only around a 20% overlap with our main sample of double registrants constructed using cell phone numbers. We reproduce Table 2 using this new sample of double registrants. The results are reported in Table 4.

Although the samples of double registrants are largely independent, the results in Table 4 are consistent with those in Table 2. For example, among near-certain voters (column 3), the *Only first swing* coefficient is 6.598 (vs. 7.715 in Table 2), and the *Only second swing* coefficient is -2.553 (vs. -4.585). Similarly, the cost-based coefficients share the same pattern across the tables.

To the extent that double-registered cellphone users are different on unobservables (we showed in Section 2 that cellphone users are similar on observables to the rest of the population), these results provide evidence using a largely independent sample of voters where we do not select on cellphone.

Dual-state homeowner sample. While ownership of property in two states is not a necessary condition to engage in CSSV, it may facilitate strategic behavior since these individuals have an active connection to each state and may be more aware of their double-registrant status. To build a sample of these individuals, we take the union of our cellphone and the unique-name samples and match it to CoreLogic’s Tax Assessment Database, a near-comprehensive dataset of properties. We match on owners’ first name, last name, and longitude and latitude of an address since these variables appear in both the L2 voter and CoreLogic property databases.³²

Table 5 presents the results. Using this sample of dual-state homeowners, we find substantially larger incentive-based CSSV, consistent with the idea that individuals can more easily

³²After matching, the cellphone sample consists of 15,038 double-registrants and the unique-name sample consists of 12,771; the union of the two is 25,308. Because these samples are relatively small, we use the union to increase precision.

vote where they own a property.³³ For example, for the near-certain general election voters in column 3, the *Only first swing* coefficient is 13.084 and the *Only second swing coefficient* is -11.654. These point estimates are substantially larger compared to those in Table 2. This analysis complements our main finding by demonstrating stronger incentive-based CSSV in a subpopulation where it is more likely to occur.

4.5.3 ADDITIONAL CONTROLS

As additional robustness, in Table A5, we report the results of alternative specifications for Table 2 column 1 that control for state-pair level election features, voter roll maintenance, migration flows, or tax rates.

Column 1 includes indicators for only the first state allowing early voting, for only the second state allowing it, and for both allowing it. In a similar fashion, columns 2, 3, and 4 include indicators for automatic mailing of absentee applications, automatic registrations, or election day registration, respectively.

Columns 5 and 6 include control variables for how well states maintain their voter rolls. In column 5, we use the 2020 Election Administration and Voting Survey (EAVS) to categorize states with an above-median removal rate and construct indicators mirroring the specification in column 1. Column 6 similarly controls for how frequently states update voter registration date, as measured by how often the state registration date matches L2's.³⁴

Column 7 controls for 2018-2019 state-pair level migration flows from the first state to the second state using data from the IRS Statistics of Income.³⁵ Column 8 accounts for differences in state tax burdens (Tax Foundation, 2020) in a similar manner to column 1.

Regardless of which controls are added, our CSSV estimates remain robust.

³³For cost-based CSSV, the prediction for how the effects should change is unclear. On one hand, it is easier to vote in person if an individual has a physical presence in each state, making auto-mailed ballots less relevant; on the other hand, if an individual owns a property, they will automatically receive a ballot in the mail if they are registered, potentially making auto-mailed ballots more relevant.

³⁴L2 attempts to retain the original date a voter was registered in a state, while many states update this date in their records as voters interact with the registration system.

³⁵In a separate and untabulated analysis, we regress migrant inflows per capita on whether only the destination state is swing, only the origin state is swing, or both are swing. The estimates on all three indicators are small and insignificant.

4.5.4 STATE SUBSETS

States vary considerably by population size, with many more double registrants being registered in a large state like California than in a small state like New Hampshire. To examine whether any one state is driving our results, Appendix Figure A4 plots the incentive-based and cost-based CSSV coefficients when we exclude one state at a time. Estimates are stable across samples.

Appendix Table A3 column 4 excludes states identified in Goel et al. (2020) as having potentially lower data quality in the 2012 presidential election due to multi-generational households. Column 5 excludes states exempt from the NVRA: Idaho, Minnesota, Wyoming, and Wisconsin (North Dakota and New Hampshire are already excluded because they do not provide information on registration date). In both columns, results are robust.

4.6 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 should be higher when it is paired with Iowa.

Figure 3 and Table 6 explore these effects. Panel (a) of Figure 3 plots the probability of voting in the first state against the first state “win gap,” defined as the absolute value of the 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 49 values for the win gap (48 states plus DC), panel (b) has 2,176 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 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 6 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 (66 pp) increase in the first state’s predicted win gap (*First win gap*) reduces a double-registrant’s likelihood of voting there by 2.8 pp (66×-0.043). A one standard deviation increase in *Second – first win gap* (6 pp) increases the likelihood of voting in the first state by 0.2 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 the auto-mailed ballot variables are largely unaffected.

Stepping back, it is worth thinking carefully about identification and possible omitted variable bias. In Table 2, we rely on state-*pairs* for identification, so a bias cannot arise just because swing states are different from non-swing states. For example, even if Arizona (a swing state) is different from other states, this difference would have to affect the likelihood of voting there *differentially* depending on whether Arizona is the first or second state in a pair. In Figure 3 (and the corresponding Table 6), we rely on relative differences in competitiveness across state pairs for identification. The results in this section require that any omitted variable would not only have to affect directionality, but also intensity.

4.7 CSSV AND DOUBLE VOTING

Finally, 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)).³⁶ We observe a mean rate of double voting of approximately 1% among double-registrants in our sample, amounting to 0.031% of voter registrations. Hence, we reach the same conclusion as Goel

³⁶National 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.

et al. (2020) that double voting is exceedingly rare.

Moreover, even this low rate is likely an over-estimate for two reasons. First, outcome 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.³⁷

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 cast 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 7 reports the results and finds evidence of CSSV. When only the first state is swing, the probability of double voting increases by 26% of the mean in column 1, and when both states are swing this rises to 41%. Double-registrants respond even more strongly to costs. Double voting increases by 111% if both states send out auto-mailed ballots. The effect is similar for *Only first auto-mailed* (122%) and smaller for *Only second auto-mailed* (26%), likely because cost considerations matter less in the state where double-registrants live.

Columns 2 and 3 condition on near-certain voters, as in Table 2. The influence of incentives increases, with the coefficient in column 3 on *Both swing* rising to 57% of the subsample mean. For cost-related variables, we see a decrease in effect sizes relative to the mean.

³⁷Meyer and Mittag (2017) show that a binary dependent variable with conditionally random measurement error attenuates coefficients in most situations.

5. 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.³⁸ While our back-of-the-envelope calculation indicates that CSSV did not affect the outcome of the 2020 election, it could be important in closer elections such as the 2000 election, where the gap was 537 votes in Florida and 366 in New Mexico. Moreover, there is the possibility that campaigns could target double-registrants – or encourage people to become 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 3.6 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.

³⁸Even 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., Braun, N., Jones, D. B., and Shi, Y. (2024). Disparate racial impacts of shelby county v. holder on voter turnout. *Journal of Public Economics*, 230:105047.
- 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.
- Brown, J. R., Enos, R. D., Feigenbaum, J., and Mazumder, S. (2021). Childhood cross-ethnic exposure predicts political behavior seven decades later: Evidence from linked administrative data. *Science Advances*, 7(24).
- 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.
- Chyn, E. and Haggag, K. (2023). Moved to vote: The long-run effects of neighborhoods on political participation. *Review of Economics and Statistics*, 105(6):1596–1605.
- 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.
- Elder, E. M., Enos, R. D., and Mendelberg, T. (2024). The long-term effects of neighborhood disadvantage on voting behavior: The “moving to opportunity” experiment. *American Political Science Review*, 118(2):988–1004.
- 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.
- Fukumoto, K. and Horiuchi, Y. (2011). Making outsiders’ votes count: Detecting electoral fraud through a natural experiment. *American Political Science Review*, 105(3):586–603.

- 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.
- Hidalgo, F. D. and Nichter, S. (2016). Voter buying: Shaping the electorate through clientelism. *American Journal of Political Science*, 60(2):436–455.
- 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.
- Ichino, N. and Schündeln, M. (2012). Deterring or displacing electoral irregularities? spillover effects of observers in a randomized field experiment in ghana. *The Journal of Politics*, 74(1):292–307.
- 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 (2012). Inaccurate, costly, and inefficient - evidence that america’s voter registration system needs an upgrade. <https://www.pewtrusts.org/~media/legacy/uploadedfiles/pcs-assets/2012/pewupgradingvoterregistrationpdf.pdf> (accessed August 12, 2024).
- 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*, 91(4):1171–1203.
- 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.

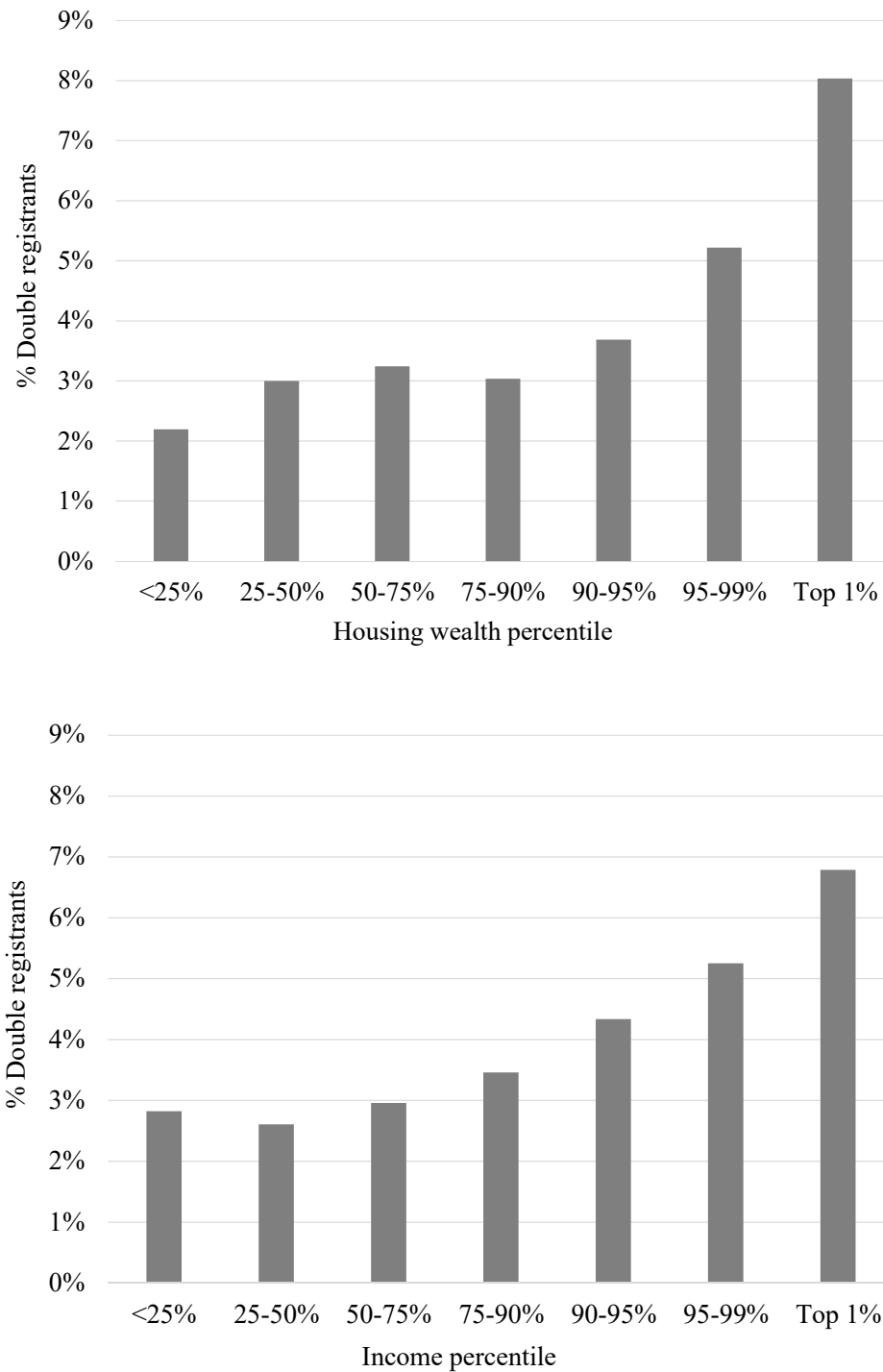


Figure 1: DOUBLE REGISTRATION RATES BY HOUSING WEALTH AND INCOME

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

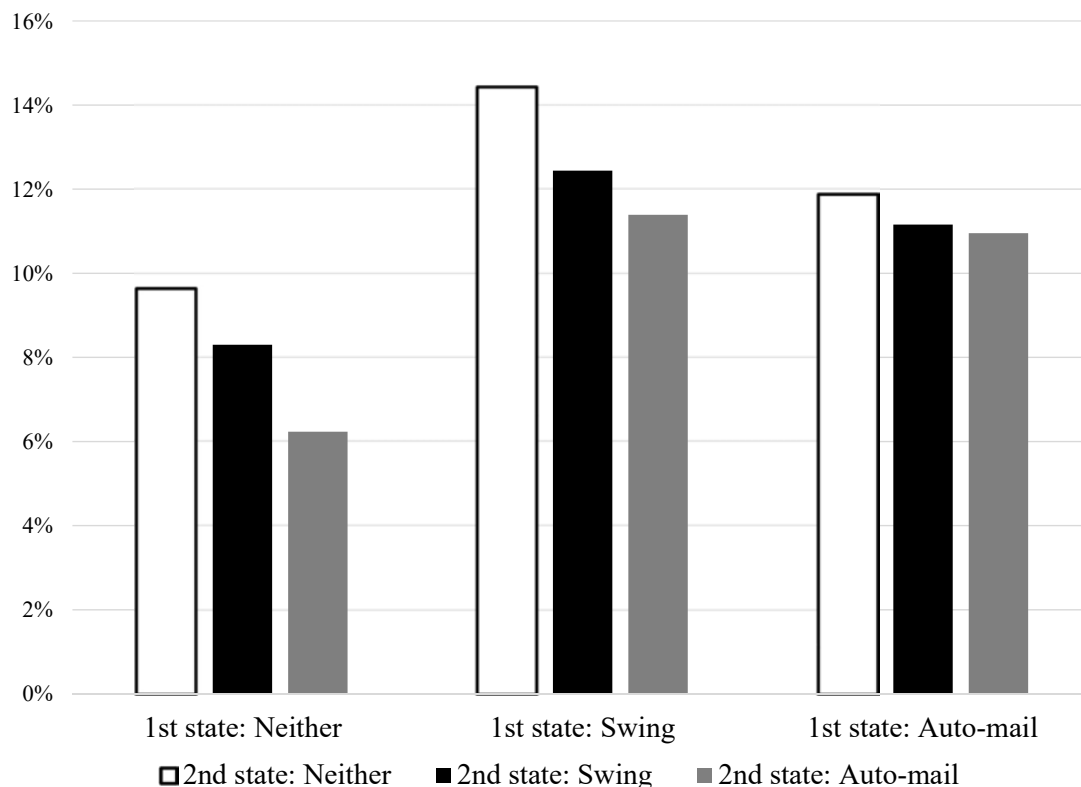


Figure 2: SHARE OF DOUBLE REGISTRANTS VOTING IN THEIR FIRST STATE OF REGISTRATION

Note: This figure shows the share of double registrants voting in their first state of registration by state-pair characteristics. We create a 3×3 matrix of first state characteristics (swing, auto-mail and neither) $\times$ second state characteristics (swing, auto-mail and neither). For example, the leftmost white bar shows the share of double registrants voting in their first state when it is neither auto-mail nor swing and when their second state is also neither swing nor auto-mail. *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); *auto-mail* indicates states that automatically sent mail-in ballots to registered voters in 2020 (National Conference of State Legislatures, 2020b). Nevada is the only swing and auto-mail state; we include it once in each group.

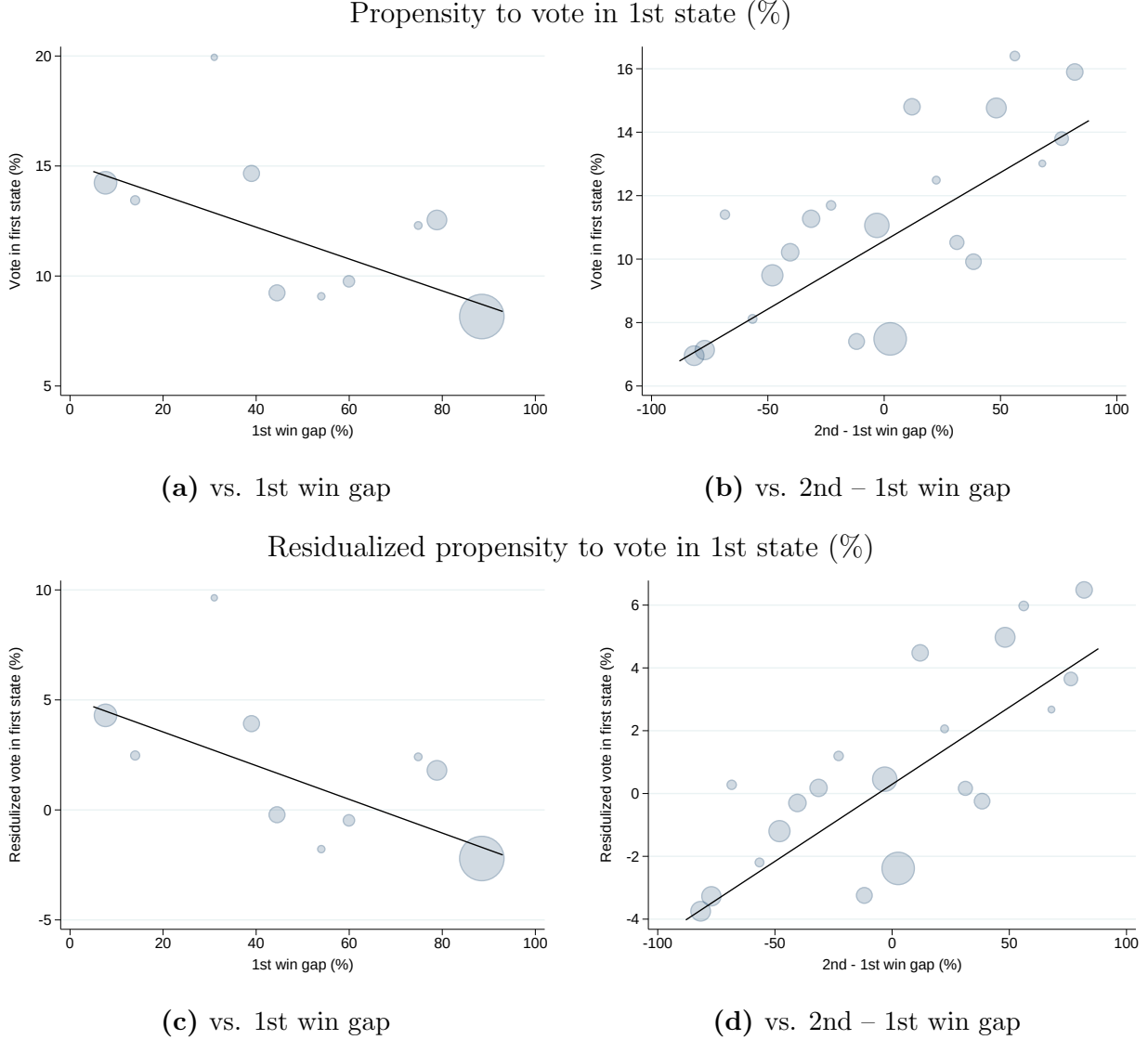


Figure 3: VOTING IN FIRST STATE OF REGISTRATION AND ELECTION CLOSENESS

Note: These plots 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 variables in Table 6. Dot sizes are proportional to the number of observations.

Table 1: SUMMARY STATISTICS FOR REGISTERED VOTERS

	All registrants	Registrants with cell phone & other info	Known double registrants
<i>2020 presidential election (%)</i>			
Vote	74.4	78.1	81.3
Vote in first state	-	-	9.8
<i>Voting history (%)</i>			
Vote in general elections	64.0	65.6	66.5
Vote in primary elections	23.0	23.7	22.3
Vote in minor elections	8.7	8.8	6.5
<i>Party & donations (%)</i>			
Republican	31.1	32.7	21.6
Democrat	41.0	41.6	46.9
Independent	28.0	25.6	31.5
FEC donor	1.5	1.8	2.7
<i>Demographics (%)</i>			
Male	47.0	46.4	52.2
White	70.9	70.4	70.7
Hispanic	13.6	14.4	14.0
Black	11.8	11.7	10.7
Asian	3.7	3.5	4.7
Low wealth	32.3	30.2	20.8
Middle wealth	49.3	49.6	49.8
High wealth	18.4	20.2	28.0
Low income	22.7	22.1	18.4
Middle income	57.6	57.1	51.5
High income	19.7	20.8	30.1
Married	40.8	43.5	42.0
Any children	35.5	38.4	44.6
Homeowner	63.8	66.8	49.5
Born pre-1955	22.9	22.7	12.8
Born 1955-64	17.7	19.7	12.9
Born 1965-74	16.0	18.2	15.6
Born 1975-84	15.6	17.7	22.2
Born 1985-94	16.8	15.5	31.2
Born post-1994	10.9	6.3	5.3
Observations	202,535,296	90,234,280	583,835

Note: Column 1 reports summary statistics for all registered voters as of October 2020; column 2 for those with the information required to match individuals across states, where the main constraint is the availability of cellphone numbers; and column 3 for registered voters 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 IN 2020

	Dependent var.: Vote in first state of registration		
	(1)	(2)	(3)
	All double-registrants	Near certain general election voters	
		Voter in primary	Predicted voter $Pr \geq 0.99$
Only first swing	4.719*** (0.794)	5.873*** (1.087)	7.715*** (1.527)
Both swing	1.421 (0.901)	1.339 (1.231)	1.703 (1.179)
Only second swing	-1.889*** (0.678)	-4.018*** (0.970)	-4.585*** (1.053)
Only first auto-mailed ballot	1.919** (0.844)	2.869*** (1.077)	1.900* (1.004)
Both auto-mailed ballot	1.932 (2.032)	0.160 (2.743)	1.933 (4.393)
Only second auto-mailed ballot	-2.983*** (0.632)	-7.557*** (0.747)	-6.747*** (1.040)
Observations	583,835	173,075	62,305
R^2	0.007	0.017	0.018
Outcome mean (%)	9.77	11.16	11.58

Note: The variables *Only first ...*, *Both ...*, and *Only second ...* indicate whether the first state, both the first and the second state, or only the second state are swing states or have auto-mailed ballots. *Swing* is an indicator for the 13 states whose PredictIt winning likelihood for the less favored party on October 3, 2020 is at least 20% (PredictIt, 2020). *Auto-mailed ballot* denotes states that automatically sent mail-in ballots to registrants in 2020 (National Conference of State Legislatures, 2020b). All independent variables are divided by 100 for presentation purposes. Column 1 uses all double-registrants in the sample; column 2 double-registrants who voted in a 2020 primary election; column 3 double-registrants whose predicted likelihood of voting in the 2020 presidential election is 0.99 or higher (see note to Figure A3 for details). Standard errors in parentheses are clustered at the state-pair level.

*** 1%, ** 5%, * 10% significance level

Table 3: EVIDENCE FROM THE 2016 ELECTION

	Dependent var.: Vote in first state of registration		
	(1)	(2)	(3)
	All double-registrants	Near certain general election voters	
		Voter in primary	Predicted voter $Pr \geq 0.95$
Only first swing	3.185*** (0.823)	6.045*** (2.114)	10.763*** (2.955)
Both swing	2.366*** (0.874)	3.797* (2.099)	2.504 (1.547)
Only second swing	0.121 (0.564)	-2.429* (1.267)	-3.976*** (1.094)
Only first early voting	0.255 (0.647)	5.349*** (1.706)	6.610*** (2.288)
Both early voting	-1.609** (0.743)	-1.663 (1.560)	-0.658 (1.659)
Only second early voting	-1.751** (0.779)	-5.825*** (1.432)	-6.123*** (1.479)
Observations	665,641	113,743	58,857
R^2	0.003	0.020	0.032
Outcome mean (%)	9.48	11.93	12.66

Note: This table replicates Table 2 using double registrants from the L2 2016 data. Double registered individuals in this sample are identified from voter records with unique names – people whose first and last names are so uncommon that their names appear only twice, in two separate states, and also share the same birthdate. The variables *Only first ...*, *Both ...*, and *Only second ...* indicate whether the first state, both the first and the second state, or only the second state are swing states or have early voting. *Swing* is an indicator for the 12 states whose PredictIt winning likelihood for the less favored party on November 1, 2016 is at least 20% (PredictIt, 2020). *Early voting* refers to states that allowed voting for at least 19 days before Nov 8, 2016 (the same cutoff used for the 2020 election) (National Conference of State Legislatures, 2020a). The sample is composed of 2016 double registrants who were still registered to vote in 2017, to ensure that they were eligible to vote in the 2016 election. All independent variables are divided by 100 for presentation purposes. Column 1 uses all double-registrants in the sample; column 2 double-registrants who voted in a 2016 primary election; column 3 double-registrants whose predicted likelihood of voting in the 2016 presidential election is 0.95 or higher. Standard errors in parentheses are clustered at the state-pair level.

*** 1%, ** 5%, * 10% significance level

Table 4: ALTERNATIVE SAMPLE FOR THE 2020 ELECTION

	Dependent var.: Vote in first state of registration		
	(1)	(2)	(3)
	All double-registrants	Near certain general election voters	
		Voter in primary	Predicted voter $\text{Pr} \geq 0.99$
Only first swing	3.448*** (0.703)	4.722*** (1.057)	6.598*** (1.304)
Both swing	0.341 (0.788)	0.231 (1.174)	1.425 (1.137)
Only second swing	-1.231 (0.828)	-3.183** (1.240)	-2.553*** (0.875)
Only first auto-mailed ballot	3.227** (1.397)	4.325*** (1.662)	1.505* (0.845)
Both auto-mailed ballot	2.359 (2.342)	1.483 (3.389)	6.138 (5.153)
Only second auto-mailed ballot	-3.421*** (0.556)	-8.483*** (0.720)	-5.219*** (0.915)
Observations	804,610	203,247	60,677
R^2	0.005	0.017	0.016
Outcome mean (%)	10.34	11.86	9.51
Fraction in cellphone sample (%)	17.57	20.61	22.99

Note: This table replicates Table 2 using an alternative sample of double registrants from the 2020 L2 data. Double registered individuals in this sample are identified from voter records with unique names – people whose first and last names are so uncommon that their names appear only twice, in two separate states, and also share the same birthdate. The overlap between the main, cellphone sample of double registrants and this sample is reported in the bottom row. Standard errors in parentheses are clustered at the state-pair level.

*** 1%, ** 5%, * 10% significance level

Table 5: SAMPLE OF DUAL-STATE HOMEOWNERS FOR THE 2020 ELECTION

	Dependent var.: Vote in first state of registration		
	(1)	(2)	(3)
	All double-registrants	Near certain general election voters	
		Voter in primary	Predicted voter $Pr \geq 0.99$
Only first swing	11.689*** (1.632)	11.502*** (2.556)	13.084*** (2.327)
Both swing	0.563 (1.838)	-3.046 (3.392)	-1.825 (2.886)
Only second swing	-6.690*** (1.092)	-15.224*** (1.960)	-11.654*** (1.651)
Only first auto-mailed ballot	4.489*** (1.110)	7.255*** (2.231)	4.727** (2.178)
Both auto-mailed ballot	5.183** (2.328)	0.116 (2.892)	4.133 (2.604)
Only second auto-mailed ballot	-0.561 (2.020)	-6.727** (2.725)	-2.238 (2.797)
Observations	25,308	8,084	13,363
R^2	0.027	0.042	0.040
Outcome mean (%)	17.63	25.01	22.01

Note: This table replicates Table 2 using double registrants who own a property in each of the two states they are registered to vote in. We identify homeownership by matching double registrants to CoreLogic Tax Assessment data (the most recent data from the 2021 file) using first name, last name, and the latitude and longitude of an address (requiring the distance between a CoreLogic address and a voter address to be no greater than 1 mile). The double-registrant sample is the union of the cellphone sample used in Table 2 and the unique-name sample used in Table 4. Everything else follows Table 2. Standard errors in parentheses are clustered at the state-pair level.

*** 1%, ** 5%, * 10% significance level

Table 6: RELATIVE CLOSENESS OF ELECTIONS AND CSSV BEHAVIOR IN 2020

	Dependent var.: Vote in first state of registration
First win gap	-0.043*** (0.014)
Second – first win gap	0.034*** (0.009)
Only first auto-mailed ballot	2.410*** (0.846)
Both auto-mailed ballot	2.456 (1.992)
Only second auto-mailed ballot	-2.909*** (0.676)
Observations	583,835
R^2	0.008
Outcome mean (%)	9.77

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. The auto-mailed ballot variables are divided by 100 for presentation purposes. Other variable definitions and sample selection follow those in Table 2 column 1. Standard errors in parentheses are clustered at the state-pair level.

*** 1%, ** 5%, * 10% significance level

Table 7: HOW INCENTIVES AND COSTS AFFECT VOTING IN BOTH STATES OF REGISTRATION IN 2020

	Dependent var.: Vote in both states of registration		
	(1)	(2)	(3)
	All double-registrants	Near certain general election voters	
		Voter in primary	Predicted voter $Pr \geq 0.99$
Only first swing	0.262** (0.125)	0.687*** (0.249)	1.805*** (0.449)
Both swing	0.415** (0.165)	1.019** (0.407)	1.854*** (0.708)
Only second swing	0.009 (0.122)	-0.124 (0.264)	-0.296 (0.412)
Only first auto-mailed ballot	1.229*** (0.211)	1.893*** (0.426)	2.556*** (0.647)
Both auto-mailed ballot	1.122*** (0.212)	0.788** (0.308)	0.506 (0.556)
Only second auto-mailed ballot	0.258** (0.111)	-0.190 (0.212)	-0.147 (0.425)
Observations	583,835	173,075	62,305
R^2	0.002	0.002	0.004
Outcome mean (%)	1.01	1.88	3.24

Note: Independent variables and samples mirror those in the corresponding columns in Table 2, including dividing independent variables by 100 for presentation purposes. Standard errors in parentheses are clustered at the state-pair level.

*** 1%, ** 5%, * 10% significance level

ONLINE APPENDIX

CROSS-STATE STRATEGIC VOTING

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

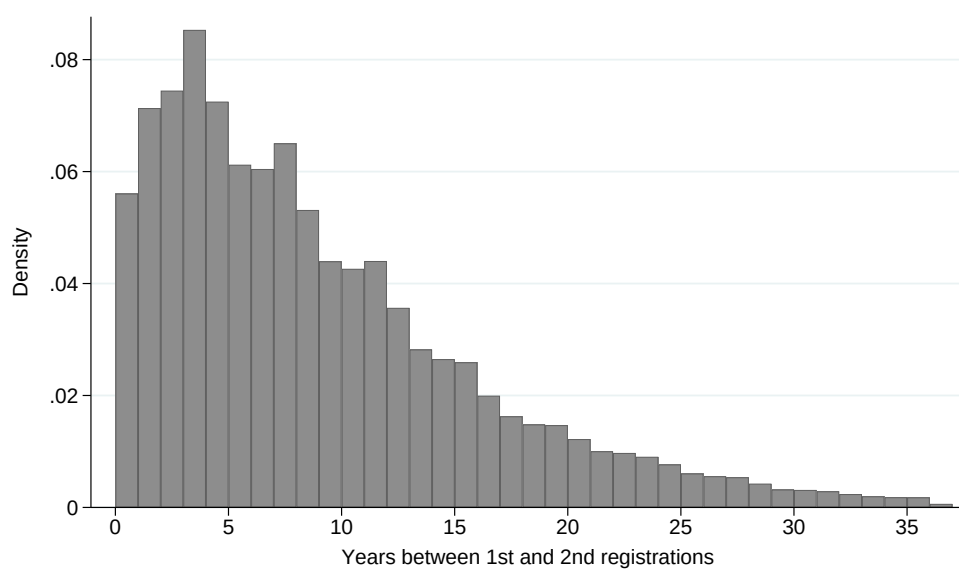


Figure A1: NUMBER OF YEARS BETWEEN FIRST AND SECOND REGISTRATIONS

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

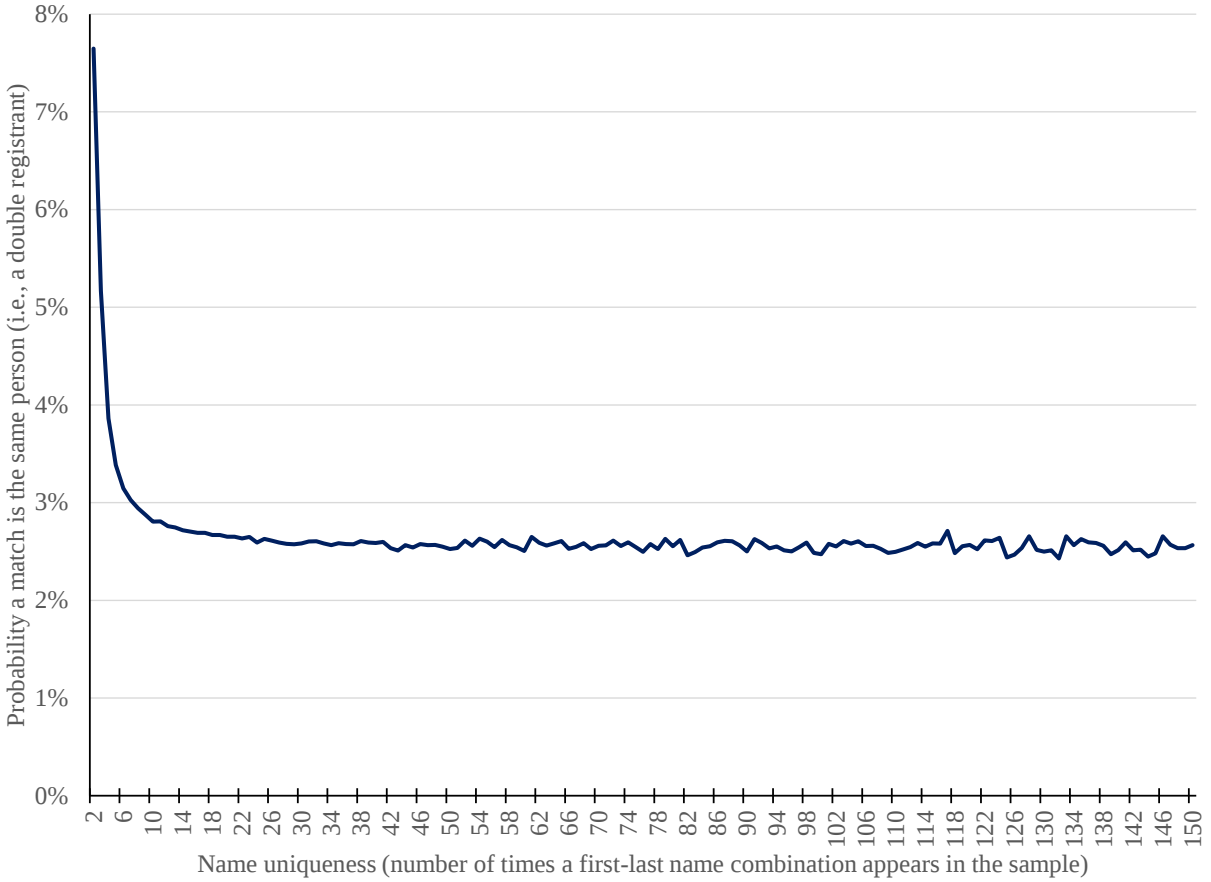


Figure A2: PROBABILITY A NAME MATCH ACROSS STATES IS THE SAME PERSON BY LEVEL OF NAME UNIQUENESS

Note: This figure plots the estimated probability that two registration 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.

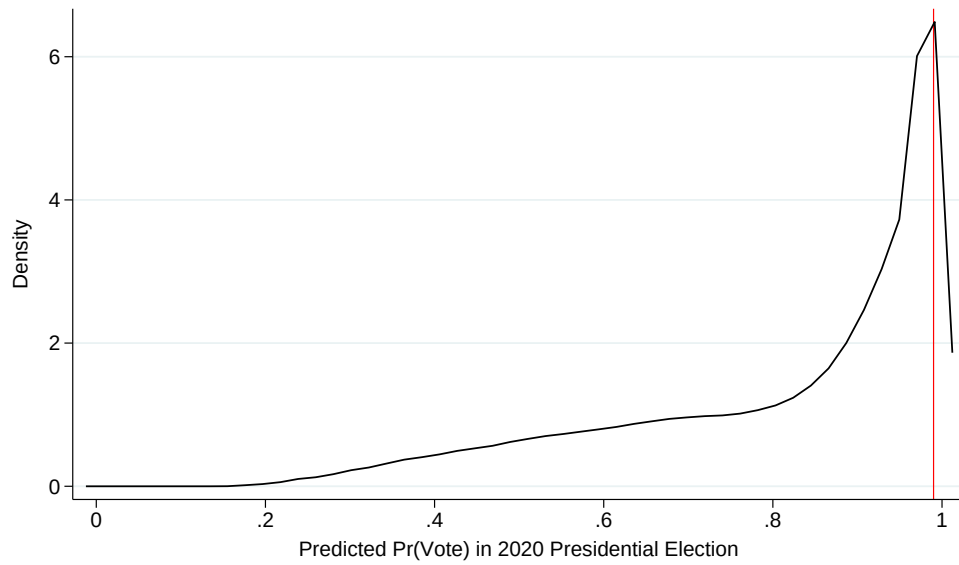
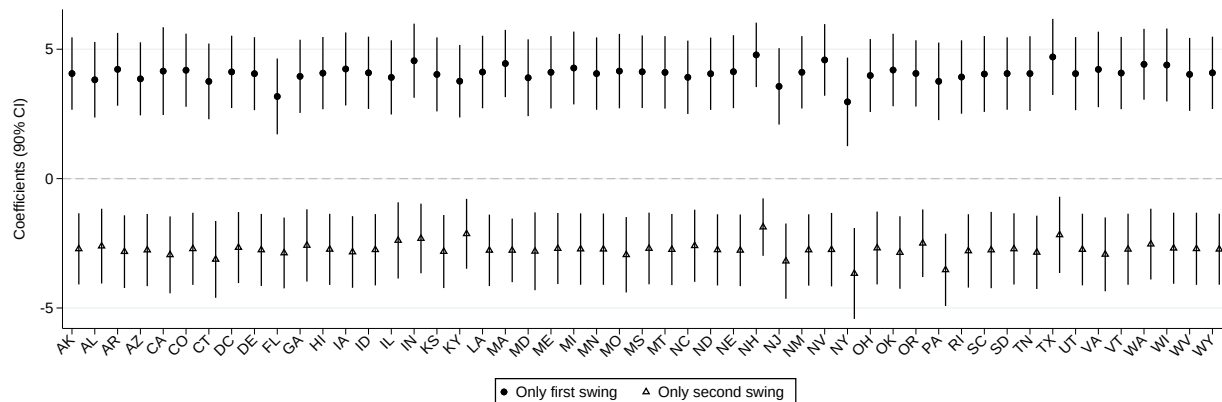
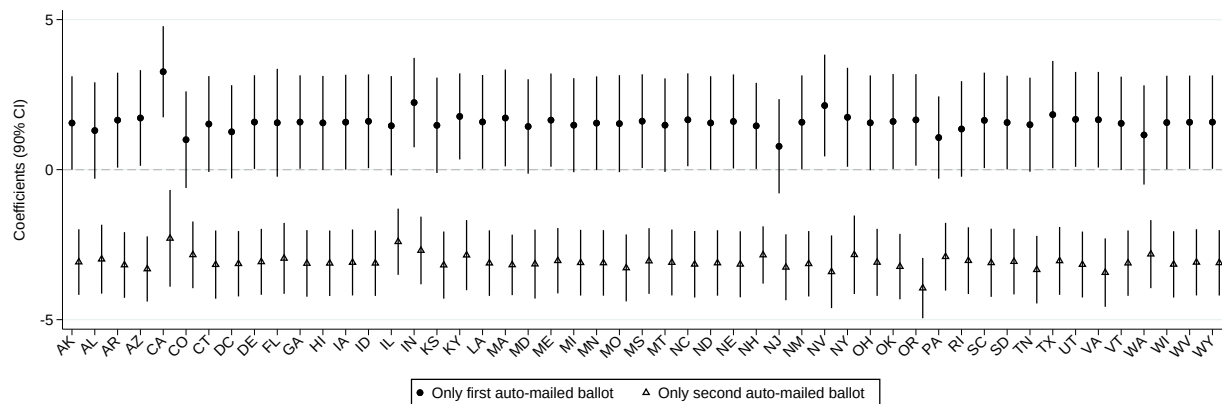


Figure A3: PREDICTED PROBABILITY OF VOTING IN THE 2020 PRESIDENTIAL ELECTION

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 the following variables as predictors: 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 7: a predicted voting probability of 0.99.



(a) CSSV incentive estimates



(b) CSSV cost estimates

Figure A4: ROBUSTNESS: EXCLUDING ONE STATE AT A TIME

Note: This figure plots the coefficient estimates (and 95% confidence intervals) for CSSV incentive and cost variables using the specification in Table 2 column 1 while excluding one state at a time. Estimates for *Only first swing* and *Only second swing* are plotted in panel (a) and those for *Only first auto-mailed ballot* and *Only second auto-mailed ballot* in panel (b). The excluded state (either first or second state) is labeled on the x-axis. Everything else follows Table 2 column 1.

Table A1: VOTER REGISTRATIONS BY STATE: L2 DATA AND EAVS

State	Total registration			Active registration			Inactive registration		
	L2	EAVS	L2/EAVS (%)	L2	EAVS	L2/EAVS (%)	L2	EAVS	L2/EAVS (%)
AK	544,865	646,093	84.3	532,119	595,647	89.3	12,746	50,446	25.3
AL	3,360,248	3,717,798	90.4	2,994,930	3,438,213	87.1	365,318	279,585	130.7
AR	1,560,875	1,831,414	85.2	1,175,076	1,408,061	83.5	385,799	423,353	91.1
AZ	4,089,871	4,728,109	86.5	3,781,982	4,275,729	88.5	307,889	452,380	68.1
CA	21,124,608	-	-	20,877,284	-	-	247,325	-	-
CO	3,667,315	4,211,528	87.1	3,578,369	3,803,762	94.1	88,946	407,766	21.8
CT	2,303,818	2,524,717	91.3	2,124,782	2,335,860	91	179,036	188,857	94.8
DC	454,482	625,683	72.6	429,570	517,890	82.9	24,912	107,793	23.1
DE	689,262	739,672	93.2	648,370	711,287	91.2	40,892	28,385	144.1
FL	13,875,629	15,218,424	91.2	12,955,508	14,517,002	89.2	920,121	701,422	131.2
GA	6,832,209	7,618,436	89.7	6,462,992	7,194,889	89.8	369,217	423,547	87.2
HI	771,937	-	-	714,106	-	-	57,831	-	-
IA	2,100,838	2,243,758	93.6	1,966,810	2,094,770	93.9	134,028	148,988	90
ID	898,418	1,029,763	87.2	839,004	1,029,763	81.5	59,414	-	-
IL	8,096,315	9,789,893	82.7	7,465,362	9,103,542	82	630,953	686,351	91.9
IN	4,236,229	4,692,091	90.3	3,652,052	4,170,353	87.6	584,177	521,738	112
KS	1,716,961	1,913,573	89.7	1,602,484	1,764,949	90.8	114,477	148,624	77
KY	3,161,547	3,565,428	88.7	2,912,392	3,319,307	87.7	249,155	246,121	101.2
LA	2,953,020	3,093,004	95.5	2,836,316	2,963,901	95.7	116,704	129,103	90.4
MA	4,439,104	4,812,909	92.2	3,922,652	4,400,254	89.1	516,452	412,655	125.2
MD	4,029,633	4,298,942	93.7	3,879,585	4,142,347	93.7	150,048	156,595	95.8
ME	983,978	1,138,576	86.4	934,708	1,135,008	82.4	49,270	3,568	1380.9
MI	7,294,117	8,105,524	90	6,761,976	7,209,300	93.8	532,141	896,224	59.4
MN	3,487,062	3,731,016	93.5	3,436,140	3,731,016	92.1	50,922	-	-
MO	3,985,784	4,338,133	91.9	3,590,033	3,963,980	90.6	395,751	374,153	105.8
MS	1,934,546	2,143,149	90.3	1,736,102	1,982,632	87.6	198,444	160,517	123.6
MT	640,332	747,439	85.7	571,625	675,971	84.6	68,707	71,468	96.1
NC	6,462,612	7,372,608	87.7	5,690,869	6,607,121	86.1	771,743	765,487	100.8
NE	1,147,639	1,266,730	90.6	1,050,614	1,168,708	89.9	97,025	98,022	99
NJ	5,851,577	6,310,564	92.7	5,454,103	5,896,836	92.5	397,474	413,728	96.1
NM	1,103,371	1,360,871	81.1	1,079,136	1,255,669	85.9	24,235	105,202	23
NV	1,753,625	2,039,162	86	1,524,037	1,835,401	83	229,588	203,761	112.7
NY	12,278,622	13,554,842	90.6	11,375,576	12,362,997	92	903,046	1,191,845	75.8
OH	7,442,327	8,073,829	92.2	6,797,392	8,073,829	84.2	644,935	-	-
OK	1,981,465	2,259,107	87.7	1,696,503	2,021,846	83.9	284,962	237,261	120.1
OR	3,079,424	2,944,588	104.6	2,756,422	2,944,588	93.6	323,002	-	-
PA	8,317,758	9,035,061	92.1	7,629,443	8,280,348	92.1	688,315	754,713	91.2
RI	755,008	809,117	93.3	694,819	735,195	94.5	60,189	73,922	81.4
SC	3,312,165	3,854,209	85.9	2,892,210	3,535,061	81.8	419,955	319,148	131.6
SD	537,360	635,256	84.6	468,534	578,683	81	68,826	56,573	121.7
TN	3,995,188	4,436,727	90	3,710,495	4,226,928	87.8	284,693	209,799	135.7
TX	15,933,926	16,955,520	94	14,121,565	15,279,870	92.4	1,812,361	1,675,649	108.2
UT	1,382,572	1,861,977	74.3	1,238,198	1,713,297	72.3	144,374	148,680	97.1
VA	5,493,823	5,975,561	91.9	5,205,669	5,763,187	90.3	288,154	212,374	135.7
VT	454,367	489,277	92.9	405,685	440,920	92	48,682	48,357	100.7
WA	4,783,984	5,255,466	91	4,462,776	4,892,871	91.2	321,208	362,595	88.6
WI	4,577,219	3,834,164	119.4	3,169,168	3,834,164	82.7	1,408,051	-	-
WV	1,145,489	1,269,024	90.3	969,037	1,062,685	91.2	176,452	206,339	85.5
WY	233,910	303,049	77.2	226,281	303,049	74.7	7,629	-	-

Note: This table reports the numbers of total, active, and inactive registrations in the L2 data and in EAVS for states in our sample. “-” indicates missing data in EAVS.

Table A2: STATE ELECTION CHARACTERISTICS

State	Swing PredictIt	Auto-mailed ballot	Swing NYT	Swing Ex-ante	Swing Ex-post	Early Voting	Auto absentee Application	Auto voter Registration	Election day Registration	Fast Removal	Frequent Update	High Tax
AK	N	N	N	Y	N	N	N	Y	N	Y	Y	N
AL	N	N	N	N	N	Y	N	N	N	N	N	N
AR	N	N	N	N	N	N	N	N	N	Y	N	Y
AZ	Y	N	Y	Y	Y	Y	N	N	N	Y	N	N
CA	N	Y	N	N	N	Y	N	Y	Y	N	Y	Y
CO	N	Y	N	N	N	N	N	Y	Y	Y	Y	N
CT	N	N	N	N	N	N	Y	Y	Y	N	Y	Y
DC	N	Y	N	N	N	N	N	Y	Y	Y	N	Y
DE	N	N	N	N	N	N	Y	N	N	N	N	Y
FL	Y	N	Y	Y	Y	N	N	N	N	N	N	N
GA	Y	N	Y	Y	Y	Y	N	Y	N	N	Y	N
HI	N	Y	N	N	N	N	N	N	Y	-	N	Y
IA	Y	N	Y	Y	Y	Y	Y	N	Y	N	Y	Y
ID	N	N	N	N	N	N	N	N	Y	-	Y	N
IL	N	N	N	N	N	Y	Y	Y	Y	N	Y	Y
IN	N	N	N	N	N	Y	N	N	N	Y	N	N
KS	N	N	N	N	N	Y	N	N	N	N	Y	N
KY	N	N	N	N	N	N	N	N	N	N	N	N
LA	N	N	N	N	N	N	N	N	N	Y	Y	N
MA	N	N	N	N	N	N	Y	Y	N	Y	Y	Y
MD	N	N	N	N	N	N	Y	Y	Y	N	Y	Y
ME	N	N	N	N	Y	Y	N	N	Y	Y	N	Y
MI	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	N
MN	Y	N	Y	Y	Y	Y	N	N	Y	N	Y	Y
MO	N	N	N	Y	N	N	N	N	N	Y	Y	N
MS	N	N	N	N	N	N	N	N	N	N	N	N
MT	N	Y	N	Y	N	Y	N	N	Y	Y	N	N
NC	Y	N	Y	Y	Y	Y	N	N	N	Y	N	N
NE	N	N	N	Y	N	Y	Y	N	N	Y	Y	Y
NJ	N	Y	N	N	N	Y	N	Y	N	N	N	Y
NM	N	N	N	N	N	N	Y	Y	N	N	Y	Y
NV	Y	Y	N	Y	Y	N	N	Y	Y	Y	N	N
NY	N	N	N	N	N	N	N	N	N	N	N	Y
OH	Y	N	Y	Y	Y	Y	Y	N	N	Y	Y	N
OK	N	N	N	N	N	N	N	N	N	Y	N	N
OR	N	Y	N	N	N	N	N	Y	N	N	Y	Y
PA	Y	N	Y	Y	Y	Y	N	N	N	Y	N	N
RI	N	N	N	N	N	Y	Y	Y	N	N	Y	Y
SC	N	N	N	Y	N	N	N	N	N	N	N	N
SD	N	N	N	N	N	Y	N	N	N	N	Y	N
TN	N	N	N	N	N	Y	N	N	N	Y	N	N
TX	Y	N	N	Y	Y	N	N	N	N	Y	Y	N
UT	N	Y	N	Y	N	N	N	N	Y	N	Y	Y
VA	N	N	N	N	N	Y	N	Y	Y	Y	N	Y
VT	N	Y	N	N	N	Y	N	Y	Y	Y	N	Y
WA	N	Y	N	N	N	N	N	Y	Y	N	N	Y
WI	Y	N	Y	Y	Y	N	Y	N	Y	N	Y	Y
WV	N	N	N	N	N	N	N	Y	N	Y	N	N
WY	N	N	N	N	N	Y	N	N	Y	Y	N	N

Note: This table lists state election characteristics (Y=yes and N=no) in the 2020 presidential election for states in our sample. See note to Tables 2, A3, and A5 for definitions of characteristics.

Table A3: ROBUSTNESS: STATE SUBSETS AND ALTERNATIVE DEFINITIONS OF SWING

	(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	Drop non-NVRA states
Only first swing	5.632*** (0.785)	4.552*** (0.839)	4.726*** (0.794)	5.242*** (0.820)	4.979*** (0.808)
Both swing	3.283*** (1.071)	1.346 (0.848)	1.439 (0.897)	1.752* (0.940)	1.630* (0.939)
Only second swing	-0.437 (0.668)	-1.958*** (0.744)	-1.896*** (0.680)	-1.873*** (0.698)	-1.925*** (0.683)
Only first auto-mailed ballot	2.122** (0.857)	1.914** (0.850)	1.933** (0.845)	1.787** (0.874)	1.903** (0.857)
Both auto-mailed ballot	3.342* (1.967)	1.813 (2.000)	1.941 (2.033)	1.892 (2.127)	1.867 (2.043)
Only second auto-mailed ballot	-2.043*** (0.662)	-3.054*** (0.654)	-2.984*** (0.632)	-3.105*** (0.644)	-3.053*** (0.636)
Observations	583,835	583,835	583,835	558,627	570,629
R^2	0.007	0.007	0.007	0.008	0.007
Outcome mean (%)	9.77	9.77	9.77	9.74	9.82

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 in columns 1-3. 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 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. Column 5 excludes states exempt from the NVRA: Idaho, Minnesota, Wyoming, and Wisconsin (North Dakota and New Hampshire are already excluded because they do not provide information on registration date). Standard errors in parentheses are clustered at the state-pair level.

*** 1%, ** 5%, * 10% significance level

Table A4: HETEROGENEITY BY INCOME AND AGE

	Dependent var.: Vote in first state of registration			
	Low income	High income	Age 18-64	Age 65-99
Only first swing	4.193*** (0.937)	5.021*** (0.708)	4.348*** (0.728)	6.941*** (1.547)
Both swing	0.948 (0.921)	1.719* (0.905)	1.243 (0.915)	1.386 (1.014)
Only second swing	-2.331*** (0.773)	-1.584*** (0.571)	-1.751*** (0.677)	-3.381*** (0.987)
Only first auto-mailed ballot	1.692* (1.018)	2.369*** (0.684)	1.915** (0.922)	2.103*** (0.706)
Both auto-mailed ballot	2.489 (2.720)	2.044 (1.560)	1.555 (1.816)	4.680 (3.936)
Only second auto-mailed ballot	-2.399*** (0.727)	-2.604*** (0.579)	-3.109*** (0.589)	-0.543 (1.434)
Observations	290,104	290,260	502,877	80,958
R^2	0.006	0.008	0.006	0.014
Outcome mean	10.78	8.62	9.47	11.6
Mean age	45.7	43.67	40.15	72.82

Note: This table reports heterogeneity analyses by income and age using the same specification as Table 2 column 1. Columns 1 and 2 split our double-registrant sample into those registering in counties with above vs. below median zipcode-level adjusted gross income in 2020 (from Individual Income Tax Statistics). Columns 3 and 4 split the sample into individuals below vs. above age 65. Standard errors in parentheses are clustered at the state-pair level.

*** 1%, ** 5%, * 10% significance level

Table A5: ROBUSTNESS: ADDITIONAL STATE CONTROLS

	Dependent var.: Vote in first state of registration							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Only first swing	4.576*** (0.761)	4.544*** (0.808)	4.531*** (0.831)	4.635*** (0.771)	3.528*** (0.867)	4.656*** (0.783)	4.742*** (0.793)	3.083*** (0.927)
Both swing	1.169 (0.758)	1.402 (0.881)	1.073 (0.906)	1.143 (0.883)	1.197* (0.715)	1.545** (0.767)	1.437 (0.902)	0.253 (1.049)
Only second swing	-1.852*** (0.604)	-1.808*** (0.696)	-2.178*** (0.692)	-2.145*** (0.685)	-1.380** (0.624)	-2.049*** (0.656)	-1.906*** (0.687)	-2.260*** (0.784)
Only first auto-mailed ballot	2.658*** (0.786)	1.538* (0.868)	2.221*** (0.836)	2.245*** (0.847)	2.526*** (0.882)	2.172*** (0.688)	1.916** (0.843)	2.491*** (0.824)
Both auto-mailed ballot	3.079 (2.033)	1.756 (2.037)	2.962 (1.928)	3.768 (2.495)	1.600 (2.111)	2.565 (2.033)	1.713 (2.081)	2.506 (2.048)
Only second auto-mailed ballot	-2.655*** (0.667)	-2.822*** (0.663)	-2.513*** (0.836)	-2.225** (0.988)	-3.688*** (0.657)	-2.877*** (0.666)	-3.012*** (0.643)	-2.797*** (0.588)
Observations	583,835	583,835	583,835	583,835	581,624	583,835	583,835	583,835
R^2	0.008	0.007	0.007	0.007	0.011	0.008	0.007	0.008
Outcome mean (%)	9.77	9.77	9.77	9.77	9.76	9.77	9.77	9.77
Early voting controls	Y	N	N	N	N	N	N	N
Auto absentee app. controls	N	Y	N	N	N	N	N	N
Auto voter reg. controls	N	N	Y	N	N	N	N	N
Election-day reg. controls	N	N	N	Y	N	N	N	N
EAVS removal rate controls	N	N	N	N	Y	N	N	N
State reg. date controls	N	N	N	N	N	Y	N	N
Migration flow controls	N	N	N	N	N	N	Y	N
Tax controls	N	N	N	N	N	N	N	Y

Note: This table reports robustness results for Table 2 column 1 by adding state-pair level controls. For all columns except column 7, we implement this by including indicators for whether only the first state equals 1 (e.g., *Only first early voting*), for whether only the second state equals 1 (e.g., *Only second early voting*), or both (e.g., *Both early voting*). Column 1 controls for state early voting status, defined as allowing voting at least 19 days before the election (the median). Columns 2, 3, and 4 control for automatic mailing of absentee ballot applications, automatic voter registration, and election-day registration, respectively. Column 5 controls for whether a state has an above-median (6.7%) voter removal rate using the 2020 Election Administration and Voting Survey (EAVS); the removal rate is calculated as the total number of voters removed from the voter registration rolls between the 2018 and the 2020 general elections, divided by the sum of the number removed plus the number remaining. Column 6 controls for whether a state has an above-median (20.3%) likelihood that the state registration date differs from the date in L2. Column 7 includes 2018-2019 state-pair level migration flows from the first to the second state per 1,000 state residents using data from the IRS Statistics of Income. Column 8 controls for whether a state has an above-median (10.8%) 2020 tax burden, as defined by the Tax Foundation, 2020. Standard errors in parentheses are clustered at the state-pair level.

*** 1%, ** 5%, * 10% significance level

Table A6: HETEROGENEITY BY PARTY

	Dep. var.: Vote in first state of registration		
	(1) Democrat	(2) Republican	(3) Independent
Only first swing	4.319*** (0.697)	4.961*** (1.234)	4.915*** (0.966)
Both swing	1.766* (0.984)	0.574 (1.380)	1.464* (0.872)
Only second swing	-1.416** (0.715)	-2.911*** (0.959)	-1.839** (0.767)
Only first auto-mailed ballot	1.471 (1.010)	2.370*** (0.841)	2.337*** (0.811)
Both auto-mailed ballot	0.572 (1.711)	4.668* (2.396)	2.551 (2.413)
Only second auto-mailed ballot	-3.454*** (0.559)	-1.103 (1.144)	-2.855*** (0.778)
Observations	273,859	126,316	183,660
R^2	0.006	0.010	0.007
Outcome mean (%)	9.07	9.72	10.84

Note: This table reports heterogeneity analyses by party affiliation using the same specification as Table 2 column 1. Columns 1, 2, and 3 consist of registered Democrats, Republicans, and Independents, respectively. Party affiliation is based on an individual's second state of registration. Standard errors in parentheses are clustered at the state-pair level.

*** 1%, ** 5%, * 10% significance level

Table A7: QUANTIFYING CSSV’S EFFECT ON THE 2020 ELECTION

	CSSV		Actual 2020 vote margin	Estimated 2020 vote gain for Democratic and Republican presidential candidates							
	induced votes										
	(1)	(2)		(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
						<i>Incentive</i>			<i>Cost</i>		<i>Total</i>
	Inflow	Outflow	Win-Loss	Dem	Rep	Dem-Rep	Dem	Rep	Dem-Rep	Dem-Rep	
AZ	2,944	1,637	10,457	1,283	1,225	58	-722	-533	-190	-131	
GA	5,444	1,982	11,779	2,339	1,682	657	-493	-196	-296	360	
WI	2,873	709	20,682	1,443	1,118	325	-341	-152	-189	136	
NV	5,653	1,347	33,596	1,474	1,384	90	924	299	626	715	
NC	4,162	1,075	74,481	1,835	1,645	191	-314	-188	-126	65	
PA	13,405	4,100	80,555	7,563	4,790	2,773	-2,322	-1,061	-1,261	1,512	
IA	2,397	551	138,611	1,176	940	236	-235	-124	-112	124	
MI	2,828	830	154,188	1,397	947	450	-342	-127	-215	236	
MN	613	327	233,012	196	199	-4	-83	-40	-42	-46	
FL	25,265	6,238	371,686	11,542	10,928	614	-2,515	-1,601	-914	-300	
OH	6,998	1,695	475,669	3,180	2,475	705	-589	-244	-345	360	
TX	17,082	5,943	631,221	7,699	7,068	632	-2,615	-1,488	-1,127	-495	

Note: This table performs a back-of-the-envelope calculation quantifying CSSV effects in swing states. It uses (i) estimates of the number of double-registrants by party for each ordered state pair, (ii) estimates of the party-specific CSSV coefficients from Table A6, and (iii) exit polls for how party affiliation translates into votes for a candidate (Schaffner et al., 2021). As an example, we first estimate that there are 5,440 (Florida, New York) Republican double-registrants, 22,164 (Florida, New York) Democrats, and 10,933 (Florida, New York) Independents using the procedure in Section 3. Second, we combine these numbers with the CSSV coefficients from Table A6 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 270 ($= 5,440 \times 4.961/100$) votes from Republican double-registrants, 957 votes from Democrats, and 161 votes from Independents flowed out of New York and into Florida. Third, exit polls in New York indicate that 87.5% of Republicans voted for Trump and 11.5% for Biden, so the 270 incentive-driven votes among (Florida, New York) Republican double-registrants led to a 236 vote gain for Trump and 31 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.

DETAILS ON QUANTIFYING CSSV’S EFFECT ON THE 2020 ELECTION

This appendix provides details on the back-of-the-envelope calculations underlying Table A7, which aims to quantify the effects of CSSV on the 2020 election. This calculation requires several assumptions, including extrapolating estimates from our regression sample to the broader voter population and inferring voting behavior from party affiliation. Columns 1 and 2 of Table A7 report 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 our 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.

Whether CSSV behavior affected the outcome of an election depends not only on how many CSSV-induced votes occur, but also on whether one party engaged in it 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. The calculated inflows and outflows in columns 1 and 2 use party-specific estimates for the number of double-registrants in each state and for the CSSV coefficients.

In Section 2 we showed that there are approximately twice as many Democratic as Republican double-registrants. Appendix Table A6 runs the regression from Table 2 column 1 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). There is some evidence that Republicans are more responsive to incentives than Democrats.

The remaining columns in Appendix Table A7 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 2,339 votes due to incentive-based CSSV, while the Republican candidate (Trump) gained 1,682, for a net difference of +657, which means Biden gained more incentive-based CSSV votes. In contrast, the net difference for cost-based CSSV is -296, which means Trump gained more cost-based CSSV votes. Combining the two effects, CSSV added only 360 votes in favor of Trump in Georgia on net, despite there being 7,426 ($=5,444 + 1,982$) total CSSV-induced votes.

The patterns observed in Georgia are present more generally across the other swing states, none of which had auto-mailed ballots except Nevada.¹ First, incentive-based CSSV generally favored Biden because of the larger number of Democratic double-registrants. Second, cost-based CSSV generally favored Trump by disproportionately pulling votes away from Biden in swing states because more Democrats reside in auto-mail states. In summary, there were a sizable number of CSSV votes in 2020 – 107,000 of which involve at least one swing state – but the *net* effect was small, because Republican and Democratic CSSV-induced votes largely canceled out.

¹Given the estimated impacts, by construction cost-based CSSV effects in swing states are negative (except for Nevada) because none of them are auto-mail states.