Voting From Jail

March 12, 2024

Abstract

We leverage new data on daily individual-level jail records and exploit the timing of incarceration to estimate the causal effects of jail incarceration on voting from jail in 2020. We find that registered voters booked into county jails for the full duration of 2020 voting days were on average 41% less likely to vote in 2020, relative to registered voters booked into the same jails within 7 – 42 days after Election Day. Black registered voters booked into county jails for the full duration of 2020 voting days were on average 68% less likely to vote in 2020, relative to Black registered voters booked into the same jails after Election Day. Jail incarceration also reduces voter registration from jail, resulting in an even larger total effect of incarceration on voting from jail. Our findings reveal the pressing need to enable voting-eligible incarcerated individuals to exercise their constitutional right to vote, and to address racial disparities in the effect of jail incarceration on the exercise of that right.

1 Introduction

In every state, otherwise voting-eligible individuals incarcerated pretrial or to serve misdemeanor sentences remain legally entitled to vote while incarcerated (Paikowsky, 2019). The vast majority of the approximately 650,000 individuals incarcerated in county jails on any given day are being held pretrial or to serve misdemeanor sentences (Sawyer and Wagner, 2022). Anecdotal evidence suggests, however, that many of those incarcerated in county jails are being given inadequate opportunity to exercise their right to vote while jailed (The Sentencing Project, 2020). Concerns have also been raised about possible racial disparities in ballot access for those incarcerated in county jails (The Sentencing Project, 2022). Yet we lack reliable causal estimates of the impact of being incarcerated in a county jail on the exercise of one's right to vote while jailed, and of any racial heterogeneity in this impact.

Prior work has estimated the post-release impacts of periods of prison and jail incarceration on individuals' exercise of their right to vote (Gerber et al., 2017; White, 2019*b*; McDonough, Enamorado and Mendelberg, 2022). Gerber et al. (2017) used state prison and sentencing records to investigate the impacts of prior periods of prison incarceration on post-release voting in the 2012 election, finding few differences in turnout between those incarcerated and released from state prison between 2008 and 2012, and either observably comparable individuals in the Pennsylvania voter files, or observably comparable convicted defendants who received sentences other than prison incarceration. Gerber et al. (2017) also found few differences in post-release turnout between those convicted defendants sentenced to jail incarceration on their most serious charge, and those receiving non-carceral sentences.

White (2019*b*) addressed potential bias arising from selection into incarceration by estimating the effects of a sentence to jail incarceration on post-release 2012 turnout using a design based on the as-if random assignment of first-time misdemeanor defendants to judges in Harris County, Texas. White (2019*b*) found a 13 percentage-point average decrease in post-release 2012 turnout among marginal Black misdemeanor defendants sentenced between 2008 and 2012, and negligible effects for marginal white defendants.

McDonough, Enamorado and Mendelberg (2022) used a similar design to estimate impacts of pretrial jail incarceration between 2008 and 2012 in Philadelphia County, Pennsylvania on post-

release 2012 turnout, leveraging the assignment of defendants to bail judges to identify causal effects. McDonough, Enamorado and Mendelberg (2022) found a 39% average reduction in post-release voter turnout among marginal defendants incarcerated pretrial before the 2012 election, with effects appearing in the year prior to the election, and with larger effects for Black defendants.[1]

Incarceration may affect turnout not only after a defendant is released, but also, for those defendants incarcerated during voting days, during periods of incarceration (Paikowsky, 2019). White and Nguyen (2022) merged state prison records with voter file records in Maine and Vermont—the two states in which individuals incarcerated for felony convictions are legally entitled to vote while incarcerated—finding that 8% and 6% of those incarcerated in state prison voted while incarcerated in 2018 in Vermont and Maine, respectively. White and Nguyen (2022) did not attempt to estimate the causal impacts of incarceration on the incidence of voting from prison. McDonough, Enamorado and Mendelberg (2022) attempted to distinguish the incapacitation effect of jail incarceration on Election Day 2012 from the post-release effects of prior periods of jail incarceration, but their design and data did not allow them to cleanly identify such an effect.

The ability to vote from jail may differ by the race of the incarcerated individual. White (2019*b*) found that Black individuals were less likely to vote after having served incarceration sentences for misdemeanor offenses, finding no effect for white individuals. McDonough, Enamorado and Mendelberg (2022) likewise found that Black individuals were less likely to vote after experiencing pretrial incarceration, finding no effect for white or Hispanic individuals. However, to date no study has investigated whether there are racially heterogeneous effects of incarceration on the ability to vote from jail.

We leverage new data on daily individual-level jail records and exploit the timing of incarceration to estimate the causal effects of jail incarceration on voting from jail in 2020. We probabilistically match 944,985 individual-level booking records from 936 jail rosters with 195,655,326 voter records from 42 corresponding statewide files for a period of 180 days centered on Election Day (November 3, 2020). We interpret matches as registered voters who were incarcerated in county jails during this time period. We then identify individuals who were booked into jail during 2020 voting days in their states (either mail-in or in-person), or whose periods of incarceration began within windows

[1]The assignment of bail judges to cases in Philadelphia County during this period is not independent of case and defendant characteristics ($p < 0.001$), raising the possibility that non-randomly assigned case and defendant characteristics may account for some of the estimated effect.

ranging from 7 to 42 days after Election Day. Our primary analyses are conducted within the sample of jailed registered voters who match voter records with probability $p > 0.75$; we also replicate these analyses after including match probability as a covariate, and after restricting the sample to match probability $p > 0.95$. We also use the full sample of jailed individuals to estimate the impacts of jail incarceration on voter registration and on turnout unconditional on registration status. We source information about jailed individuals from both jail and voter records for the sample of jailed registered voters, and from jail records for the full sample of jailed individuals.

Our identification strategy rests on the assumption that individuals booked into the same jail within narrow temporal windows, who are indistinguishable on observed characteristics, are also indistinguishable on unobserved characteristics. If this assumption is correct, we can use the 2020 turnout of individuals booked into county jails just after the last 2020 voting day as a valid counterfactual for the turnout we likely would have observed among individuals booked into the same jails during 2020 voting days, had they not been incarcerated during those voting days. We choose post-Election Day control group windows ranging between between 7 and 42 days to balance competing concerns about the comparability of treatment and control groups and sample size. We conduct balance tests on individual-level and booking-level characteristics to confirm that individuals booked into county jails during 2020 voting days were observably indistinguishable from individuals booked into the same jails during post-Election Day control group windows ranging between 7 and 42 days.

Turnout among registered voters in 2020 is estimated by the U.S. Census Bureau to have been 91.9%.[2] Registered voters booked into county jails just after the final 2020 voting day had much lower average turnout, underscoring the importance of identifying an appropriate counterfactual turnout rate for those jailed during 2020 voting days. For the sample with match probability $p > 0.75$, turnout among registered voters who were booked into county jails within 7 – 42 days after 2020 Election Day ranged between 34.9% and 35.7% in our samples, or between 61% – 62% less than the registered voter turnout observed in the U.S. population. Without an appropriate control group, we might mistakenly attribute some of the decreased turnout among those jailed prior to the election to the effects of jail incarceration, rather than to a lower propensity to vote.

We estimate the impacts of (1) being booked into a county jail during 2020 voting days (either

[2] https://www.census.gov/content/dam/Census/library/publications/2022/demo/p20-585.pdf.

mail-in or in-person) for any length of time and (2) being booked into a county jail during 2020 voting days for some proportion of voting days on the probability that a registered voter voted in the 2020 election.

In models that include all individual- and booking-level covariates along with jail and week fixed effects, we find that registered voters who were booked into county jails during 2020 voting days for any length of time were on average 3.0 – 3.3 percentage points or 8.6% – 9.2% less likely to have voted in 2020, relative to registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 8.9%; $p < 0.01$; see Figure A.1).

The impacts of incarceration on voting from jail in 2020 increased with the duration of incarceration. Including all individual- and booking-level covariates along with jail and week fixed effects, registered voters who were booked into county jails during 2020 voting days for the full span of voting days in their state were on average 13.7 – 15.1 percentage points or 38.4% – 43.3% less likely to have voted in 2020, relative to registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 40.6%; $p < 0.01$; see Figure A.2). In a series of placebo tests, we find no consistent effects of the timing of individuals' 2020 bookings on their probabilities of having voted in the 2012 or 2016 elections.

Estimates of the effects of jail incarceration are larger in the smaller samples of records that match with probability $p > 0.95$. In these samples, registered voters who were booked into county jails during 2020 voting days in their state for any length of time were on average 4.1 – 4.4 percentage points or 13.7% – 14.1% less likely to have voted in 2020, relative to registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 13.9%; $p < 0.01$). Registered voters in these samples who were booked into county jails during 2020 voting days in their state for the full duration of voting days were on average 20 - 22.1 percentage points or 64.1% – 73.7% less likely to have voted in 2020, relative to registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 68.6%; $p < 0.01$).

We also explore racial heterogeneity in the effects of incarceration on voting from jail in 2020, finding that the effect of jail incarceration was significantly larger for Black registered voters, relative to white registered voters. Including all individual- and booking-level covariates along with jail and week fixed effects, white registered voters who were booked into county jails during 2020

voting days for the full set of voting days were on average 12.1 – 13.1 percentage points or 31.2% – 34.6% less likely to have voted in 2020, relative to white registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 32.8%; $p < 0.01$). The estimated negative effect of incarceration on voting from jail was an additional 7.2 – 8.4 percentage points or 57.9% – 64.1% larger for Black registered voters (avg. effect magnitude = 60.5%; $p < 0.01$). The total estimated effect of jail incarceration for Black registered voters booked into county jails during 2020 voting days for the full span of voting days was a 19.3 – 21.5 percentage point or 62.7% – 73.4% decrease in turnout, relative to Black registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 67.6%; see Figures A.1 and A.2). Estimates of racially heterogeneous effects are of comparable magnitude when we include the states for which race is predicted by L2.

We also explore the effects of jail incarceration on voter registration behavior, and on voting behavior unconditional on registration status. Jail incarceration in 2020 had small effects on voter registration for those detained for any length of time, but larger effects for those detained for longer periods of time. Including all covariates and fixed effects, individuals booked into county jails for the full span of 2020 voting days in their state were on average 9.3 – 11 percentage points or 31.7% – 37.9% less likely to have been registered to vote before Election Day 2020, relative to individuals booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 34.5%; $p < 0.01$). The effects of jail incarceration on turnout unconditional on registration status are correspondingly larger than the effects conditional on registration, because they include the negative effects of jail incarceration on registration. Including all covariates and fixed effects, individuals booked into county jails for the full duration of 2020 voting days in their state were on average 7.9 – 9 percentage points or 91.8% – 111.11% less likely to have voted in 2020, unconditional on registration status, relative to individuals booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 100.1%; $p < 0.01$).

The nature of the data we use prevent us from exploring the precise mechanisms generating both the overall decreases in turnout for individuals booked into county jails during 2020 voting days, and the substantially larger decreases for Black individuals. Nevertheless, at a minimum our findings reveal a pressing need for states and counties to take steps to ensure that voting-eligible incarcerated individuals are given adequate opportunities to exercise their constitutional right to

vote, and to address troubling racial disparities in the exercise of that right.

2 Voting From Jail

The Supreme Court ruled in 1974 that otherwise voting-eligible individuals incarcerated in county jails may not be denied access to the means to exercise their constitutional right to vote (O'Brien v. Skinner, 414 U.S. 524 (1974)). The vast majority of those held in county jails are being held pretrial or for the purpose of serving a sentence following a misdemeanor conviction, neither of which is disenfranchising in any state.[3] However, incarcerated individuals are not at liberty to freely investigate registration and voting eligibility and procedures, take steps to register to vote and secure ballot access, and cast their ballots. Those who are incarcerated require the assistance of correctional authorities to complete each of these steps. But correctional authorities may not provide this critical assistance.

In the first instance, correctional authorities may not be aware that many of the individuals incarcerated in county jails retain legal eligibility to register and vote. Even if they are aware of this fact, they may not be aware of their county's procedures and deadlines for providing registration and ballot access to the incarcerated. Even if they are aware of registration and voting eligibility and procedures, they may not take proactive steps to provide the requisite information and materials to those incarcerated. They may attempt to provide the requisite information and materials, but instead provide inaccurate and/or incomplete information and materials. They may also refuse to provide this information and materials to those who request it (Paikowsky, 2019; The Sentencing Project, 2020, 2022).

The empirical question is whether and by how much jail incarceration interferes with an individual's exercise of the right to vote. Those who are incarcerated in county jails during voting days may already have been unlikely to investigate registration and voting procedures, take steps to register to vote and secure ballot access, and cast their ballots, even had they not been incarcerated. Further, those who are not incarcerated may not provide an adequate comparison group to assess the impact of jail incarceration on voting. To date we lack any reliable causal estimates of whether

[3] Like individuals not incarcerated in county jails, those incarcerated in county jails may be ineligible to register and vote for reasons unrelated to their current incarceration, including citizenship status and prior felony convictions. For more information on enfranchisement status within incarcerated populations, see The Sentencing Project (2020).

and by how much being incarcerated in a county jail interferes with an individual's exercise of the right to vote.

3 Data

3.1 Jail Incarceration Data

We source daily individual-level jail incarceration data from the Jail Data Initiative (JDI).[4] Daily JDI jail records are scraped from county and municipal websites that post public rosters of all incarcerated individuals on (at least) a daily basis. By Fall 2020, the JDI project was scraping daily jail rosters from more than 1,100 county jails. Jails are not included in the JDI data if they do not post online jail rosters (e.g., all jails in the state of Alaska), or if their states have unified correctional systems (e.g., Connecticut and Hawaii).

The Jail Data Initiative converts daily jail roster records into booking records by identifying unique individuals and unique periods of incarceration within each facility. Unique individuals are identified using a variety of fields including booking number, person identification number, name, and date of birth, depending on field availability. Unique periods of incarceration or booking spells are identified using the dates on which unique individuals enter and exit jail rosters (unique individuals may have multiple booking records in the same county). Demographic and bail/bond fields are standardized across facilities. Charge descriptions contained in text fields are classified into charge categories using the Text-Based Offense Classification algorithm developed by the Criminal Justice Administrative Records System at the University of Michigan (CJARS) (Choi et al., 2022).

The JDI bookings database contains daily jail incarceration data unprecedented in granularity and national scope. While the Bureau of Justice Statistics produces a 5-year Census of Jails—most recently in 2019—and an intervening Annual Survey of Jails, these data are only based on single-day point-in-time snapshots of jail populations (Zhen Zeng and Minton, 2021).

For our main analyses of the effects of jail incarceration on the turnout of registered voters, we include as covariates JDI data on the length of each booking in days, the number of charges associated with each booking, and indicators for whether the most serious CJARS charge category associated with each booking was violent, property, drug, public order, or DUI, with criminal

[4] https://jaildatainitiative.org/.

traffic as the excluded charge category.[5] For our analyses of the effects of jail incarceration on voter registration status, and on turnout unconditional on registration status, we additionally include as covariates JDI data on gender, age, and race. Further information on JDI data collection and standardization is provided in Appendix C.

3.2 Voter File Data

We source voter file data from the L2 voter database. These data contain voting history and demographic information for registered voters.[6] L2 data were pulled in early 2021. For our main analyses of the effects of jail incarceration on the turnout of registered voters, we include as covariates L2 data on age, gender, race/ethnicity, and party registration.

L2 race and ethnicity categories are both reported by states to L2 and, where not reported by states, predicted by L2 from other features available in the data. For our main estimates of racial heterogeneity in the effect of jail incarceration on voting we use only states that directly report race data to L2; we replicate these analyses using the full L2 sample including both reported and predicted race in Appendix A. We map L2 race/ethnicity values to non-Hispanic white (European), Black (Likely African-American) and Other (East and South Asian, Hispanic and Portuguese, or Other).

L2 party categories are likewise both reported directly by states and, where not reported by states, predicted by L2 from other features available in the data. We map the L2 party categories to Democratic, Republican, and Non-Partisan or Other (a group that includes Non-Partisan, Other, and 54 other minor party descriptions). Further information on L2 voter data is provided in Appendix C.

3.3 Matching Jail Records to Voter File Records

We initially constructed a sample of 944,985 JDI jail booking records from 936 counties for individuals who were booked into county jails on dates within a +/- 90-day window around Election Day 2020 (August 5, 2020 to February 1, 2021). To make probabilistic record linkage tractable, we

[5] The CJARS Text-Based Offense Classification algorithm uses the Uniform Crime Classification Standard (UCCS) to categorize charges into charge categories, and determines most serious charge using a hierarchical ranking of top-level UCCS codes based on the following order of severity, from most to least serious: violent, property, drug, public order, DUI, criminal traffic (Choi et al., 2022).

[6] https://l2-data.com/our-data/.

created age strata for JDI bookings and blocked on L2 within-state matching pools of +/- 2 years. Where age was missing for JDI booking records, we matched across full-state L2 voter files. We further restricted blocks to L2 surname soundex matches of incarcerated individuals.

Data completeness varies by jail roster. We employed a Java implementation of the probabilistic matching algorithm used by McDonough, Enamorado and Mendelberg (2022), itself an instance of the Fellegi-Sunter/Expectation Maximization linkage model as developed by Enamorado, Fifield and Imai (2019). In lieu of labeled matches we produced a baseline averaged Jaro-Winkler distance between name components, and considered all scores above a threshold to be matches for parametric initialization. We then vectorized similarities of name, gender, and in-block age as ternary (to account for missingness and variation in reporting of name components), and FIPS codes as binary (within-state). We ran samples from this pool of match-candidate pairs through a standard expectation maximization process (Winkler, 2000) to stabilize parameters, and then conducted our final linkage, keeping only highest-scoring matches per booked individual and discarding low-probability matches.

We followed a similar strategy of probability match score re-weighting based on first and last name frequencies to that employed by Enamorado, Fifield and Imai (2019) and McDonough, Enamorado and Mendelberg (2022), and then further thresholded matched results. We also removed any rows indicating multiple voter matches per booked individual, where multiple matched bookings exhibited temporal overlap, or where booking records indicate that individuals were younger than 18 years old. Our primary analyses are conducted within the sample with match probability $p > 0.75$; we also replicate these analyses after including match probability as a covariate, and after restricting the sample to match probability $p > 0.95$. We also use the full sample of jailed individuals to estimate the impacts of jail incarceration on voter registration and on turnout unconditional on registration status.

3.4 2020 Voting Days

In response to the COVID-19 pandemic, many states enacted emergency measures that impacted voting days and methods during the 2020 general election. Some states suspended requirements for requesting absentee mail ballots, instead permitting no-excuse absentee ballot requests; others automatically sent out mail-in absentee ballots to registered voters for the first time, or extended

early in-person voting days (Weiser et al., 2020). There was a 105% increase in voting by mail from the 2016 general election to the 2020 general election, and a 37% increase in early in-person voting (Scherer, 2021).

Mail-in ballot and early in-person voting days varied both across and within states. For each state we record the earliest of the two dates on which voters could begin (1) returning the mail-in ballots mailed out by county boards of elections, or (2) voting in person. Earliest mail-in ballot dates range from September 4, 2020 (60 days prior to Election Day) in North Carolina to October 16, 2020 (19 days prior to Election Day) in Washington state. Earliest in-person voting dates range from September 14, 2020 (50 days prior to Election Day) in Pennsylvania to Election Day itself in several states (National Conference of State Legislatures, 2022; Higgins and Rattner, 2020). By taking the earlier of these two dates in each state, we ensure that our treatment windows cover the complete legal voting periods of each state.

4 Analysis

Individuals incarcerated in jails or prisons likely differ along a number of dimensions, relative to unincarcerated individuals. These differences may be related to individuals' propensity to vote. Of particular concern, individuals at high risk for incarceration may have low propensity to vote, leading researchers to overestimate the impacts of incarceration on voting.

Gerber et al. (2017) sought to address selection into incarceration by controlling for observed characteristics correlated with both risk of incarceration and propensity to vote. However, this strategy does not address the possible presence of unobserved characteristics related to both risk of incarceration and propensity to vote. White (2019*a*) and McDonough, Enamorado and Mendelberg (2022) addressed selection into jail incarceration by leveraging the potentially as-if random assignment of cases to trial and bail judges. However, this strategy does not permit clean identification of the effects of incarceration on voting from jail, since whether an individual remains incarcerated during legal voting periods is affected by factors other than a judge's bail determination or sentencing decision (McDonough, Enamorado and Mendelberg, 2022).

We leverage the timing of jail bookings to address selection into jail incarceration. Individuals booked into county jails during legal voting days in 2020 are likely very similar to individuals booked

into those same jails just after the last legal voting day (November 3, 2020 in all states). In our design, incarcerated individuals are effectively randomized into treatment and control groups based on the timing of their jail incarceration, with treated individuals booked into county jails during legal voting days, and control individuals booked into those same jails just after Election Day. This identification strategy has been successfully deployed in a number of recent papers (White, 2019*b*; Mueller-Smith and Schnepel, 2021; Anker, Doleac and Landersø, 2021). Following this literature, we report balance tests to confirm the similarity of treatment and control groups.

We initially restrict our attention to estimating the effect of jail incarceration on the voting behavior of registered voters booked into county jails during our period of interest, using a match probability of $p > 0.75$ as indicating a jailed registered voter. We replicate analyses after including the match probability as a covariate, and after restricting the sample to a match probability of $p > 0.95$. We then estimate the effects of jail incarceration on both registration and voting behavior unconditional on registration status.

4.1 Construction of Treatment and Control Samples

We expect that registered voters booked into county jails during 2020 voting days (our treatment group) will have individual- and booking-level characteristics that are similar to those of registered voters booked into the same jails just after Election Day (our control group). However, there may be a trade-off between the comparability of our treatment and control groups and the size of the control group. For example, registered voters booked into jail within the first week after Election Day might be most comparable to voters booked into county jails during 2020 voting days, but at the expense of the control group sample size. We might fail to reject the hypothesis of no differences between the treatment and control groups simply because our control group sample size would be quite small (Imai, King and Stuart, 2008). Conversely, the further from Election Day we expand the control group window, the less likely we are to spuriously "pass" a balance test because of small control group sample size, but also the less comparable our treatment and control samples may become.

We address this trade-off by choosing three post-election control group windows that span different versions of the comparability/sample size trade-off: 7, 21, and 42 days post-election. Table 1 reports estimates from models predicting whether a booking occurred during the treatment

window (1) or one of these three control windows (0) using a variety of individual- and booking-level characteristics, including age (from L2, in years), reported race/ethnicity (from L2, indicators for Black and non-Hispanic white), party registration (from L2, indicators for Democratic and Republican registration), gender (from L2, indicator for male), number of charges (from JDI), and indicators for whether the top charge against the booked individual was a DUI, drug, property, public order, or violent charge (from JDI). All models include jail-level fixed effects; robust standard errors clustered on jails are reported in parentheses. Samples include linked records with match probability $p > 0.75$, excluding observations missing data on any covariates, or where voting records indicated that an individual did not register to vote until after November 3, 2020. All coefficients are very close to zero and joint F-test p-values are all greater than 0.10.

Table 1: Balance Tests

Control Group Window:	**7 days**	**21 days**	**42 days**
Age (L2)	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Black (L2)	0.008	0.010	0.002
	(0.006)	(0.007)	(0.007)
White (L2)	0.000	0.001	-0.001
	(0.005)	(0.005)	(0.005)
Male (L2)	0.002	0.006	0.004
	(0.003)	(0.004)	(0.004)
Democrat (L2)	-0.007*	-0.002	-0.001
	(0.004)	(0.004)	(0.004)
Republican (L2)	-0.004	-0.006	-0.005
	(0.004)	(0.005)	(0.005)
Number of Charges (JDI)	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.001)
Top Charge: Drug (JDI)	0.009	-0.002	-0.007
	(0.010)	(0.010)	(0.010)
Top Charge: DUI (JDI)	0.003	-0.009	-0.011
	(0.010)	(0.011)	(0.011)
Top Charge: Property (JDI)	0.005	-0.005	-0.009
	(0.009)	(0.010)	(0.010)
Top Charge: Public Order (JDI)	0.007	-0.002	-0.002
	(0.009)	(0.010)	(0.010)
Top Charge: Violent (JDI)	0.007	-0.007	-0.016*
	(0.008)	(0.010)	(0.010)
Observations	63000	79854	103091
Joint F-Test p-value	0.614	0.372	0.182

Note: This table reports estimates of models predicting whether a booking occurred during the treatment window (1) or a control window (0), using a variety of individual- and booking-level characteristics, for post-Election Day control windows ranging between 7 and 42 days and a treatment window containing legal voting days in 2020, for the sample of matched records with match probability $p > 0.75$. All models include jail-level fixed effects; robust standard errors clustered on jails are reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.10.

We show in Appendix A that, for these three control group windows, the choice to include all 2020 voting days in the treatment group is not consequential for the comparability of the treatment and control groups. Figure A.3 reports estimates of balance on individual- and booking-level characteristics for treatment windows containing the full set of voting days in each state in 2020 and all subsets of those voting days. For each control group window, we estimated the models in Table 1 after sequentially dropping the earliest voting day in the sample and repeating the balance test, continuing until we reached a treatment window of 7 days prior to Election Day. Control and treatment groups are balanced not only for the full set of 2020 voting days, but also for all subsets of those voting days.

4.2 Descriptive Statistics

Table 2 reports summary statistics for the treatment and control groups for each of the estimation samples identified in Table 1. The average age in our sample ranges between 37.2 – 37.4 years. 24.6% - 27.1% of our samples are Black and 60.1% – 60.4% are white. 73.3% – 73.6% of our samples are male. 20% - 20.4% of our samples are Republicans, 41% – 41.9% are Democrats, and 37.8% – 38.9% are nonpartisan or other party. The average number of charges is 2.5 charges across all samples. 28.3% – 29.8% of those in our samples are incarcerated for public order top charges; 27.5% – 29.2% are incarcerated for violent top charges; 19.4% – 20% are incarcerated for property top charges; 13.1% – 13.8% are incarcerated for drug top charges; 6.2% – 6.7% are incarcerated for DUI top charges; 3.2% – 3.5% are incarcerated for criminal traffic top charges. The average length of jail incarceration ranges from 35.1 – 37.5 days. The average proportion of 2020 voting days during which individuals in the treatment group are confined is 21.5%.

Turnout among registered voters who were booked into county jails within 7 – 42 days after 2020 Election Day ranged between 34.9% and 35.7% in our samples, or between 61% and 62% less than the registered voter turnout observed in the U.S. population.[7]. As noted earlier, without an appropriate control group, we might mistakenly attribute some of the lower turnout of those incarcerated during 2020 voting days to the effects of jail incarceration, rather than to a lower propensity to vote.

[7] https://www.census.gov/content/dam/Census/library/publications/2022/demo/p20-585.pdf.

Table 2: Summary Statistics

Treatment/Control:	T	C	C	C
Control Group Days:		7d	21d	42d
Age (L2)	37.18	37.37	37.28	37.23
Black (L2)	0.271	0.246	0.247	0.253
White (L2)	0.601	0.604	0.603	0.601
Male (L2)	0.734	0.736	0.733	0.734
Democrat (L2)	0.417	0.419	0.410	0.410
Republican (L2)	0.200	0.203	0.204	0.201
Non-Partisan or Other Party (L2)	0.383	0.378	0.386	0.389
Number of Charges (JDI)	2.507	2.539	2.501	2.514
Top Charge: Criminal Traffic (JDI)	0.033	0.035	0.032	0.032
Top Charge: Drug (JDI)	0.138	0.131	0.131	0.131
Top Charge: DUI (JDI)	0.062	0.067	0.066	0.065
Top Charge: Property (JDI)	0.194	0.200	0.199	0.198
Top Charge: Public Order (JDI)	0.298	0.288	0.287	0.283
Top Charge: Violent (JDI)	0.275	0.279	0.285	0.292
Length of Stay (days) (JDI)	37.46	37.21	35.10	35.56
Confined During Voting Days	1.000	0.000	0.000	0.000
Proportion of Voting Days Confined	0.215	0.000	0.000	0.000
Turnout	0.323	0.349	0.354	0.357
Observations	53995	9005	25859	49096

Note: This table reports mean values of individual- and booking-level characteristics for the samples reported in Table 1, comprising matched records with match probability $p > 0.75$, by treatment and control groups.

4.3 Estimating Average Treatment Effects

Table 3 reports OLS estimates of the effects of (1) being booked into a county jail during 2020 voting days (either mail-in or in-person) for any length of time on a registered voter's probability of voting in the 2020 election, with and without covariates, and (2) being booked into a county jail during 2020 voting days for some proportion of voting days on a registered voter's probability of voting in the 2020 election, with and without covariates. For example, if a registered voter was incarcerated in North Carolina from September 4, 2020 to January 1, 2021, the proportion of voting days confined = 1.0, as the individual was incarcerated during 60 out of all 60 possible voting days in North Carolina. If a registered voter was incarcerated in Washington state from November 3, 2020 to December 10, 2020, the proportion of voting days confined = 0.053, as the individual was incarcerated during 1 out of 19 possible voting days in Washington state. Models are estimated on the samples reported in Table 1. All models include jail and week-of-year fixed effects; robust standard errors two-way clustered at the jail and week-of-year level are reported in parentheses.[8]

[8]In our design, paired clusters of individuals booked into the same jail are assigned to treatment or control based on the timing of the booking (before or after 2020 election days). The current econometric recommendation

Table 3: The Effect of Incarceration on Voting From Jail

Control Group Window:	**7 days**	**21 days**	**42 days**
Confined During Voting Days	-0.029***	-0.031***	-0.032***
	(0.002)	(0.002)	(0.002)
with covariates	-0.030***	-0.032***	-0.033***
	(0.002)	(0.002)	(0.002)
Proportion of Voting Days Confined	-0.169***	-0.166***	-0.162***
	(0.021)	(0.020)	(0.019)
with covariates	-0.151***	-0.142***	-0.137***
	(0.018)	(0.018)	(0.018)
Mean Proportion of Voting Days Confined	0.215	0.215	0.215
Max. Proportion of Voting Days Confined	1.000	1.000	1.000
Mean Control Turnout	0.349	0.354	0.357
Observations	63000	79854	103091

Note: This table reports OLS estimates of (1) the effect of being booked into a county jail during 2020 voting days for any length of time on a registered voter's probability of voting in the 2020 election, with and without covariates; and (2) the effect of the proportion of 2020 voting days during which a registered voter booked into a county jail during voting days was confined on their probability of voting in the 2020 election, with and without covariates, for samples of matched records (registered voters booked into county jails during the specified time periods) with match probability $p > 0.75$. Models are estimated on samples as described in the notes to Table 1. All models include jail and week-of-year fixed effects; robust standard errors two-way clustered at the jail and week-of-year level are reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.10.

As reported in Table 3, in models that include all individual- and booking-level covariates along with jail and week fixed effects, we find that registered voters who were booked into county jails during 2020 voting days in their state for any length of time were on average 3.0 – 3.3 percentage points or 8.6% – 9.2% less likely to have voted in 2020, relative to registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 8.9%; p < 0.01). These estimates imply that registered voters who were booked into county jails during 2020 voting days for the full span of voting in their state voted from jail at rates ranging between 19.8% and 22.0%.

The impacts of incarceration on voting from jail in 2020 increased with the duration of incarceration. As reported in Table 3, including all individual- and booking-level covariates along with jail and week fixed effects, registered voters who were booked into county jails during 2020 voting days in their state for the full duration of voting days were on average 13.7 - 15.1 percentage points or 38.4% – 43.3% less likely to have voted in 2020, relative to registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 40.6%; p $<$

in this context is to cluster standard errors at the level of the pair (the jail in our setting) (de Chaisemartin and Ramirez-Cuellar, 2024).

0.01). These estimates imply that registered voters who were booked into county jails during 2020 voting days for the full span of voting in their state voted from jail at rates ranging between 19.8% and 22.0%. Table A.1 replicates these models including match probability as a covariate, reporting very similar estimates.

4.4 Placebo Tests

Table 4 reports placebo tests of the effects on 2016 and 2012 voter turnout of a booking occurring during 2020 voting days, relative to a post-Election Day control window. The 2016 placebo samples include registered voters in our estimation samples who were registered to vote on both Election Day 2020 and Election Day 2016 (November 8, 2016). The 2016 placebo tests report OLS estimates of the effect of a 2020 booking occurring during 2020 voting days on a registered voter's probability of voting in the 2016 election, relative to a post-Election Day control window. The 2012 placebo samples include registered voters in our estimation samples who were registered to vote on both Election Day 2020 and Election Day 2012 (November 6, 2012). The 2012 placebo tests report OLS estimates of the effect of a 2020 booking occurring during 2020 voting days on a registered voter's probability of voting in the 2012 election, relative to a post-Election Day control window.

As reported in Table 4, and as expected, there are no consistent effects of pre-Election Day jail incarceration in 2020 on registered voters' turnout in 2016 or 2012. The effects of 2020 jail incarceration on 2016 turnout are very slightly positive but consistently insignificant. The effects of 2020 jail incarceration on 2012 turnout are very slightly negative but generally insignificant, reaching $p < 0.05$ in only one sample. By contrast, the effects of 2020 jail incarceration on 2020 turnout among registered voters who were also registered to vote in 2016 and/or 2012 are consistently negative, large, and significant, ranging between 2 – 3.6 percentage points or 5.0% – 7.8% decreases in turnout ($p < 0.01$).

4.5 Racially Disparate Effects of Jail Incarceration

Prior work has found that Black individuals are often treated differently in the criminal justice system, relative to individuals of other races. Black individuals experience less effective responses to their reports of crime victimization (Harvey and Mattia, 2022), are subjected to greater use of force by white police officers (Hoekstra and Sloan, 2022), are less likely to receive discounts on speeding

Table 4: Placebo Tests

	2016 Placebo		**2012 Placebo**	
Control Group Window: 7 days	2016	2020	2012	2020
Confined During Voting Days	0.005	-0.020***	-0.012	-0.024***
	(0.007)	(0.007)	(0.008)	(0.009)
Mean Control Turnout	0.442	0.399	0.554	0.450
Observations	40693	40693	28388	28388
Control Group Window: 21 days	2016	2020	2012	2020
Confined During Voting Days	0.004	-0.027***	-0.009	-0.033***
	(0.005)	(0.005)	(0.006)	(0.006)
Mean Control Turnout	0.441	0.406	0.544	0.459
Observations	51471	51471	35862	35862
Control Group Window: 42 days	2016	2020	2012	2020
Confined During Voting Days	0.001	-0.030***	-0.010**	-0.036***
	(0.004)	(0.004)	(0.005)	(0.005)
Mean Control Turnout	0.445	0.410	0.546	0.463
Observations	66386	66386	46045	46045

Note: This table reports placebo tests of the effects of 2020 jail incarceration on 2016 and 2012 voter turnout. The 2016 placebo samples include registered voters in each of our estimation samples who were registered to vote on both Election Day 2020 and Election Day 2016 (November 8, 2016). The 2012 placebo samples include registered voters in each of our estimation samples who were registered to vote on both Election Day 2020 and Election Day 2012 (November 6, 2012). All models include jail fixed effects; robust standard errors clustered at the jail level are reported in parentheses. ***p $<$ 0.01, **p $<$ 0.05, *p $<$ 0.10.

tickets (Goncalves and Mello, 2021), are more likely to be detained pretrial (Arnold, Dobbie and Yang, 2018), receive longer sentences from Republican judges (Cohen and Yang, 2019), are more likely to receive incarceration sentences from trial judges nearing reelection (Park, 2017), are more likely to have their appeals rejected by appellate judges nearing reappointment (Harvey and Yntiso, 2021), and are less likely to receive relief on death row appeals from appellate panels with no Black members (Kastellec, 2021). These findings suggest that Black individuals incarcerated in county jails during election days may be given differential access to the means to vote, resulting in racially disparate impacts of jail incarceration on voting from jail.

Black individuals may also respond differently to the experience of jail incarceration, relative to non-Black individuals. White (2019*b*) found that Black individuals were less likely to vote after having served incarceration sentences for misdemeanor offenses, finding no effect for white individuals. McDonough, Enamorado and Mendelberg (2022) likewise found that Black individuals were less likely to vote after experiencing pretrial incarceration, finding no effect for white or Hispanic

individuals. If Black individuals subjected to incarceration have higher rates of pre-incarceration turnout, then one mechanism driving these findings could be that there is more room for incarceration to depress Black turnout (White, 2019*b*). Alternatively, Black individuals may experience incarceration as having been more unjustly imposed than white civilians, leading to greater demobilizing effects for incarcerated Black individuals (White, 2019*b*; McDonough, Enamorado and Mendelberg, 2022).

We explore racial heterogeneity in the effects of incarceration on voting from jail in 2020 by interacting our two treatment variables (a binary indicator for having been booked into a county jail during 2020 voting days for any length of time, and the proportion of 2020 voting days during which a registered voter booked into a county jail during 2020 voting days was incarcerated) with the indicator for Black registered voters from the L2 voter files. We restrict the samples to only those states reporting race to L2. For ease of interpretation we also retain only non-Hispanic white and Black voters from the matched samples. All models include all covariates along with jail and week-of-year fixed effects; robust standard errors are two-way clustered at the jail and week-of-year level.

Table 5 reports estimates. While there is no discernible racial heterogeneity in the effects of jail incarceration on voting when using the binary measure of incarceration, a large racial gap in the effects of jail incarceration on voting emerges when we take into account the duration of incarceration during voting days. Including all individual- and booking-level covariates along with jail and week fixed effects, white registered voters who were booked into county jails during 2020 voting days for the full set of voting days were on average 12.1 – 13.1 percentage points or 31.2% – 34.6% less likely to have voted in 2020, relative to white registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 32.8%; $p < 0.01$). These estimates imply that white registered voters booked into county jails during 2020 voting days for the full set of voting days voted from jail at rates ranging between 24.8% – 26.7%.

The estimated negative effect of incarceration on voting from jail was an additional 7.2 – 8.4 percentage points or 57.9% – 64.1% larger for Black registered voters (avg. effect magnitude = 60.5%; $p < 0.01$). The total estimated effect of jail incarceration for Black registered voters booked into county jails during 2020 voting days for the full span of voting days was a 19.3 – 21.5 percentage point or 62.7% – 73.4% decrease in turnout, relative to Black registered voters who were booked

into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 67.6%; see Figure A.2). These estimates imply that Black registered voters booked into county jails during 2020 voting days for the full set of voting days voted from jail at rates ranging between only 7.8% and 11.5%.

Table 5: Racial Disparities in the Effect of Incarceration on Voting From Jail (With Covariates) States Reporting Race

Control Group Window:	**7 days**	**21 days**	**42 days**
Confined During Voting Days	-0.044***	-0.047***	-0.045***
	(0.008)	(0.008)	(0.007)
Confined × Black	0.008	0.008	0.003
	(0.008)	(0.010)	(0.009)
Proportion of Voting Days Confined	-0.131***	-0.126***	-0.121***
	(0.029)	(0.028)	(0.027)
Proportion Confined × Black	-0.084***	-0.073***	-0.072***
	(0.020)	(0.021)	(0.021)
Mean Proportion of Voting Days Confined Black	0.197	0.197	0.197
Mean Proportion of Voting Days Confined White	0.210	0.210	0.210
Mean Control Turnout Black	0.293	0.298	0.308
Mean Control Turnout White	0.379	0.385	0.388
Observations	28382	35843	45868

Note: This table reports OLS estimates of (1) the effect of being booked into a county jail during 2020 voting days for any length of time on a registered voter's probability of voting in the 2020 election, interacted with an indicator for whether the individual is Black (0 = non-Hispanic white), with covariates; and (2) the effect of the proportion of 2020 voting days during which a registered voter booked into a county jail during voting days was confined on their probability of voting in the 2020 election, interacted with an indicator for whether the individual is Black, with covariates. Models are estimated on samples as described in the notes to Table 1, including only Black and white individuals, and including only states that report voter race to L2. All models include jail and week-of-year fixed effects; robust standard errors two-way clustered at the jail and week-of-year level are reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.10.

The L2 data also include predicted race and ethnicity categories for states that do not report race. Table A.2 expands the analyses reported in Table 5 to include states for which race and ethnicity are predicted by L2 using other features available in the data. Effect magnitudes and significance levels in Table A.2 are similar to those reported in Table 5, for both white and Black registered voters.

Although in both White (2019*b*) and McDonough, Enamorado and Mendelberg (2022) incarcerated Black individuals had higher rates of pre-incarceration turnout, that is not the case in our samples. White registered voters booked into county jails just after Election Day 2020 had turnout rates ranging between 37.9% and 38.8%, while Black registered voters booked into county jails just after Election Day had turnout rates ranging between 29.3% and 30.8%. It is thus unlikely that

this mechanism is driving the racially heterogeneous effects of incarceration on voting from jail reported in Table 5. However, our data do not permit us to distinguish between racially disparate treatment of incarcerated Black individuals, and racially disparate responses to the experience of incarceration.

4.6 Match Probability $p > 0.95$

The linked records used in our main analyses have match probabilities that exceed 0.75. In Appendix B we replicate our primary analyses using only linked records with match probabilities that exceed 0.95. Table B.1 reports summary statistics for these treatment and control groups, which are again closely comparable on individual- and booking-level characteristics. Table B.2 reports estimates of the effect of jail incarceration on the voting behavior of registered voters for these samples. Effect magnitudes are larger than those reported for the sample of linked records with match probabilities exceeding 0.75. In models that include all individual- and booking-level covariates along with jail and week fixed effects, registered voters in these samples who were booked into county jails during 2020 voting days in their state for any length of time were on average 4.1 – 4.4 percentage points or 13.7% – 14.1% less likely to have voted in 2020, relative to registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 13.9%; $p < 0.01$). Registered voters in these samples who were booked into county jails during 2020 voting days in their state for the full duration of voting days were on average 20 - 22.1 percentage points or 64.1% – 73.7% less likely to have voted in 2020, relative to registered voters who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 68.6%; $p < 0.01$).

Table B.3 reports no consistent effects of pre-Election Day jail incarceration in 2020 on registered voters' turnout in 2016 or 2012, for the sample of linked records with match probabilities exceeding 0.95. The effects of 2020 jail incarceration on 2016 turnout are very slightly positive but consistently insignificant; the effects of 2020 jail incarceration on 2012 turnout are very slightly negative but again consistently insignificant. The effects of 2020 jail incarceration on 2020 turnout among registered voters who were also registered to vote in 2016 and/or 2012 are consistently negative, large, and significant.

4.7 The Effect of Jail Incarceration on Registration

Jail incarceration may also affect voting behavior by reducing the ability of jailed individuals to register to vote. If this was the case in 2020, then our estimates of the effect of jail incarceration on the voting behavior of registered voters understate the total effect of jail incarceration on voting from jail.

In Appendix C we replicate our analyses for the full sample of jailed individuals, including both registered and unregistered voters, defining a registered voter as a jailed individual who matched to L2 records with match probability $p > 0.95$ and was registered to vote before Election Day 2020. Individual- and booking-level characteristics for this sample are sourced only from the JDI database in order to have coverage on both unregistered and registered voters. Table C.1 reports summary statistics for the treatment and control samples, which are again quite comparable on all individual- and booking-level characteristics. Pre-Election Day voter registration rates among individuals who were booked into county jails within 7 – 42 days after 2020 Election Day ranged between 29% and 29.4% in our samples. 2020 turnout rates in these same post-Election day samples, unconditional on registration status, ranged between 8.1% and 8.6%.

Table C.2 reports estimates of the effect of jail incarceration during 2020 voting days on the probability that a jailed individual was registered to vote by Election Day 2020, for the three control group windows. In models that include all individual- and booking-level covariates along with jail and week fixed effects, there is a very small positive effect of being jailed for any length of time before Election Day on the probability of being registered to vote by Election Day (avg. effect magnitude = 0.02%; p < 0.01). The sign of the effect flips when incarceration is measured as proportion of voting days incarcerated. Including all individual- and booking-level covariates along with jail and week fixed effects, individuals booked into county jails for the full span of 2020 voting days in their state were on average 9.3 – 11 percentage points or 31.7% – 37.9% less likely to have been registered to vote before Election Day 2020, relative to individuals booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 34.5%; p < 0.01).

Table C.3 reports estimates of the effect of jail incarceration during 2020 voting days on turnout unconditional on registration status. Effect magnitudes for the impacts of being incarcerated for any length of time before Election Day are comparable to those estimated only on samples of

registered voters. In models that include all individual- and booking-level covariates along with jail and week fixed effects, individuals in this sample who were booked into county jails during 2020 voting days in their state for any length of time were on average 0.6 – 0.7 percentage points or 7.4% – 8.2% less likely to have voted in 2020, relative to individuals who were booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 7.9%; $p < 0.01$). Effect magnitudes for the impacts of being incarcerated for the full duration of voting days on turnout, unconditional on registration status, are larger than those estimated for registered voters only, because they incorporate the additional effect of jail incarceration on registration. Including all individual- and booking-level covariates along with jail and week fixed effects, individuals booked into county jails for the full duration of 2020 voting days in their state were on average 7.9 – 9 percentage points or 91.8% – 111.11% less likely to have voted in 2020, relative to individuals booked into the same jails within 7 – 42 days after Election Day (avg. effect magnitude = 100.1%; $p < 0.01$).

5 Conclusion

On any given day, approximately 650,000 individuals are being held in local jails to await trial or to serve a misdemeanor sentence (Sawyer and Wagner, 2022). In no state does incarceration pretrial or to serve a misdemeanor sentence deprive an individual of their constitutional right to vote, including while they are incarcerated. Yet incarcerated individuals are not free to investigate registration and voting rules and procedures, take the steps necessary to register to vote and to secure ballot access, and return their ballots. They require the assistance of correctional authorities to complete each of these steps. But correctional authorities may not provide this critical assistance.

Whether and by how much jail incarceration interferes with an individual's exercise of the right to vote are empirical questions, ones for which we have lacked good estimates. We leverage new data on daily individual-level jail records and exploit the timing of incarceration to estimate the causal effects of jail incarceration on voting from jail in 2020. We find that registered voters booked into county jails for the full duration of 2020 voting days were on average 41% less likely to vote in 2020, relative to registered voters booked into the same jails within 7 – 42 days after Election Day. The estimated negative effect of incarceration on voting from jail was much larger for Black registered voters, who were on average 68% less likely to vote in 2020 if booked into county jails

for the full duration of 2020 voting days, relative to Black registered voters booked into the same jails just after Election Day. Placebo tests indicate no effects of 2020 jail incarceration on the 2012 or 2016 turnout of registered voters. We also find that jail incarceration additionally reduces voter registration, which amplifies the effect of incarceration on voting from jail. Our findings reveal a pressing need to enable voting-eligible incarcerated individuals to exercise their constitutional right to vote, and to address troubling racial disparities in the effect of jail incarceration on the exercise of the right to vote.

References

Anker, Anne Sofie Tegner, Jennifer L. Doleac and Rasmus Landersø. 2021. "The Effects of DNA Databases on the Deterrence and Detection of Offenders." *American Economic Journal: Applied Economics* 13(4):194–225.

Arnold, David, Will Dobbie and Crystal S Yang. 2018. "Racial Bias in Bail Decisions." *The Quarterly Journal of Economics* 133(4):1885–1932.

Choi, Jay, David Kilmer, Michael Mueller-Smith and Sema A. Taheri. 2022. Hierarchical Approaches to Text-based Offense Classification. Working paper University of Michigan.

Cohen, Alma and Crystal S. Yang. 2019. "Judicial Politics and Sentencing Decisions." *American Economic Journal: Economic Policy* 11(1):160–91.

de Chaisemartin, Clément and Jaime Ramirez-Cuellar. 2024. "At What Level Should One Cluster Standard Errors in Paired and Small-Strata Experiments?" *American Economic Journal: Applied Economics* 16(1):193–212.

Enamorado, Ted, Benjamin Fifield and Kosuke Imai. 2019. "Using a Probabilistic Model to Assist Merging of Large-Scale Administrative Records." *American Political Science Review* 113(2):353–371.

Gerber, Alan S., Gregory A. Huber, Marc Meredith, Daniel R. Biggers and David J. Hendry. 2017. "Does Incarceration Reduce Voting? Evidence about the Political Consequences of Spending Time in Prison." *The Journal of Politics* 79(4):1130–1146.

Goncalves, Felipe and Steven Mello. 2021. "A Few Bad Apples? Racial Bias in Policing." *American Economic Review* 111(5):1406–41.

Harvey, Anna and Sidak Yntiso. 2021. "Judicial Accountability and Racial Disparity in Criminal Appeals." *The Journal of Legal Studies* 50(2):261–302.

Harvey, Anna and Taylor Mattia. 2022. "Reducing racial disparities in crime victimization: Evidence from employment discrimination litigation." *Journal of Urban Economics* .

Higgins, Tucker and Nate Rattner. 2020. "When does vote by mail and early voting start? A state by state guide." https://www.cnbc.com/2020/09/04/election-early-absentee-mail-voting-every-state.html.

Hoekstra, Mark and CarlyWill Sloan. 2022. "Does Race Matter for Police Use of Force? Evidence from 911 Calls." *American Economic Review* 112(3):827–60.

Imai, Kosuke, Gary King and Elizabeth A. Stuart. 2008. "Misunderstandings between experimentalists and observationalists about causal inference." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171(2):481–502.

Kastellec, Jonathan P. 2021. "Race, Context, and Judging on the Courts of Appeals: Race-Based Panel Effects in Death Penalty Cases." *Justice System Journal* 42(3-4):394–415.

McDonough, Anne, Ted Enamorado and Tali Mendelberg. 2022. "Jailed While Presumed Innocent: The Demobilizing Effects of Pretrial Incarceration." *The Journal of Politics* 84(3).

Mueller-Smith, Michael and Kevin Schnepel. 2021. "Diversion in the Criminal Justice System." *The Review of Economic Studies* 88(2):883–936.

National Conference of State Legislatures. 2022. "Table 7: When States Mail Out Absentee/Mail Ballots." https://www.ncsl.org/research/elections-and-campaigns/vopp-table-7-when-states-mail-out-absentee-ballots.aspx.

Paikowsky, Dana. 2019. "Jails as Polling Places: Living Up to the Obligation to Enfranchise the Voters We Jail." *Harvard Civil Rights-Civil Liberties Law Review* 54(2):829–873.

Park, Kyung H. 2017. "The Impact of Judicial Elections in the Sentencing of Black Crime." *Journal of Human Resources* 52(4):998–1031.

Sawyer, Wendy and Peter Wagner. 2022. "Mass Incarceration: The Whole Pie 2022." https://www.prisonpolicy.org/reports/pie2022.html.

Scherer, Zachary. 2021. "Majority of Voters Used Nontraditional Methods to Cast Ballots in 2020." https://www.census.gov/library/stories/2021/04/what-methods-did-people-use-to-vote-in-2020-election.html.

The Sentencing Project. 2020. "Voting in Jails." https://www.sentencingproject.org/publications/voting-in-jails/.

The Sentencing Project. 2022. "Voting in Jails: Advocacy Strategies to #UnlocktheVote." https://www.sentencingproject.org/publications/voting-in-jails-advocacy-strategies-to-unlockthevote/.

Weiser, Wendy R., Eliza Sweren-Becker, Dominique Erney and Anne Glatz. 2020. "Mail Voting: What Has Changed in 2020." https://www.brennancenter.org/our-work/research-reports/mail-voting-what-has-changed-2020.

White, Ariel. 2019*a*. "Family Matters? Voting Behavior in Households with Criminal Justice Contact." *American Political Science Review* 113(2):607–613.

White, Ariel. 2019*b*. "Misdemeanor Disenfranchisement? The Demobilizing Effects of Brief Jail Spells on Potential Voters." *American Political Science Review* 113(2):311–324.

White, Ariel and Avery Nguyen. 2022. "How Often Do People Vote While Incarcerated? Evidence from Maine and Vermont." *Journal of Politics* 84(1):568–572.

Winkler, William E. 2000. Using the EM Algorithm for Weight Computation in the Fellegi-Sunter Model of Record Linkage. In *Proceedings of the Section on Survey Research Methods, American Statistical Association*. pp. 667–671.

Zhen Zeng, Ph.D. and Todd D. Minton. 2021. Jail Inmates in 2019. Technical Report NCJ 255608 U.S. Department of Justice: Bureau of Justice Statistics.

A Appendix A: Registered Voters; Match Probability $p > 0.75$

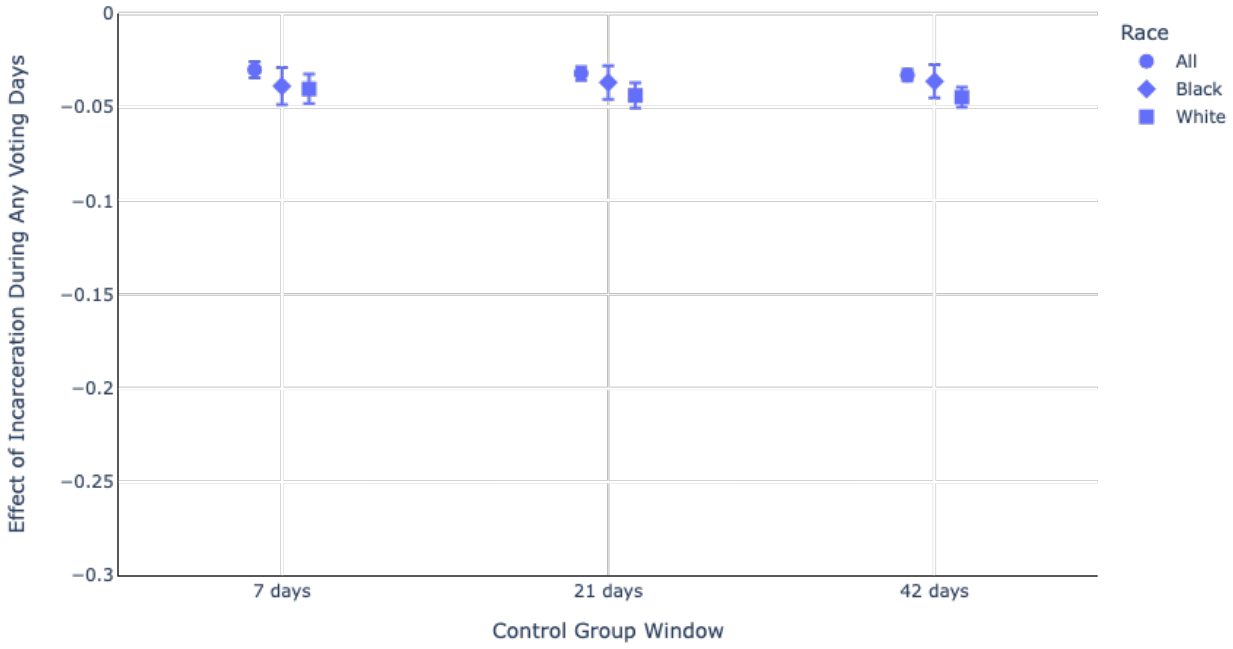

Figure A.1: The Effect of Incarceration During Any Voting Days

Note: This figure reports OLS point estimates and 95% confidence intervals for the effect of being booked into a county jail during 2020 voting days for any length of time on a registered voter's probability of voting in the 2020 election, with all covariates included, for samples with match probability $p > 0.75$, as described in the notes to Table 1. All models include jail and week-of-year fixed effects; robust standard errors two-way clustered at the jail and week-of-year level are reported in parentheses.

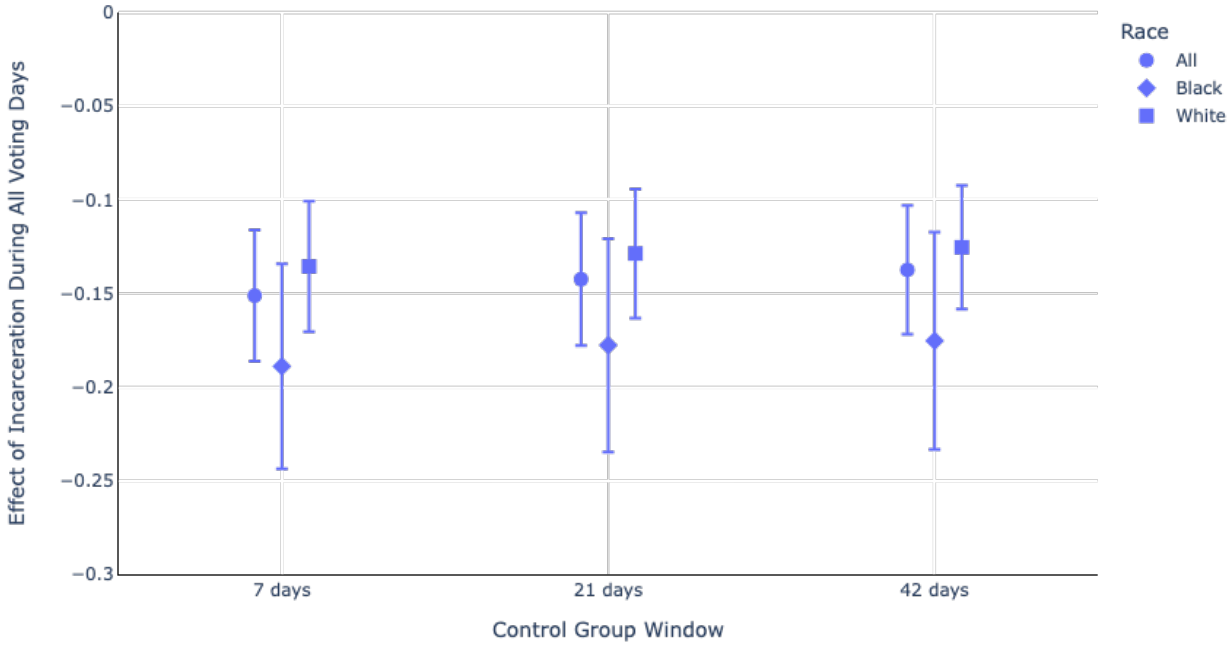

Figure A.2: The Effect of Incarceration During All Voting Days

Note: This figure reports OLS point estimates and 95% confidence intervals for the effect of being booked into a county jail during all 2020 voting days on a registered voter's probability of voting in the 2020 election, with all covariates included, for samples with match probability $p > 0.75$, as described in the notes to Table 1. All models include jail and week-of-year fixed effects; robust standard errors two-way clustered at the jail and week-of-year level are reported in parentheses.

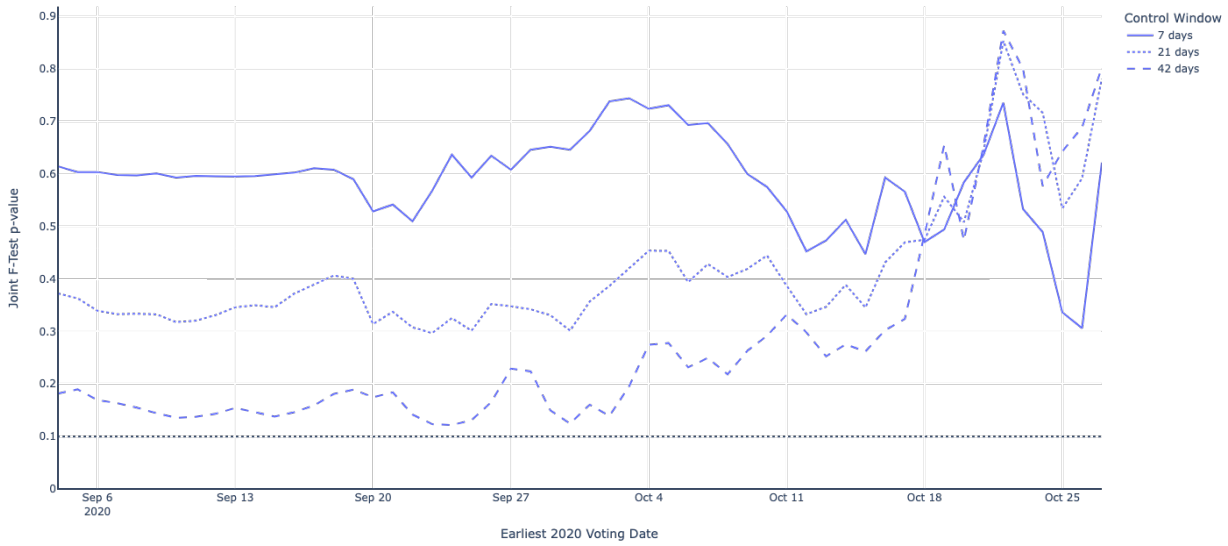

Figure A.3: Balance Tests (Joint F-test P-values)

Note: This figure reports estimates of balance on individual- and booking-level characteristics for post-Election Day control windows ranging between 7 and 42 days and treatment windows containing legal voting days in 2020, for the sample of matched records with match probability $p > 0.75$. For each post-Election Day control window we initially tested for balance using treatment windows containing the full set of voting days in each state in 2020, sequentially dropping the earliest voting day in the sample and repeating the balance test until we reached a treatment window of 7 days prior to Election Day.

Table A.1: The Effect of Incarceration on Voting From Jail
Including Match Probability as Covariate

Control Group Window:	**7 days**	**21 days**	**42 days**
Confined During Voting Days	-0.029***	-0.031***	-0.032***
	(0.002)	(0.002)	(0.002)
Match Probability	-1.429***	-1.391***	-1.351***
	(0.041)	(0.043)	(0.043)
Proportion of Voting Days Confined	-0.168***	-0.157***	-0.151***
	(0.018)	(0.019)	(0.018)
Match Probability	-1.449***	-1.407***	-1.362***
	(0.044)	(0.046)	(0.045)
Mean Proportion of Voting Days Confined	0.215	0.215	0.215
Mean Control Turnout	0.349	0.354	0.357
Observations	63000	79854	103091

Note: This table reports OLS estimates of (1) the effect of being booked into a county jail during 2020 voting days for any length of time on a registered voter's probability of voting in the 2020 election, with covariates; and (2) the effect of the proportion of 2020 voting days during which a registered voter booked into a county jail during voting days was confined on their probability of voting in the 2020 election, with covariates, for the sample of matched records with match probability $p > 0.75$, including match probability as a covariate. Models are estimated on samples as described in the notes to Table 1. All models include jail and week-of-year fixed effects; robust standard errors two-way clustered at the jail and week-of-year level are reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.10.

Table A.2: Racial Disparities in the Effect of Incarceration on Voting From Jail Including Reported and Predicted Race

Control Group Window:	**7 days**	**21 days**	**42 days**
Confined During Voting Days	-0.039***	-0.042***	-0.042***
	(0.004)	(0.005)	(0.004)
Confined × Black	-0.003	-0.001	-0.003
	(0.006)	(0.008)	(0.006)
Proportion of Voting Days Confined	-0.135***	-0.128***	-0.125***
	(0.017)	(0.017)	(0.016)
Proportion Confined × Black	-0.060**	-0.055**	-0.054**
	(0.024)	(0.023)	(0.022)
Mean Proportion of Voting Days Confined Black	0.209	0.209	0.209
Mean Proportion of Voting Days Confined White	0.218	0.218	0.218
Mean Control Turnout Black	0.278	0.285	0.292
Mean Control Turnout White	0.377	0.380	0.383
Observations	54741	69084	89010

Note: This table reports OLS estimates of (1) the effect of being booked into a county jail during 2020 voting days for any length of time on a registered voter's probability of voting in the 2020 election, interacted with an indicator for whether the individual is Black (0 = non-Hispanic white), with covariates; and (2) the effect of the proportion of 2020 voting days during which a registered voter booked into a county jail during voting days was confined on their probability of voting in the 2020 election, interacted with an indicator for whether the individual is Black, with covariates. Samples include states for which L2 predicts race and ethnicity categories. Models are estimated on samples as described in the notes to Table 1, including only Black and non-Hispanic white individuals. All models include jail and week-of-year fixed effects; robust standard errors two-way clustered at the jail and week-of-year level are reported in parentheses. ***p $<$ 0.01, **p $<$ 0.05, *p $<$ 0.10.

B Appendix B: Registered Voters; Match Probability $p > 0.95$

Table B.1: Summary Statistics
Match Probability $p > 0.95$

Treatment/Control:	T	C	C	C
Control Group Days:		7d	21d	42d
Age (L2)	36.65	36.76	36.77	36.70
Black (L2)	0.289	0.257	0.260	0.267
White (L2)	0.587	0.594	0.592	0.589
Male (L2)	0.716	0.717	0.715	0.717
Democrat (L2)	0.422	0.424	0.414	0.415
Republican (L2)	0.182	0.186	0.187	0.185
Non-Partisan or Other Party (L2)	0.396	0.391	0.400	0.401
Number of Charges (JDI)	2.489	2.553	2.505	2.501
Top Charge: Drug (JDI)	0.135	0.129	0.129	0.127
Top Charge: DUI (JDI)	0.065	0.070	0.069	0.068
Top Charge: Criminal Traffic (JDI)	0.033	0.036	0.032	0.032
Top Charge: Property (JDI)	0.190	0.193	0.193	0.192
Top Charge: Public Order (JDI)	0.292	0.286	0.282	0.277
Top Charge: Violent (JDI)	0.284	0.286	0.295	0.303
Length of Stay (days) (JDI)	35.45	35.35	33.65	34.12
Confined During Voting Days	1.000	0.000	0.000	0.000
Proportion of Voting Days Confined	0.206	0.000	0.000	0.000
Turnout	0.269	0.300	0.308	0.312
Observations	40962	6899	19872	37749

Note: This table reports mean values of individual- and booking-level characteristics for records with match probability $p > 0.95$, by treatment and control groups.

Table B.2: The Effect of Incarceration on Voting From Jail
Match Probability $p > 0.95$

Control Group Windows:	**7 days**	**21 days**	**42 days**
Confined During Voting Days	-0.040***	-0.042***	-0.042***
	(0.002)	(0.002)	(0.001)
with covariates	-0.041***	-0.043***	-0.044***
	(0.002)	(0.002)	(0.002)
Proportion of Voting Days Confined	-0.261***	-0.256***	-0.251***
	(0.023)	(0.023)	(0.022)
with covariates	-0.221***	-0.209***	-0.200***
	(0.020)	(0.020)	(0.020)
Mean Proportion of Voting Days Confined	0.206	0.206	0.206
Mean Control Turnout	0.300	0.308	0.312
Observations	47861	60834	78711

Note: This table reports OLS estimates of (1) the effect of being booked into a county jail during 2020 voting days for any length of time on a registered voter's probability of voting in the 2020 election, with and without covariates; and (2) the effect of the proportion of 2020 voting days during which a registered voter booked into a county jail during voting days was confined on their probability of voting in the 2020 election, with and without covariates, for the sample of matched records with match probability $p > 0.95$. All models include jail and week-of-year fixed effects; robust standard errors two-way clustered at the jail and week-of-year level are reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.10.

Table B.3: Placebo Tests
Match Probability $p > 0.95$

	2016 Placebo		**2012 Placebo**	
Control Group Window: 7 days	2016	2020	2012	2020
Confined During Voting Days	0.009	-0.025***	-0.009	-0.027***
	(0.009)	(0.008)	(0.010)	(0.010)
Mean Control Turnout	0.403	0.342	0.526	0.389
Observations	30014	30014	20415	20415
Control Group Window: 21 days	2016	2020	2012	2020
Confined During Voting Days	0.008	-0.035***	-0.003	-0.041***
	(0.005)	(0.005)	(0.007)	(0.007)
Mean Control Turnout	0.404	0.355	0.515	0.403
Observations	38117	38117	25914	25914
Control Group Window: 42 days	2016	2020	2012	2020
Confined During Voting Days	0.003	-0.037***	-0.003	-0.041***
	(0.004)	(0.004)	(0.005)	(0.005)
Mean Control Turnout	0.409	0.359	0.516	0.406
Observations	49243	49243	33339	33339

Note: This table reports placebo tests of the effects of 2020 jail incarceration on 2016 and 2012 voter turnout, for the sample of matched records with match probability $p > 0.95$. The 2016 (2012) placebo samples include registered voters who were registered to vote on both Election Day 2020 and Election Day 2016 (Election Day 2012). The 2016 (2012) placebo tests report OLS estimates of the effect of a 2020 booking during 2020 voting days on a registered voter's probability of voting in the 2016 (2012) election. All models include jail fixed effects; robust standard errors clustered at the jail level are reported in parentheses. ***p < 0.01, **p < 0.05 , *p < 0.10.

C Appendix C: Registered and Unregistered Voters

Table C.1: Summary Statistics
Registered and Unregistered Voters

Treatment/Control:	T	C	C	C
Control Group Days:		7d	21d	42d
Age (JDI)	35.70	35.76	35.71	35.65
Black (JDI)	0.336	0.326	0.329	0.333
White (JDI)	0.597	0.594	0.591	0.588
Male (JDI)	0.769	0.776	0.773	0.774
Number of Charges (JDI)	2.750	2.766	2.764	2.885
Top Charge: Drug (JDI)	0.152	0.145	0.145	0.143
Top Charge: DUI (JDI)	0.044	0.045	0.046	0.045
Top Charge: Criminal Traffic (JDI)	0.029	0.032	0.030	0.030
Top Charge: Property (JDI)	0.202	0.206	0.207	0.205
Top Charge: Public Order (JDI)	0.323	0.317	0.314	0.310
Top Charge: Violent (JDI)	0.250	0.255	0.258	0.267
Length of Stay (days) (JDI)	45.06	44.44	42.06	42.16
Confined During Voting Days	1.000	0.000	0.000	0.000
Proportion of Voting Days Confined	0.235	0.000	0.000	0.000
Registered by Election Day	0.290	0.290	0.294	0.293
Voted	0.074	0.081	0.085	0.086
Observations	106836	17477	50229	95311

Note: This table reports mean values of individual- and booking-level characteristics for both registered and unregistered voters, defining a registered voter as a jailed individual with match probability $p > 0.95$, by treatment and control groups.

Table C.2: The Effect of Incarceration on Registering From Jail

Control Group Window:	**7 days**	**21 days**	**42 days**
Confined During Voting Days	0.011***	0.010***	0.010***
	(0.001)	(0.001)	(0.001)
with covariates	0.007***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)
Proportion of Voting Days Confined	-0.164***	-0.160***	-0.158***
	(0.013)	(0.013)	(0.012)
with covariates	-0.110***	-0.100***	-0.093***
	(0.011)	(0.011)	(0.010)
Mean Proportion of Voting Days Confined	0.235	0.235	0.235
Mean Control Registration Rate	0.290	0.294	0.293
Observations	124313	157065	202147

Note: This table reports OLS estimates of (1) the effect of being booked into a county jail during 2020 voting days for any length of time on an individual's probability of being registered to vote, with and without covariates; and (2) the effect of the proportion of 2020 voting days during which an individual booked into a county jail during voting days was confined on their probability of being registered to vote, with and without covariates, where registration is defined as matching to L2 records with match probability $p > 0.95$. All models include jail and week-of-year fixed effects; robust standard errors two-way clustered at the jail and week-of-year level are reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.10.

Table C.3: The Effect of Incarceration on Voting From Jail
Unconditional on Registration Status

Control Group Window:	**7 days**	**21 days**	**42 days**
Confined During Voting Days	-0.005***	-0.005***	-0.005***
	(0.000)	(0.000)	(0.000)
with covariates	-0.006***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)
Proportion of Voting Days Confined	-0.112***	-0.109***	-0.108***
	(0.011)	(0.011)	(0.010)
with covariates	-0.090***	-0.083***	-0.079***
	(0.010)	(0.010)	(0.010)
Mean Proportion of Voting Days Confined	0.235	0.235	0.235
Mean Control Turnout	0.081	0.085	0.086
Observations	124313	157065	202147

Note: This table reports OLS estimates of (1) the effect of being booked into a county jail during 2020 voting days for any length of time on an individual's probability of voting in the 2020 election, with and without covariates; and (2) the effect of the proportion of 2020 voting days during which an individual booked into a county jail during voting days was incarcerated on their probability of voting in the 2020 election, with and without covariates, unconditional on registration status. All models include jail and week-of-year fixed effects; robust standard errors two-way clustered at the jail and week-of-year level are reported in parentheses. ***p $<$ 0.01, **p $<$ 0.05, *p $<$ 0.10.

D Appendix D: Data and Matching

D.1 Jail Data Initiative Booking Records

Our data on jail bookings are sourced from the Jail Data Initiative (JDI) NoSQL database of booking records.[9] We initially sampled all individuals in the JDI database with jail booking dates from August 5, 2020 to February 1, 2021. These dates span 90 days on each side of 2020 Election Day (November 3, 2020) and include all legal voting days in 2020 (the earliest of which was September 4, 2020 in North Carolina, for absentee mail-in voting, 60 days prior to Election Day), and a post-Election Day period encompassing all control windows we employ. This sample includes 944,985 bookings. For each of these bookings we retrieved individual- and booking-level standardized fields from the JDI database as detailed below.

Name Jail rosters may report a single name string, and/or first name, middle name, middle initial, and/or surname. In order to compare name components as part of the match process, we split names into components using the TupiLabs Java HumanNameParser package.[10] Additionally, we create soundex codes for surnames, using the Apache Soundex Java class.[11]

Length of Stay Length of stay is used as a booking-level covariate in analyses of the effects of jail incarceration. Among the initial sample of 944,985 bookings, length of stay ranges from 1 to 736 days, and the mean length of stay is 44.48 days.

Age Jail rosters may report age, year of birth, and/or date of birth. The JDI booking data report a standardized age field constructed from these original fields. Among the initial sample of 944,985 bookings, age ranges from 13 to 95, and the mean age is 35.7. The proportion without reported age is 0.19. Age is used as an individual-level covariate in analyses that include both unregistered and registered voters. We exclude from those samples booking records without reported age.

Gender Jail rosters may report sex and/or gender. The JDI database reports a standardized sex-gender field from these original fields. Values (e.g., "F", "Female", "FEM", etc.) are manually encoded as male, female, trans, non-binary or unknown. Among the initial sample of 944,985

[9] For detailed documentation of the JDI pipeline and database architecture, see https://jaildatainitiative.org/documentation.

[10] https://github.com/tupilabs/HumanNameParser.java.

[11] https://commons.apache.org/proper/commons-codec/apidocs/org/apache/commons/codec/language/Soundex.html.

bookings, the proportion without reported or encoded sex-gender is 0.24. Among those with encoded sex-gender, the proportion male is 0.78 and the proportion female is 0.23. 17 records (0.00) are encoded as non-binary or trans; these are re-categorized as unknown for matching with L2 voter records. Encoded sex-gender is used as an individual-level covariate in analyses that include both unregistered and registered voters. We exclude from those samples booking records without encoded sex-gender.

Number of Charges Jail rosters may or may not report charges for each booking; charges may appear as a single string in an unparseable format. Among the initial sample of 944,985 bookings, the proportion with no determinable number of charges is 0.10. Among those with parseable charges, the mean number of charges per booking is 2.6. Number of charges is used as a booking-level covariate in analyses of the effects of jail incarceration. We exclude from all samples booking records without parseable charge records.

Top Charge For each booking with parseable charge description(s) as specified above, the JDI database reports the highest-probability-match Uniform Crime Classification Standard (UCCS) code from the Text-Based Offense Classification (TOC) algorithm developed by the Criminal Justice Administrative Records System at the University of Michigan.[12] In some cases, the charge string is not sufficiently legible for the model to produce a match classification. The JDI database assigns a "top" charge type for the booking from among the least granular classification codes ("offense type" codes, with values 1-6 corresponding to the charge categories outlined below), using a hierarchical ordering of classification codes by severity. This order of severity, from most to least serious, is: violent, property, drug, public order, DUI, criminal traffic. Among the initial sample of 944,985 bookings, the proportion with no determinable top charge type is 0.2. Among those with encoded top charge types, the proportions of each type are: violent (0.21), property (0.17), drug (0.12), public order (0.25), DUI (0.04), and criminal traffic (0.03). Top charge indicators are used as booking-level covariates in analyses of the effects of jail incarceration. We exclude from all samples booking records without parseable charge records.

County FIPS Code The JDI database reports the unique five-digit Federal Information Processing Standard Publication 6-4 (FIPS) county-identifying code from the Bureau of Justice

[12]For detailed information about the CJARS TOC model, see https://cjars-toc.isr.umich.edu/.

Statistics 2013 Census of Jails corresponding to the county in which a jail is located.[13] Among the initial sample of 944,985 bookings, there are 985 unique county FIPS codes.

D.2 L2 Voter Records

We sourced voter records from L2 voter files. The records used here were pulled in Spring 2021. The voter records from which we derive match candidates include 195,655,326 registered voters in the 42 states for which we have jail booking records.

Age Among the initial sample of 195,655,326 voters, age ranges from 18 to 100, and the mean age is 50.07. The proportion without reported age is 0.02. L2 age is used as an individual-level covariate in analyses of the effects of jail incarceration on the voting behavior of registered voters. We exclude from those samples records without L2 age.

Gender Among the initial sample of 195,655,326 voters, the proportion without reported gender is 0.01. Among those with reported gender, the proportion male is 0.47 and the proportion female is 0.53. L2 gender is used as an individual-level covariate in analyses of the effects of jail incarceration on the voting behavior of registered voters. We exclude from those samples records without L2 gender.

Race Among the initial sample of 195,655,326 voters, the broadest L2 race/ethnicity field may take one of the following values: "East and South Asian," "European," "Hispanic and Portuguese," "Likely African-American," or "Other." In states where these fields are not reported, L2 uses an algorithmic process to predict race/ethnicity, relying on a proprietary database of name-ethnicity mappings, with additional census block and secondary in-block surname assessment. We re-categorize L2 race/ethnicity categories into one of three possible values: (non-Hispanic) "white" if L2 ethnicity value is "European," (non-Hispanic) "Black" if L2 ethnicity value is "Likely African-American," and "Other" if L2 ethnicity value is "East and South Asian," "Hispanic and Portuguese" or "Other." Among our sample of voters, the proportion with unknown/un-predicted race is 0.09. Among those with known/predicted race, the proportion white is 0.68, the proportion Black is 0.12, and the proportion Other is 0.20. L2 race/ethnicity is used as an individual-level covariate in analyses of the effects of jail incarceration on the voting behavior of registered voters. We exclude

[13] https://nvlpubs.nist.gov/nistpubs/Legacy/FIPS/fipspub6-4.pdf, https://doi.org/10.3886/ICPSR36128.v4.

from those samples records without L2 race/ethnicity.

Party L2 reports voter party registration data where available. Where party registration data are not available, L2 models "likely" party registration (https://www.l2datamapping.com/help?ds=VM_US#data__party). The party registration field may take one of 57 values or "unknown." We re-categorize the party registration field into one of three possible values: "Democratic," "Republican", and "Non-Partisan or Other". Among our sample of voters, the proportion with unknown party registration is 0.00. Among those with known/predicted party, the proportion Democratic is 0.41, the proportion Republican is 0.32, and the proportion Non-Partisan or Other is 0.28. L2 party registration is used as an individual-level covariate in analyses of the effects of jail incarceration on the voting behavior of registered voters. We exclude from those samples records without L2 party registration.

Registration Date Voter registration date is used for sample inclusion in turnout analysis. L2 reports a calculated registration date field that reduces missingness in reported registration date, and captures the earliest registration date for an individual, since official registration dates are often reset when individuals move between voting districts. Among the initial sample of 195,655,326 voters, the mean calculated voter registration date is March 6, 1990.

Voting Indicators L2 reports binary fields that indicate whether registered voters voted. We record such fields for the 2020, 2016 and 2012 general elections. Among the initial sample of 195,464,995 registered voters, the mean registered voter turnout for the 2020 general election was 0.75, the mean registered voter turnout for the 2016 general election was 0.59, and the mean registered voter turnout for the 2012 general election was 0.50.

County FIPS Code Among the initial sample of 195,655,326 voters there are 3,004 unique county FIPS codes.

D.3 Probabilistic Record Linkage

Using the pools of booking and voter records described above, we conduct probabilistic record linkage using the Fellegi-Sunter/Expectation Maximization linkage model enumerated by McDonough, Enamorado and Mendelberg (2022) and Enamorado, Fifield and Imai (2019).

Match-candidate booking rows appear as follows (note: full name and soundex code are used for blocking and initialization, but not for matching):

Booking ID	FIPS	Age	Gender	First	Middle	Last	Full Name	Soundex
12a3b...	11217	30		Jane	A	Doe	Jane A Doe	D000

Match-candidate voter rows appear as follows:

Voter ID	FIPS	Age	Gender	First	Middle	Last	Full Name	Soundex
LALAL1...	11217	29	M	John	Adam	Doe	John Adam Doe	D000

We create blocks for matching as below. Importantly, we only run our match program within-state; if an individual is booked in Alabama but is registered in the Wyoming voter file, we do not consider them for matching.

1. For each unique age value among match-candidate bookings, we select only match-candidate registered voters such that the absolute value of the difference (L1 norm) of their age and the current age is <= 2 years (a blocked age range of 5 years). The intuition for this age bounding leniency is that jail rosters may not consistently update recorded age for booked individuals over time, and the scope of the JDI database is 2 years. For bookings with unknown age, we do not block on voter age.

2. For each match-candidate booking with the current age value, we select only match-candidate registered voters with equivalent surname soundex codes.

Our sampling process runs over a random selection of within-state match-candidate pairs. We compute a binary match indicator field for parametric initialization by calculating the average of the Jaro-Winkler similarities of first name, last name, and full name, considering matches to be any rows with resultant average > 0.88. We also output numeric similarity vectors for each match-candidate pair. These are defined by the following logic:

1. FIPS: 1 if FIPS values are equal; else 0.

2. Age: 2 if the absolute value of the difference of age between booking and registered voter is <= 1; 1 if either age is missing; else 0.

3. Gender: 2 if booking and registered voter gender values are equal; 1 if either gender is missing; else 0.

4. First Name: 2 if Jaro-Winkler similarity of booking and registered voter first name is > 0.94; 1 if Jaro-Winkler similarity of booking and registered voter first name is > 0.88 and $<= 0.94$; else 0.

5. Middle Name: 2 if (a) one value is an initial and the other is a complete middle name, and the initial equals the first letter of the complete middle name or (b) both values are complete middle names, and their Jaro-Winkler similarity is > 0.94; 1 if (a) either middle name is missing or (b) both values are complete middle names, and their Jaro-Winkler similarity is > 0.88 and $<= 0.94$; else 0.

6. Last Name: 2 if Jaro-Winkler similarity of booking and registered voter last name is > 0.94; 1 if Jaro-Winkler similarity of booking and registered voter last name is > 0.88 and $<= 0.94$; else 0.

For the synthetic booking and registered voter row examples above, sampling output would appear as follows:

Booking ID	Registered Voter ID	FIPS	Age	Gender	First	Middle	Last	Average	Match
12a3b...	LALAL1...	1	2	1	0	2	2	0.83	0

Sampling results are used to initialize parameters for Fellegi-Sunter calculation. Subsequently, iterative expectation maximization is conducted to stabilize parameter estimates from 50 re-samples of 1 million match-candidate pairs each. Following the full linkage process, we retain only the registered voter match(es) for each booking with the highest probability score(s), and discard all matches with probability scores < 0.5. We then conduct in-block/in-state term frequency re-weighting of first and last names on match records as described by Enamorado, Fifield and Imai (2019) and Winkler (2000). The final steps of our matching process are exclusions using the following logic:

1. We exclude any rows where re-weighted match probability is < 0.75.

2. We exclude all rows for any booking that is matched to more than one unique voter ID.

3. We exclude all rows for any registered voter ID that is matched to overlapping bookings (where booking start and end dates are determined as described in Section 3.1 of this report).

4. We exclude all rows where booking-reported age is < 18.

5. We exclude all rows where voter registration date is after Election Day (November 3, 2020).

In total, out of our initial sample of 944,985 jail bookings we match to 331,703 registered voter records. As a final step to enable measurement of covariates related to charges against booked individuals, we subset this pool of matches to include only the 305,239 individuals detained in facilities whose rosters report charge information.