Supplemental Information

I here present more information on how the full data set used in the body of this paper was constructed.

Geocoding

Voters' home addresses were converted to latitudes and longitudes using a geocoder provided by SmartyStreets. I then used the statistical software R to map these latitudes and longitudes to census block groups, census tracts, and city council districts using shapefiles publicly available from the Census Bureau and the City of New York. This geocoder is not perfect: among individuals registered to vote in New York City, the geocoder failed to determine the latitude and longitude of the addresses of 1% of registered voters. The geocoder was slightly less successful when it came to lost voters; 1.6% of these individuals were not geocoded. Voters who were not successfully geocoded are dropped from the dataset; however, because so few observations went uncoded, it is unlikely to affect the analysis.

Matching

I construct the primary dataset by matching the administrative criminal supervision records with the registered voter file. I match using first, middle, and last names, as well as dates of birth. Capitalization in names in both datasets is standardized, and punctuation and spaces are removed. If a middle name is missing in one dataset but not in the other, I consider this a match. Similarly, if one dataset includes only a middle initial which matches to the first letter of the middle name the other data set, this is considered a match. First and last names, and dates of birth, must match exactly between the two datasets.

To test for false positive matches, I employ the test developed in Meredith and Morse (2013). I slightly alter the date of birth reported in the NYSDOCCS dataset to create false records. Comparing the number of matches between these "fake" records and the voter file with the number of matches between the "true" records and the voter file provides an estimate of how frequently false positives occur. Table 5 shows the results of true matches, as well as matches using a set of fake records created by adding or subtracting 35 days from an individual's birthdate. This analysis indicates that false positives account for between 0.4 and 0.5% of all matches, a share that is likely too small to have any material impact on the overall analysis. The numbers in Table 5 are derived by matching (and modifying) all individuals who were incarcerated or on parole on Election Day in 2017 with the registered voter file from April of 2018.

Table 5. Results of Shifting Birthdates.

Group	Number of Matches Between DOCCS and Voter File Records
Actual birthdate	20,955
Birthdate + 35 days	105
Birthdate – 35 days	92

Note. DOCCS = Department of Corrections and Community Supervision.

Testing for false negatives is more challenging. If an individual marries and changes her name after being discharged from parole, for instance, I will not identify her using my matching methodology. For two reasons, however, false positives are unlikely to pose major challenges: firstly, women are far more likely to change their last names than men, and women make up barely 6% of individuals who have been discharged from felony parole. Secondly, because both parolee discharge and voter registration are legal records, individuals are likely to be recorded using their full names (that is to say, an individual is unlikely to be "Alex" in one set of records and "Alexander" in the other).

References

Meredith, Marc, and Michael Morse. 2013. "Do Voting Rights Notification Laws Increase Ex-Felon Turnout?" *The ANNALS of the American Academy of Political and Social Science* 651 (1): 220–49. doi:10.1177/0002716213502931