Early Voting Laws, Voter Turnout, and Partisan Vote Composition: Evidence from Ohio[†]

By Ethan Kaplan and Haishan Yuan*

We estimate effects of early voting on voter turnout using a 2010 homogenization law from Ohio that forced some counties to expand and others to contract early voting. Using voter registration data, we compare individuals who live within the same 2 × 2 mile square block but in different counties. We find substantial positive impacts of early voting on turnout equal to 0.22 percentage points of additional turnout per additional early voting day. We also find greater impacts on women, Democrats, independents, and those of child-bearing and working age. We simulate impacts of national early day laws on recent election outcomes. (*JEL* D72, K16)

In recent years, political parties have been very active in passing legislation at the state level expanding or limiting ease of access to voting. State-level legislative activity regulating voting has been primarily concentrated in four areas: legal changes affecting the ease of voter registration, laws expanding or contracting the ability of felons to vote, laws that tighten or loosen identification requirements at the ballot box, and laws expanding or contracting the prevalence of early voting availability.

Early voting in particular and preelection voting in general have become common forms of voting. Though preelection voting began first in California back in 1976 and then in Texas in 1987, most of the rollout of early voting happened in the 1990s and 2000s (Biggers and Hanmer 2015). As of 1992, 7 percent of individuals cast their ballots using some form of preelection voting (McDonald 2016). By 2008, the beginning of our main sample, preelection voting had expanded to over 30 percent of ballots cast nationally; these numbers rose to 34.5 percent by 2016. Initially, preelection voting was primarily in the form of mail balloting. However, in recent years, the

importance of early voting[1] has risen. In 2016, over 47 million of the approximately 136.5 million ballots cast used some form of preelection voting; 23 million of these were cast in person. Substantial differences have emerged across states in early voting availability. On the one hand, 13 states currently have no in-person early voting.[2] At the other extreme, Minnesota provided 46 days of early voting. Early voting is potentially important in the United States because Election Day is neither a national holiday nor a weekend day as it is in many developed countries.

Of course, it is not clear that expanding opportunities to vote will actually increase voting. Some political scientists who have studied early voting have estimated positive effects on turnout (Glynn and Kashin 2017, Herron and Smith 2012, 2014); others have found no systematic overall impact upon turnout (Gronke, Galanes-Rosenbaum, and Miller 2007), and yet others have found that early voting expansion has reduced turnout (Burden et al. 2014). The idea that early voting may reduce turnout may sound strange at first. However, there is a well-documented effect that people vote in part to tell others (DellaVigna et al. 2017). It is also possible that voters turn out in order to be seen voting and that early voting, by spreading voting across many weeks, reduces the link between being seen and voting. Burden et al. (2014) have a similar explanation for their seemingly perverse findings. They claim that early voting weakens a sense of common solidarity, which is important for motivating high turnout.

Unfortunately, given the importance of the subject, there are surprisingly few studies on the impacts of differences in state voting laws in general and early voting in particular. Moreover, what studies exist suffer from plausible endogeneity bias. So far, there have been three main approaches to estimating early voting impacts: time series estimates using individual-level administrative data (Gronke, Galanes-Rosenbaum, and Miller 2007; Gronke and Miller 2012; Herron and Smith 2014), differenced estimates between treated and control individuals (Glynn and Kashin 2017), and cross-state differences in differences (Burden et al. 2014).

The time series estimates assume that changes in individual turnout are solely due to the effect of changes in early voting. This reduces endogeneity arising from characteristics of the electorate but does not control well for endogeneity due to characteristics of the election or differential reactions to the characteristics of the election across demographic groups. Burden et al. (2014) uses county-level voting data as well as individual-level data from the CPS voter supplement to run cross-state difference-in-differences regressions with county-level (or individual) demographic as well as state-level controls. The identifying variation comes from state-level legal changes in the availability in early voting. However, these are confounded by state-level and demographic voting trends including election characteristics. Glynn and Kashin (2017) use voter registration data from Florida and difference 2008 turnout rates between two groups: 2006 early voters who did not

---

[1] Voting early in person.

[2] The 13 states without in-person early voting are Alabama, Connecticut, Delaware, Kentucky, Michigan, Mississippi, Missouri, New Hampshire, New York, Pennsylvania, Rhode Island, South Carolina, and Virginia. Note that we are *not* counting Colorado, Oregon, and Washington State, which are "all mail balloting" states where individuals are mailed ballots weeks before the elections and then may either mail in their ballots or drop them off in person before the election at polling stations.

vote early in 2008 and absentee voters in 2006. However, both groups had access to early voting in 2008 and voting trends plausibly differ between prior early voters and absentee voters.

In this paper, we estimate the impact of early voting on voter turnout. We do this using voter registration data from Ohio and look at turnout before versus after Ohio homogenized early voting availability across counties. Our paper, rather than trying to add in covariates to control for unobservables, tries to construct treatment and control groups that are similar. We do this using geographical discontinuities in treatment across county borders. We thus follow the literature using spatial neighbors with differential spatial treatment (Dube, Lester, and Reich 2010; Snyder and Strömberg 2010; Spenkuch and Toniatti 2018). We also add to a growing literature within economics estimating the impact of electoral interventions using more credible research designs (Braconnier, Dormagen, and Pons 2017; Naidu 2012; Pons 2018). Since we use individual level data from the Ohio voter registration database, even with our limited geographical discontinuity sample, we have tight standard errors. However, our spatial discontinuity approach also allows for credible identification.

Our best specification estimates effects within 2×2 mile blocks that straddle county lines where counties differentially changed early voting availability due to the change in state law. Besides looking at aggregate turnout effects, we estimate differential turnout effects for weekend days, same day registration days, and days where polls were open until 7 PM or later. We also estimate models where we allow for nonlinearities in treatment. We not only show estimates by different types of treatment but also by different types of voters. We estimate the impacts differentially by sex, party, and age. Overall, we find that an extra day of early voting increases turnout by 0.22 percentage points.

Finally, we use our estimates of partisan effects to linearly simulate the impact of hypothetical national early voting election laws. We find no impacts on majority control of the House of Representatives in either 2012 or 2016. However, we find that a national law of 23 days of early voting (the current level in Ohio) would have led to Democratic control of the Senate in 2012 as well as the presidency in 2016. We moreover find that 46 days of national early voting (the current level in Minnesota) would additionally have led to Democratic control of the Senate in the 2016 election.

In Section I, we give an overview of the electoral law change we use in the state of Ohio. In Section II, we describe our data. In Section III, we present our methodology. In Section IV, we discuss our main estimates. In Section V, we show the results of simulations of national electoral law changes on election outcomes. Finally, in Section VI, we conclude.

I. Ohio Election Law Changes

Like many states, Ohio saw large expansions of early voting in the 2000s. In 2002, 6.8 percent of voters cast preelection absentee ballots. In 2005, Ohio passed legislation allowing for in-person early voting. By 2008, 29.7 percent of the electorate voted preelection (Kaltenthaler 2010). In the general election of 2012, the percent making use of in-person voting before the election was 10.6 percent and by

2016, it had risen to 11.8 percent.[3] The contraction in early voting availability in urban areas happened during a period of increased popularity of early voting.

The expansion in early voting was differential across counties. Urban, Democratic areas expanded early voting at a faster rate than rural, Republican ones. By 2008, rural Pickaway County was open for 109 hours of early voting, spanning a total of 11 days including only 1 weekend day, 2 days of same-day registration and no weekend days of same-day registration or Sunday voting days. By contrast, urban Franklin County, which contains the city of Columbus, was open for a total of 340 hours spread over 35 days including 7 days of same-day registration voting,[4] 10 weekend days of early voting including 2 same-day registration weekend voting days, and 5 Sundays.

In November 2010, Republican John Kasich defeated incumbent Democrat Ted Strickland for the governorship. In addition, the state senate remained majority Republican and the state House of Representatives switched majority control to the Republican Party. Under unified Republican control, the government passed State Bill 295, which homogenized early voting across counties. Each county's early voting station was required to be open the exact same hours on the exact same days as those of all other counties. This meant that Cuyahoga County with a population of 1.266 million in 2012 ended up with identical hours of early voting as rural Pickaway County with population 56,000. The law eliminated early voting for the three days prior to the election. Thus, early voting in the weekend before the election was eliminated from all counties. The total number of days was changed to 26 with 4 same-day registration days, no weekend days of same-day registration and 2 Saturdays though no Sundays.

Large pre-2012 discrepancy across counties within Ohio led to large differential changes due to the state policy changes implemented in 2012. On the one hand, Cuyahoga, Franklin, and Summit Counties all saw reductions of nine days of early voting. This reduction was largely due to reduced weekend voting. In each case, eight of the nine days were Saturdays or Sundays. Moreover, Cuyahoga's total hours were reduced by 56.5; Franklin's and Summit's each by 94 hours. By contrast, Wyandot and Pickaway both increased their weekend early voting by one day. Though Wyandot's total number of days of early voting availability did not increase, Pickaway's did by 15 days. Wyandot's total hours of early voting increased by 100 and Pickaway's by 137.

The contracting counties were quite different from the expanding ones in terms of political orientation. Cuyahoga, which contains Cleveland, is a large urban area with a 1.28 million population as of the 2010 census. It is 30.3 percent African American and had a 68.8 percent vote share for Obama in 2008. Franklin County, containing

---

[3]These numbers on the prevalence of in-person early voting were obtained from the Ohio secretary of state website: https://www.sos.state.oh.us/SOS/elections/Research/electResultsMain.aspx.

[4]All states except North Dakota require registration in order to vote. Ten states and Washington, DC, allow registration on Election Day. All other states require pre-registration. In Ohio, registration must occur 30 days or more before the election in order to participate in a national general election. In practice, this has meant 28 days before the election because 30 days before the election has been a Sunday and the next day has been Columbus Day. The state of Ohio always extends the deadline to the next business day if it falls on a weekend or a holiday. In years where early voting extended before the deadline, citizens could register to vote and vote at the same time in an early voting station. This is called same-day registration.

Columbus, Ohio, is 21.2 percent African American, has a population of 1.16 million and had a 60.1 percent Obama vote share in 2008. Summit County, containing Akron, Ohio, has a population of 540,000, is 13.2 percent African American and had a 56.7 percent Obama vote share in 2008. By contrast, Pickaway, which is a rural county without a major city, has a population of less than 60,000, is 3.7 percent African American and had an Obama vote share of 39.8 percent; similarly, Wyandot, another rural county, has a population slightly above 20,000, is 0.4 percent African American and had a 38.6 percent Obama vote share.

Of course, comparing the counties that contracted versus expanded early voting risks strong endogeneity bias due to the correlation of differences in demographics and thus voting trends with the magnitudes and signs of early voting changes. Our main strategy thus relies on finding locations with differential contractions and expansions but similar demographics and thus voting trends.

In Figure 1, we show the changes in early voting days between 2008 and 2012 by county. We see large reductions in early voting days in 2012 relative to 2008 for the large urban counties of Cuyahoga, Summit, and Franklin. We also show changes of early voting by six other different measures across counties between 2008 and 2012 in online Appendix Figure A.1. These measures of early voting access are the numbers of weekend days, days allowing same-day registration, days open late, weekdays, Saturdays, and Sundays. For most measures, urban counties experienced large reductions of early voting in 2012, while rural counties saw increases, no changes, or relatively small decreases.

II. Data

Our main data source is the voter registration database from the state of Ohio. We collected the data in November 2014. The database contains full name, exact date of birth, date of registration, individual voting history dating back to the year 2000, address of residence including county, precinct, and party for those who have participated in primaries.[5]

Ohio is an open primary state. Therefore, the data does not contain party registration but instead records the party of the primaries the voter participated in. We record an individual as a Republican if the most recent primary they participated in was a Republican Party primary, a Democrat if the most recent primary they participated in was a Democratic Party primary, and an independent for those who have never participated in a primary within the time span of our data. A total of 43.1 percent of registered individuals are listed as independent in our sample, 30.4 percent are listed as Democrats, and 26.5 percent as Republicans.[6]

---

[5] The registration data for all of Morgan County is missing from the files that we obtained from the secretary of state of Ohio. Morgan County is one of the smallest population counties in Ohio. It has a total of 14,904 residents out of state with 7.6 million registered voters. Thus, Morgan County voters comprise less than 0.1 percent of all Ohio voters. Morgan County neighbors five counties in total; no pairs containing Morgan County appear in our sample. However, since Morgan and all of its neighbors all retained their exact number of early voting days between 2008 and 2012, the omission of these pairs is not consequential to our estimation.

[6] In our data, which goes back to the year 2000 and covers eight national primaries, only 7.2 percent of registered voters voted in a Republican primary for one election and a Democratic primary for another election.
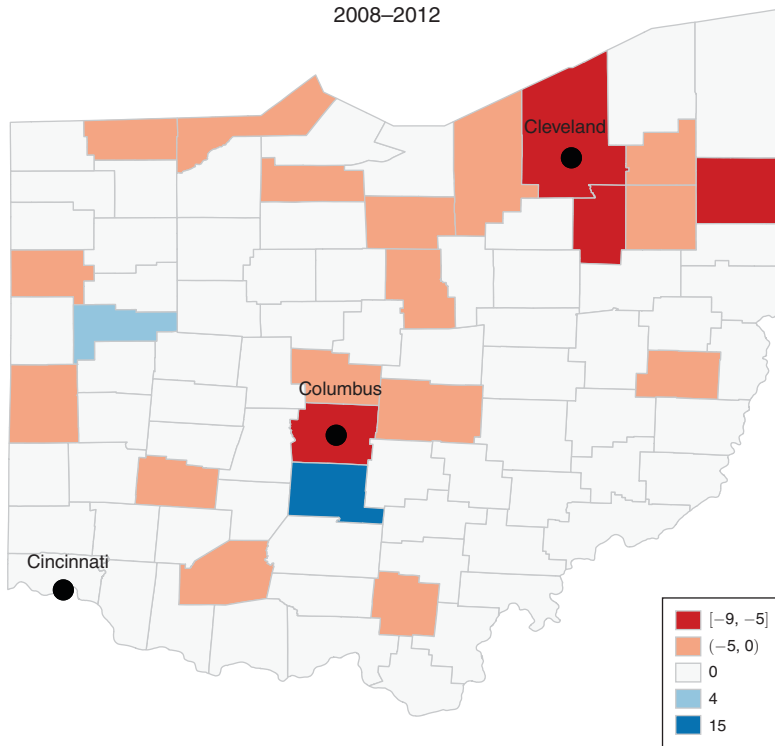
FIGURE 1. CHANGES IN EARLY VOTING DAYS

Using ArcGIS and Google Maps, we geocode each individual registration address into longitude and latitude. We then divide the state of Ohio into a mutually exclusive and exhaustive set of equal-sized square geographical blocks.

We additionally use the geocoded locations to merge demographic information on education, and income at the census block-group level and race at the census block level to each individual. Thus, each individual within a census block group has a set of demographic variables that do not vary across individuals within the same census block group.[7] These sets of variables include percent white, percent black, percent Hispanic, median household income, percent high school dropouts, and percent college graduates.

Next, for each of Ohio's 88 counties, we obtain from each individual county secretary of state the exact hours of early voting availability for each day of early voting. We do this for the years 2008, 2010, 2012, and 2014. We use this data to compute our main treatment variable: number of days of early voting by county for each election. We also compute other treatment variables, which we use for estimating heterogeneity in the treatment effect by type of treatment. These additional variables are number of hours, number of weekend days, number of Saturdays, number

---

[7]One exception being neighborhood racial composition, for which census block-level data is available.

of Sundays, number of weekdays, number of days of same-day registration, and number of days where polls were open until 7 PM or later.

We also compute for each individual the probability that they are female based upon first name in the voter registration file. Ohio voter registration data does not record sex. However, the Social Security Administration keeps a registry of all baby first names by sex. These lists are maintained by year. For confidentiality reasons, the data are truncated. Names with fewer than five occurrences in a given year for a given sex are not reported. As an example, in the year 1980, 94.8 percent of births in the United States are in our national baby name list. We obtained both the national lists as well as the lists for the state of Ohio. For each year and for each of the two lists (national and Ohio), we compute the probability that a name is female as the proportion of babies with that name who are female. If a name is not listed for a particular gender, we assume that zero babies were born with that name for that gender. We use the probability that a baby is female as our sex variable. We drop unmatched observations. A total of 95.6 percent of individuals in our voter registration file match to one of the first names in the national baby name file in their birth year; 89.9 percent match to one of the first names in the Ohio state baby name file in their birth year.

Finally, we use the self-reported ideology question and the party affiliation question from the 2016 Cooperative Congressional Election Study (CCES). The CCES is a stratified sample survey, administered by YouGov, which links questionnaire answers by respondents to actual voting records. We use 40,784 observations from the CCES to show ideological differences by party affiliation.

III. Methodology

We employ four empirical strategies to estimate the impact of restrictive voting laws upon voter turnout. The last of these is our preferred strategy. The first is the county-level difference-in-differences estimator where we regress voter turnout on treatment controlling for a county fixed effect and a time fixed effect. Our main treatment variable is the number of days of early voting. However, we also estimate models where we are interested in the heterogeneity of the treatment effect across different types of treatment (i.e., weekdays versus weekends, etc.). In these cases, we simultaneously regress upon multiple regressors. Our estimation equation is given by

$$(1) \qquad\qquad V_{ict} \; = \; \alpha_t + \phi_c + \mathbf{T}'_{ct}\boldsymbol{\beta} + \epsilon_{ct} + \theta_{ict},$$

where V_{ict} is a binary variable equal to 100 if voter i turns out in county c for the general election in time period t and zero otherwise, α_t is an election-year fixed effect, ϕ_c is a county level fixed effect, $\mathbf{T}_{ct}$ is a vector of treatment variables, ϵ_{ct} is a mean zero serially correlated county-specific random term which is independent across counties, and θ_{ict} is an idiosyncratic individual-level random term. We choose our dependent variable to take on the values of 100 or zero so that our estimates are expressed in units of percentage point effects per unit of treatment. We cluster standard errors for equation (1) at the county level. This specification assumes that

aggregate voting trends by county are uncorrelated with treatment. In particular, it assumes that trends in voter turnout in urban counties that saw large reductions in early voting would have been the same as in rural counties whose early voting access stayed constant or increased absent the early voting changes.

Our second main specification replaces the county-level fixed effects ϕ_c from equation (1) with individual fixed effects γ_i. Since there are no covariates in these regressions, the switch to individual fixed effects operates by dropping those who were not registered continuously over the time period. First-time registrants include those who were previously too young to register, those who were not too young but had never registered or voted, and those who moved to Ohio from out of state.[8] The individual fixed effects identification strategy relies upon weaker assumptions than the identification strategy assumed by the best related papers in the observational methods literature such as Card and Moretti (2007). Our individual fixed effects model, by contrast, correctly estimates treatment effects for those whose registration did not change across elections. However, this is still under the maintained assumption that voting trends for registered individuals across counties are uncorrelated with treatment. Our model of turnout, in this case, is given by

$$(2) \qquad V_{ict} \, = \, \alpha_t + \gamma_i + \mathbf{T}'_{ct}\boldsymbol{\beta} + \epsilon_{ct} + \theta_{ict}.$$

We next restrict our sample to individuals living within k miles of county borders, excluding borders that coincide with Ohio state borders. We refer to such samples as k-mile samples and re-estimate equation (2) using the k-mile sample with standard errors still clustered at the county level. We restrict the sample because our fourth and baseline estimation strategy requires restriction to individuals near county borders, and we separately estimate on that sample using equation (2) in order to isolate the impact of the geographical discontinuity design method. Our benchmark block size is 2×2 mile square blocks, though we also show estimation with block sizes ranging from 0.1×0.1 miles to 20×20 miles. Individuals living within two miles of counties' borders inside Ohio are marked by red dots in Figure 2.

Our final and preferred specification is a geographic discontinuity design. We divide up the state of Ohio into a mutually exclusive and exhaustive set of $k \times k$-mile square blocks (i.e., k^2 square mile blocks). Each individual then belongs to a unique block. We regress the change in turnout between 2008 and 2012 upon the change in early voting days, using the k-mile sample and controlling for geographical block fixed effects. We thus estimate

$$(3) \qquad \Delta V_{ibc} \, = \, \Delta \mathbf{T}'_c\boldsymbol{\beta} + \rho_b + \epsilon_c + \theta_{ic},$$

where ρ_b is a geographical block fixed effect.[9] Notice that the first differencing eliminates any individual fixed effect, and the geographical block fixed effect accounts

---

[8] The data is already purged of those who have passed away. If there is measurement error in reporting of death, it does not impact our estimation as long as the error is not systematically correlated with treatment.

[9] As a robustness check, we also estimate a model where instead of block fixed effects, we put in border segment fixed effects, which incorporate a following k-mile band on each side of each pair of counties following Dell (2010). The equation for this specification is the same as in equation (3) with b denoting border segments rather than blocks.
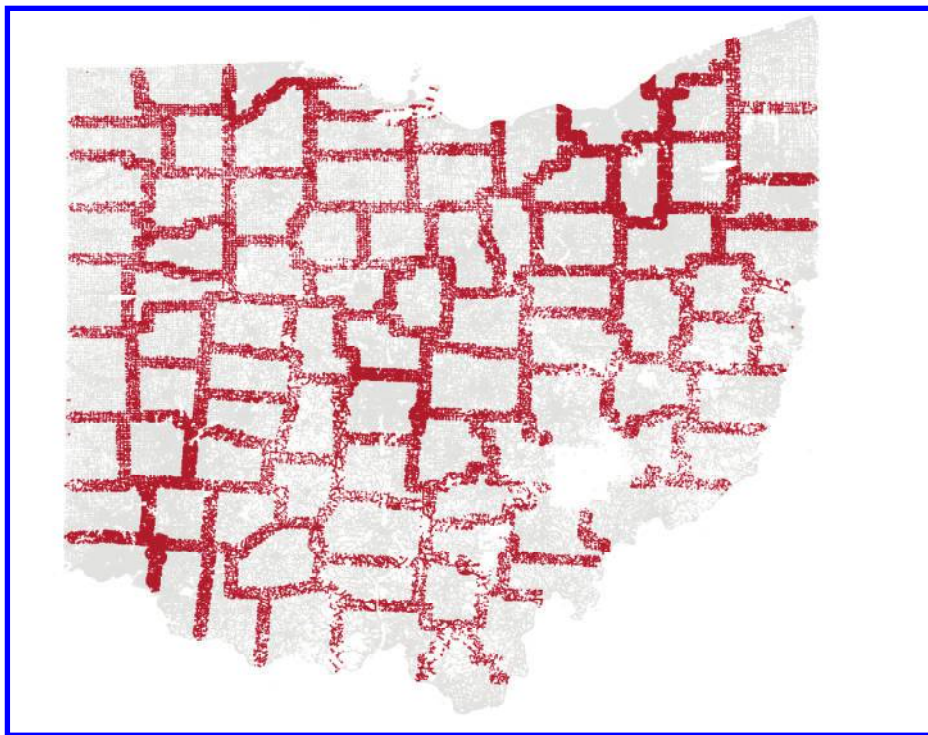
FIGURE 2. FULL SAMPLE OF OHIO REGISTERED VOTERS AND TWO-MILE COUNTY BORDER SAMPLE

for any year-specific local geographical/demographic effects that are constant within small areas across county lines. This specification is our most taxing and is thus the specification that requires the weakest identification assumption. Our maintained assumption under this identification strategy is that turnout trends for individuals are not correlated with change in treatment ΔT_c, within small geographical blocks.[10]

In addition to running regressions with voter turnout as our dependent variable, we also put placebo variables on the left-hand side. Placebo variables measured at the individual level include age, imputed gender, and party affiliation (Democrat, Republican, and independent). However, we also put in census aggregate variables, which come from matching individuals to census block groups (census block for racial composition). For variables measured at the individual level, we also estimate our geographical fixed effects model interacted with variables for subgroups of the population. We do this for Democrats, Republicans, and independents as well as

---

[10] This estimation strategy derives from the geographical discontinuity design literature, which initially arose in the context of the empirical literature on the minimum wage (Card and Krueger 1994; Dube, Lester, and Reich 2010). Here, instead of comparing counties within pairs that straddle state lines and have different minimum wage levels over time, we are comparing individuals within small geographical blocks who live in different counties with differential changes in the availability of early voting over time. Our estimation strategy would be analogous to the minimum wage literature if we put in block × county fixed effects instead of first differencing by individual. However, since we only have two data points per individual, first differencing our data by individual is identical to putting in individual fixed effects, and putting in individual fixed effects is a more stringent specification than putting in block × county fixed effects. The first differencing is computationally preferable to the fixed effects approach due to the large sample of individuals.

for the estimated probability of being female. In this case, we estimate interactive models given by

$$(4) \qquad \Delta V_{ibc} \ = \ \rho_b + \beta \Delta T_c + \mathbf{D}'_i \Delta \mathbf{T}_c \boldsymbol{\gamma} + \epsilon_c + \theta_{ibc}$$

where $\mathbf{D}_i$ is a vector of demographic variables measured either at the individual level or the block-group level.

We also separately estimate equation (3) by five-year age groups where we break up our sample into mutually exclusive sets of people born within the same set of five contiguous years.

Finally, we additionally estimate models where we allow for nonlinearities in treatment, in which case we estimate

$$(5) \qquad \Delta V_{ibc} \ = \ \rho_b + \beta \Delta T_c + \theta \Delta T_c^2 + \epsilon_c + \theta_{ibc},$$

where ΔT_c^2 is treatment squared (i.e., squared changes in number of days).

IV. Results

In this section, we discuss our main results. We first present means of voter registration variables and census covariates for counties with below and above mean change in early voting days, respectively. We then present our main aggregate turnout effects. We additionally show robustness of our main turnout effects by bandwidth followed by covariate balance tables by bandwidth. We then show placebos where we estimate the impact of our 2008–2012 treatment in other windows of time. We then break down our results by age, sex, and party. We end the main results section with models that are nonlinear in the number of days of early voting available.

In Table 1, we show the potential endogeneity issues of cross-county comparisons. We do this by comparing demographic and voting history characteristics of counties with above-mean versus below-mean change in number of early voting days. In online Appendix Table A.1, we also break down counties by above versus below mean change between presidential election years in hours, days open late (7 PM or later), weekend days, Sundays, and days with same-day registration, respectively.

The results for changes in number of days, as reported in Table 1, are broadly similar to those for the other treatment variables reported in online Appendix Table A.1. We also show average demographic characteristics from the census as well as average individual characteristics from the voter registration data in 30 rows (15 characteristics for each of expanding and contracting counties). At the bottom of the table, we show the numbers living in counties with expanding versus contracting early voting. A substantial majority of individuals saw expansions in hours, declines in days, expansions in weekend days, and declines in days with same-day registration.

Important for our identification strategy, there are substantial political and demographic differences that correlate strongly with the size and magnitude of the changes in early voting days. The distribution of changes in days is left-skewed. As shown in Figure 1, between 2008 and 2012, only 2 counties increased the number of days of

TABLE 1—SAMPLE MEANS OF REGISTERED OHIO VOTERS BY CHANGE IN EARLY VOTING DAYS

| | 2008–2012 changes in number of days | | |
	All	+/−	Mean
Black (percent)	11.6	+	7.6
		−	17.9
Hispanic (percent)	2.6	+	2.0
		−	3.6
White (percent)	83.7	+	88.6
		−	76.0
Democrat	30.4	+	27.4
		−	35.1
Independent	43.1	+	42.4
		−	44.1
Republican	26.5	+	30.2
		−	20.8
College graduate (percent)	25.9	+	23.2
		−	30.1
HS dropout (percent)	11.7	+	12.0
		−	11.3
Med. household income	55.3	+	54.6
		−	56.5
Age in 2008	44.6	+	45.1
		−	43.8
Distance to early voting site (miles)	10.9	+	10.7
		−	11.1
Voted in 2008	86.2	+	86.4
		−	85.9
Voted in 2010	59.9	+	60.3
		−	59.2
Voted in 2012	76.3	+	76.7
		−	75.6
Voted in 2014	43.7	+	45.0
		−	41.9
Observations	6,559,589	+	3,998,136
		−	2,561,453

Notes: Each row reports means of one variable indicated by the first column. Column "+/−" indicates a subsample of counties with above (+) or below (−) mean changes of early voting days between 2008 and 2012. Variable "Med. household income" is the median household income of a registered voter's census block group in thousands of dollars. "Distance to early vote site" is measured in miles. "Age in 2008" is measured in years as of the general Election Day in 2008. All other variables are in percentage points.

early voting, one by 4 days and the other by 15. In contrast, 20 counties decreased their early voting, 4 by between 5 days and 9 days. Counties with below-mean change were fully 12.5 percentage points less white. Counties with larger reductions were unsurprisingly also more African American. Though median household income varies by less than 5 percent across above-mean and below-mean counties, the college graduation rate in below-mean counties is 30 percent higher than in above-mean counties. Registered voters in above-mean counties are 9.4 percentage points more likely to have most recently participated in a Republican primary and 7.7 percentage points less likely to have participated in a Democratic primary. We geocoded polling stations and computed distance to early voting polling station for each individual

based upon their registration address. Average distance is approximately 10 miles and does not differ substantially across above-mean and below-mean counties. We also show turnout for 2008, 2010, 2012, and 2014, respectively. There are larger drops in turnout in counties with larger drops in number of days of early voting. However, demographic and political differences across expanding and contracting counties should give us pause in interpreting those differential changes in turnout as causally attributable to changes in early voting policy.

A. *Aggregate Turnout Effects*

We present our main effects in Table 2. The estimates are very tight in large part because the sample size is so large. Our estimates range from an increase in turnout of 0.0549 percentage points per additional day of early voting for the county fixed effects model to an increase of 0.218 percentage points for the baseline geographical fixed effects model. Three of the four models have relatively similar coefficients and are all statistically significant at a 95 percent level of confidence or higher. However, the county difference-in-differences model is substantially smaller and is statistically insignificant. As shown in Table 1, the places which expanded early voting hours were more Republican counties. The lower numbers in the county fixed effects model reflects declining support and thus lower turnout for President Obama in the more rural, Republican areas of Ohio. The Obama vote share remained largely stable in urban areas but declined by a couple of percentage points in rural areas where pro-Obama voters were less energized to turn out in 2012. Though we could control for voter demographics in the the county difference-in-differences model, bias is a problem if the statistical model does not include all relevant variables correlated with treatment and also if the functional form of the relationship between turnout and controls is not correctly specified. The geographic discontinuity model does not, by contrast, rely upon correctly specifying covariates or upon finding the correct functional form of the relationship between turnout and covariates.

The estimates are sizable. To put the estimates into context, a 16-day expansion of early voting would increase turnout by more than the difference between the maximum and the minimum turnout rates over the past four presidential elections (3.3 percentage points). Two days of early voting would increase voting by more than the addition of a newspaper in the late nineteenth and early twentieth century (Gentzkow, Shapiro, and Sinkinson 2011). Strömberg's (2004) estimates of the impact of radio on turnout suggest that a 10 percent increase in radio penetration in the 1930s is roughly equivalent to an increase in early voting of 6 days. The size of our effects also imply that the decline in turnout due to television from its introduction through 1970 (Gentzkow 2006) would be equivalent to a two-week reduction in early voting.

The standard errors for the individual fixed effects model restricted to the two-mile border sample are roughly the same magnitude as the full sample county difference-in-differences despite the fact that the sample size drops by slightly more than 95 percent. This is likely, at least in part, because the border samples are a more homogeneous sample so that the reduction in sample size does not come at the expense of higher standard errors. The standard errors rise with the final geographical

TABLE 2—EARLY VOTING EFFECTS ON TURNOUT

	Full sample		2-mile border sample	
	(1)	(2)	(3)	(4)
Number of days	0.055	0.201	0.137	0.218
	(0.037)	(0.077)	(0.039)	(0.077)
Year fixed effects	Y	Y	Y	
County fixed effects	Y			
Individual fixed effects		Y	Y⁻	Y⁻
Year-specific geographic fixed effects				Y
Observations	11,532,916	11,532,916	1,141,271	1,141,237

Notes: Each cell reports one coefficient estimate of an OLS regression. In all regressions, the dependent variable is a binary variable equal to 100 if a registered voter turns out to vote in a general election, and zero otherwise. The rows indicate the key explanatory variable in each regression. In columns 1 and 2, the samples include the full sample of registered Ohio voters and the specifications include county fixed effects and individual fixed effects, respectively. In columns 3 and 4, the samples include individuals living within two miles of a county border. Both specifications include individual fixed effects estimated taking four-year differences of the regression equations. Y^- indicates the allowance of individual fixed effects through four-year differencing. Column 4 additionally includes 2×2 mile geographic fixed effects. All samples include general elections in 2008 and 2012, and all specifications include year (election) fixed effects. Standard errors in columns 1 to 3 regressions are clustered by county. Standard errors in column 4 are clustered two-way by county and by county-border segment.

fixed effects model because they are clustered two-way on county and county-pair rather than just on county.[11] This tells us again that the comparison across county borders is apt because the increase in standard errors comes from accounting for positive correlation within a county-pair in addition to controlling for within-county correlation.

B. *Empirical Model Validation*

In this section, we describe a number of tests that we perform to validate our benchmark model. In Table 3, we show covariate balance by bandwidth for 12 covariates. In Table 4, we perform placebo estimates of our treatment effects in other time periods. In Table 5, we show robustness of our estimates to choice of bandwidth.

In the prior section, we presented geographical discontinuity estimates with a bandwidth of two miles. In this section, we motivate our bandwidth choice by running placebo estimates for a range of different bandwidths. We regress placebo outcomes on the change in early voting days between 2008 and 2012 conditional upon square block fixed effects for bandwidths ranging from 0.1 miles $\times$ 0.1 miles to 20 miles $\times$ 20 miles. Overall, we include 8 different block sizes including our benchmark block size of 2×2 miles. These results are shown in Table 3. Our individually measured placebo variables are dummy variables for independents, Democrats, and Republicans, age in 2008, sex, and distance to early voting station. We also put in census variables, measured at the individual's census block group, as additional placebos. These include percent college graduates, percent high school dropouts, median

---

[11] If we estimate the geographical discontinuity model and cluster only on county, then the standard errors are smaller.

TABLE 3—TESTS OF COVARIATE BALANCE BY AREA OF GEOGRAPHIC FIXED EFFECTS

	0.1	0.5	1	1.5	2	3	5	10	20
Independent	0.195	0.362	0.100	0.074	0.192	−0.128	0.064	−0.016	−0.198
	(0.253)	(0.206)	(0.212)	(0.219)	(0.361)	(0.176)	(0.346)	(0.277)	(0.339)
Republican	−0.085	−0.186	0.066	0.085	0.004	0.283	0.285	0.559	1.056
	(0.204)	(0.128)	(0.143)	(0.181)	(0.254)	(0.193)	(0.325)	(0.277)	(0.307)
Democrat	−0.110	−0.176	−0.166	−0.159	−0.196	−0.155	−0.350	−0.543	−0.858
	(0.196)	(0.137)	(0.121)	(0.119)	(0.148)	(0.139)	(0.145)	(0.209)	(0.239)
Age in 2008	−0.052	−0.007	0.043	0.042	0.038	0.072	0.058	0.067	0.131
	(0.112)	(0.063)	(0.062)	(0.086)	(0.073)	(0.071)	(0.089)	(0.065)	(0.071)
Female	0.265	−0.036	−0.020	−0.000	−0.047	−0.070	−0.072	−0.089	−0.083
	(0.143)	(0.051)	(0.045)	(0.052)	(0.042)	(0.065)	(0.076)	(0.062)	(0.063)
Distance to early voting site	0.001	0.004	0.005	0.012	0.006	0.020	0.013	0.045	0.218
	(0.001)	(0.004)	(0.010)	(0.019)	(0.021)	(0.034)	(0.062)	(0.080)	(0.155)
College grad. (percent)	−0.239	−0.337	−0.208	−0.194	−0.503	−0.156	−0.263	−0.145	−0.474
	(0.269)	(0.274)	(0.200)	(0.251)	(0.348)	(0.278)	(0.377)	(0.313)	(0.338)
HS dropout (percent)	0.074	0.049	0.064	0.078	0.162	0.059	0.126	0.002	−0.101
	(0.103)	(0.105)	(0.086)	(0.079)	(0.112)	(0.115)	(0.130)	(0.135)	(0.107)
Med. HH. income	−0.261	−0.512	−0.269	0.037	−0.577	0.289	0.258	0.427	0.872
	(0.342)	(0.412)	(0.411)	(0.547)	(0.549)	(0.525)	(0.798)	(0.604)	(0.627)
Hispanic (percent)	0.038	0.042	−0.011	0.007	−0.031	−0.014	−0.031	−0.046	−0.097
	(0.021)	(0.023)	(0.019)	(0.020)	(0.032)	(0.023)	(0.041)	(0.043)	(0.058)
Black (percent)	0.019	0.012	−0.061	−0.159	0.159	−0.339	−0.291	−0.564	−1.353
	(0.092)	(0.089)	(0.073)	(0.321)	(0.240)	(0.366)	(0.405)	(0.374)	(0.440)
White (percent)	0.025	0.019	0.112	0.184	−0.014	0.385	0.401	0.718	1.600
	(0.114)	(0.131)	(0.106)	(0.348)	(0.271)	(0.387)	(0.445)	(0.398)	(0.462)
Omnibus test (F-statistic)	1.01	1.23	0.61	0.42	0.86	0.80	0.96	1.82	5.15
Omnibus test (p-value)	0.436	0.258	0.835	0.958	0.587	0.651	0.485	0.040	0.000
Observations	53,444	273,582	562,230	855,736	1,141,237	1,703,635	2,669,114	4,260,076	4,322,569

Notes: Each cell reports the estimated coefficient of the number of early voting days in an OLS regression. In all regressions, the dependent variables are the same for each row and are indicated by the first column. "Med. HH. income" is the median household income of a registered voter's census block group in thousands of dollars. "Distance to early vote site" is measured in miles. "Age in 2008" is measured in years as of the general Election Day in 2008. "Democrat" is equal to 100 if the most recent primary the registered voter participated in since 2000 is a Democratic primary, and zero otherwise. "Republican" is similarly defined by Republican primary participation. "Independent" is defined as registered voters who have not voted in the primary of either party between 2000 and 2008. "Female" is the percentages of females who, according to the Social Security Administration, were born in the year of birth of the registrant and share the registrant's first name. Black, white, and Hispanic are census block-level population shares. Other demographic variables are measured in percentages at the block-group level. The sample is restricted to individuals living within k miles of a county border, where k is indicated by the column headings. Standard errors are clustered two-way by county and by county border.

household income, percent Hispanic, percent black, and percent white. Out of our 12 placebos, none are statistically significant for one-, two-, or three-mile blocks. We pool the estimates and test for joint significance across covariates. The *p*-value for joint significance on our 2×2 mile benchmark is 0.587. Only bandwidths 10 and 20 are statistically significant at even a 10 percent level. The larger bandwidths are analogous to a neighboring county-pair design; the failure of placebos for these large bandwidths also suggest additional endogeneity problems for the two-way fixed effects design.

We also consider the possibility that people might not differ systematically across county borders but that counties with early voting contraction are counties that were more generous in early voting access as well as more limited in purging of voters. Counties that purge inactive voters from the voter rolls more liberally might also

TABLE 4—PLACEBO TESTS

Dependent variable	Change in turnout between							
	2008−2010		2010−2012		2008−2012		2010−2014	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
△ days: 2008–2010	0.120 (0.165)	0.200 (0.175)			0.083 (0.129)			
△ days: 2008–2012		−0.111 (0.130)		0.180 (0.178)	0.185 (0.080)	0.269 (0.120)	0.067 (0.189)	
△ days: 2010–2012			0.261 (0.146)	0.117 (0.215)		−0.083 (0.129)		
△ days: 2010–2014							0.080 (0.297)	0.134 (0.188)
Observations	1,141,237	1,141,237	1,141,237	1,141,237	1,141,237	1,141,237	1,141,237	1,141,237

Notes: Each column reports one placebo test. The dependent variables are the change in voter turnout between years that are indicated by the column header. The dependent variables are scaled by 100 so that they take a value of −100, 0, or 100. Each row reports the estimated coefficients of the change in early voting days between years that are indicated by the first column. In all regressions, the sample is limited to individuals living within two miles of a county border. All regressions include a set of 2×2 mile geographic fixed effects. Individual fixed effects for turnout propensity are differenced out. Standard errors are clustered two-way by county and by county border.

TABLE 5—THE TURNOUT EFFECTS OF EARLY VOTING LAWS: BY SPATIAL BANDWIDTH

Bandwidth (b)	0.1	0.5	1	1.5	2	3	5	10	20
	$b \times b$ geo fixed effects								
Number of days	0.204 (0.117)	0.189 (0.108)	0.241 (0.114)	0.214 (0.120)	0.218 (0.077)	0.197 (0.094)	0.224 (0.112)	0.235 (0.094)	0.326 (0.098)
Observations	53,444	273,582	562,230	855,736	1,141,237	1,703,635	2,669,114	4,260,076	4,322,569
	Border segment fixed effects								
Number of days	0.185 (0.099)	0.174 (0.115)	0.186 (0.097)	0.218 (0.093)	0.221 (0.097)	0.197 (0.099)	0.199 (0.097)	0.235 (0.102)	0.247 (0.105)
Observations	55,584	274,718	562,616	855,835	1,141,271	1,703,645	2,669,117	4,260,076	4,322,569

Notes: The bandwidth b in the top row indicates that the estimates in the column below are estimated from a border sample that includes individuals living within b miles of internal county borders. Each cell reports the estimated coefficient of the number of early voting days from one OLS regression. In the upper panel, the samples include individuals living within b miles of the county border, and the specification includes $b \times b$ geo fixed effects. In the lower panel, the samples include individuals living within b miles of the county border and the specification includes border segment fixed effects. The dependent variable is a binary variable equal to 100 if a registered voter turns out to vote in a general election, and zero otherwise. Individual fixed effects have been differenced out. Standard errors are clustered two-way by county and by county border. Each sample includes 87 counties and 222 county-border segments.

have lower turnout unrelated to early voting policy. Though we see no correlation between pre-existing generosity of voting and average age of registered voters in our placebos, it is possible that more restrictive counties purge both younger voters who move and older voters who move or pass away. Thus, it is possible that our null effect on age differences above is consistent with differential purging across counties but of both younger and older individuals with no net effect on mean age. We thus consider additional balance placebos using: date last voted; date last voted before the 2008 general election; a dummy for never having voted; a dummy for never having voted prior to the 2008 election; and 7 dummies, one for each decade of age (20, 30, 40, 50, 60, 70, and 80). Of these 11 dummies, 2 are statistically

significantly different at the 10 percent level and none at the 5 percent level. We present these results in online Appendix Table A.2.

If contracting counties were differentially purged by county voter registrars, we should then see fewer total registrants in those counties. We test for this explicitly in online Appendix Table A.3. We compute the number of voters in each census block. We then regress, at the census block level, numbers of registrants on the change in early voting between 2008 and 2012 controlling for 2×2 mile block fixed effects. We do this for number of individuals registered before 2008, number of individuals registered before 2012, and the difference between the number of registrants by 2012 and the number by 2008. Our coefficients are uniformly small and statistically insignificant.

Though we find no systematic differences correlated with the change in early voting within 2×2 mile blocks, we are nonetheless concerned about unobservable differences across counties even within small geographical areas that drive changes in turnout and are correlated with early voting changes. As a result, we perform additional placebo tests. There were differential changes across counties in early voting availability between 2008 and 2010, as well as uniform changes across all counties between 2012 and 2014. We add to our baseline regressions these additional changes. We find that our baseline estimates remain of a similar magnitude and retain statistical significance, while the placebo treatments show up as statistically insignificant. We also estimate placebo effects of the 2008–2012 changes in early voting on the change in turnout between 2008 and 2010, 2010 and 2012, and 2010 and 2014. We estimate these specifications without any other treatment variables as well as conditional upon the actual changes in early voting hours during the actual time periods. In all cases, we find small and statistically insignificant effects of placebo turnout on voting.

Finally, we augment our validation by showing that our estimates are robust to choice of bandwidth. In Table 5, we show our estimates. The estimates are remarkably stable across bandwidths. Except for the largest (20 mile) bandwidth, all 7 other bandwidth choices yield estimates within 15 percent of our main estimated treatment effect. This reflects an absence of endogeneity bias as seen in the stability of covariate balance across bandwidths for the changes in days of early voting. However, it also reflects stability of the treatment effects across bandwidths, which suggests that the external validity of our results do not suffer from the sample restriction. We also show, in a second panel, that our results are robust to the method of Dell (2010), which estimates with border segment (county-pair segments) fixed effects instead of block fixed effects. Except for the largest bandwidth, the border segment fixed effect model yields very similar estimates to the block fixed effect model, and thus the estimates also are robust across bandwidths. Overall, we find substantial evidence that our estimates are plausibly causal and that they are also robust to specification.

C. *Party Effects*

Typically, the Democratic Party has fought to expand early voting and the Republican Party has fought to reduce it (Biggers and Hanmer 2015). We now

TABLE 6—THE TURNOUT EFFECTS OF EARLY VOTING LAWS: BY PARTY

	(1)	(2)
Number of days	−0.097	−0.059
	(0.068)	(0.053)
Days × Independent	0.718	0.612
	(0.113)	(0.095)
Days × Democrat	0.129	0.109
	(0.021)	(0.020)
Individual fixed effects	Y	Y
Year-specific geographic FE	Y	Y
Subsample (18 years old by 2000)		Y
Observations	1,081,750	1,029,446

Notes: Each column reports the coefficient estimates of an OLS regression. The dependent variable is the four-year difference of a binary variable, which equals to 100 if a registered voter turns out to vote in a general election, and zero otherwise. The rows indicate the explanatory variables in each regression. In all regressions, the sample is limited to individuals living within two miles of a county border. Column 2 is restricted to individuals who had turned 18 by the year 2000. Party affiliation is identified by the most recent primary vote before the 2008 general election. All regressions include a set of 2×2 mile geographic fixed effects. Individual fixed effects for turnout propensity are differenced out. Standard errors are clustered two-way by county and by county border.

ask whether political parties are acting in a way that is consistent with their own interest. Of course, parties acting in their own interest may also be ideologically motivated. In this section, we will estimate the partisan impacts of early voting expansion and contraction for Democrats, Republicans, and independents. To be clear, we are not estimating the causal impact of party on the treatment effect of early voting expansion. Party preferences are correlated with gender, race, education, and many other determinants of political preferences. We do not try to isolate the pure impact of party. However, the differential impact by party (and age and gender) is of great importance both politically and legally. We thus focus on estimation of differential impacts by party (and in other sections, by age and gender).

In order to estimate early voting impacts by partisan affiliation, we first measure partisanship at the individual level. For those who have participated in a primary, we record their partisanship as the party whose primary they most recently voted in. For those who have never voted in a primary or for the very small number of individuals who have most recently voted in a third-party primary, we record them as independents. We also consider estimates on a sample of those who turned 18 by the year 2000 and thus had greater chance to declare partisan leanings through primary participation by the year 2008. We consider this second sample our preferred one due to better measurement of partisanship. We then separately estimate the impact of an additional day of early voting upon voter turnout for Democrats, Republicans, and independents. Our results are reported in Table 6.

We regress change in turnout from 2008 to 2012 on change in days, controlling for geographical block fixed effects. We also regress on change in early voting interacted with a dummy for Democrat as well as a dummy for independent. The baseline change in days coefficient can, therefore, be interpreted as the effect for

Republicans, whereas the other two coefficients reflect the additional effects upon Democrats and independents.

An additional day of early voting is estimated to have a -0.097 effect on Republican turnout in presidential elections; however, the coefficient is not statistically significantly different from zero. In the sample of those who were 18 years of age by 2000, the coefficient is -0.059. The coefficient for Democrats is slightly more than 0.1 percentage points higher and is statistically different from zero with greater than a 99 percent level of confidence. The effect for independents is very large at 0.621. The large size of the impact upon independents underscores that independents are more weakly attached to politics, and in presidential elections, increasing the availability of voting has a large impact. The way we measure independents is by their participation in primaries. This is the only measure available to us because Ohio is an open primary state. Having said that, our measures of Democrats, Republicans, and independents roughly correspond to what is found in closed primary states such as Florida, North Carolina, and California.

If we view independents as more politically moderate, then early voting has a de-polarizing impact upon the vote in presidential elections. In online Appendix Figure A.2, we show that independents are much more likely to identify themselves as ideologically moderate as opposed to conservative or liberal than either registered Republicans or registered Democrats.

Republicans seem to be the most reliable voters. A higher fraction of Democrats and Republicans both turn out for presidential elections than do independents. The marginal voters are thus independents who are more politically indifferent in presidential elections. Easing access to voting differentially impacts Democrats but impacts independents to an even greater degree during presidential election years.

There are three caveats that limit the interpretation of our estimates on differential effects by partisan leaning as effects upon the partisan vote share. First, we do not know that those who have voted in a party's primary will vote for that party in the general election. Second, we do not know how independents vote. However, as shown in online Appendix Table A.4, the correlation between our measure of Democratic vote share and the precinct-level vote share is 0.571 for 2008. The correlations are surprisingly high given that different people turn out from election to election. Finally, in order to compute partisan vote share impacts, we need to weight Republicans and Democrats by their voter registration shares. We do this in Section VI.

D. *Effects by Age*

The heterogeneity in the effect of early voting expansion by partisan affiliation is interesting in large part because it is informative about the impact on the partisan vote share. We next turn toward estimation of differential effects by age. These estimates are interesting not only inherently but also because they are informative about who the marginal voter is and what that tells us about the costs and benefits of voter turnout. We next estimate the heterogeneity in the effect of early voting expansion by age. Age heterogeneity tells us about the age profile of the marginal voter and thus about the life-cycle determinants of turnout.
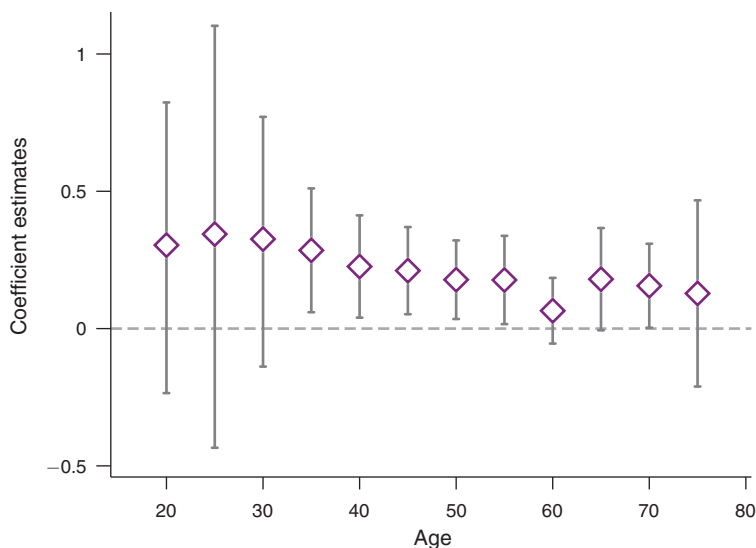
FIGURE 3. HETEROGENEOUS TREATMENT EFFECTS OF EARLY VOTING ON AGE GROUPS

Notes: The graph above plots age-group-specific coefficients with 95 percent confidence intervals from a model with turnout regressed upon early voting days. Data from the 2008 and 2012 general elections are included. Each age group spans five years. Individual fixed effects are differenced out and age-group-specific 2×2 mile geographic fixed effects are included. The dependent variable is a binary variable equal to 100 if an individual voted, and zero otherwise.

We use our main identification border discontinuity design strategy to estimate the effect of an additional day of early voting by age. Since there are not many registered voters of a given age within a 2×2 mile block, we group individuals into bins by five-year age groups starting with the group 18–22. Each group is centered around a multiple of 5: 20, 25, 30, etc. The final group we use is the one centered around 75 years of age. After the 75-year-old group, the numbers become too thin to estimate effects upon.

We show our estimates graphically in Figure 3. We list the estimated treatment effect for a group on the *y*-axis and the median age of the age group on the *x*-axis. The first thing that we note is that the effects are positive for all of the 12 age groups. Second, we note that all age group pairs have overlapping 95 percent confidence intervals. We do not have the statistical power to differentiate the heterogeneity of effect by age group. We also estimate effects solely using cross-county variation. We present this graph in online Appendix Figure A.3. The effects show a similar age pattern but are more pronounced.

In Figure 3, we see that early voting has a greater impact on younger voters. The four age groups with the largest estimated effects are the four youngest with the peak effect being for 23–27-year-olds. The largest effects are approximately 0.344 percentage points per additional early voting day for 23–27-year-olds. The effect for the four youngest age groups is a statistically insignificant 0.140 percentage points per day higher than the effect for those older than 37. This age group is largely comprised of working parents with infants and young children. The median age of first birth for women in Ohio was 25 in 2006 (Matthews and Hamilton 2009)

and nationally, the first age of first birth for men is two years more than that for women.

E. *Effects by Gender*

We also estimate the impact of early voting expansion by gender. In contrast to age, which Ohio voting records measures directly, Ohio does not record gender or sex on voter registration forms. Therefore, we only indirectly measure gender. We impute gender probabilities for each individual in our dataset by matching first names by year of birth to lists of first names by gender and year of birth from the Social Security Administration as described in Section II. For uncommon first names (those with less than five individuals of a given sex born in a given year for both genders), we cannot match them to the Social Security files and we drop them. This is 4.6 percent of our sample. For the remaining sample, we estimate equation (4). We do this in two ways. First we interact our treatment variable with the estimated probability that an individual is female. Second, we create a binary variable taking on the value of 1 if the probability of being female is at least 95 percent and 0 if the probability of being female is less than or equal to 5 percent. For this second specification, we drop all observations with a probability of being female in between 5 percent and 95 percent. As shown in Table 7, using the binary variable drops the sample size by only 4.5 percent, reflecting that most names are either definitively male or definitively female.

In addition to estimating models with continuous and binary gender measures, we also estimate the impacts using gender imputed by national Social Security lists as well as state of Ohio Social Security lists. The state of Ohio lists are smaller and thus fewer names can be matched. However, if gender specificity in naming varies by state, the Ohio data is probably more accurate for the Ohio voting population. Using the state lists lowers the sample size by 4.2 percent for the continuous measure of gender and 2.8 percent for the binary measure of gender. In the text, we report estimates using the continuous measure of gender and from the national sample. However, all estimates of differential effects by gender are very similar. In all specifications, switching from national to state or from continuous to binary impacts the estimated coefficient by less than 5 percent.

We find robust evidence that there is a differential effect across the genders. For men, an additional day of early voting increases turnout by 0.174 percentage points. This coefficient is statistically significant with more than a 95 percent level of confidence. There is an additional 0.054 impact for women, which is statistically significant at more than a 99 percent level. The tightness of the standard errors on the gender coefficient reflects the uniformity of systematic differences in voting behavior across the sexes. The effect of early voting laws on female turnout is roughly 31 percent higher than that for men.

F. *Effects by Age and Gender*

We also show estimates by age group broken down by gender. Since we have small numbers of men and women respectively in many of our geographical cells, we

Table 7—The Turnout Effects of Early Voting Laws: By Gender

	Inferring gender national first names		Inferring gender Ohio first names	
	(1)	(2)	(3)	(4)
Number of days	0.174 (0.077)	0.174 (0.077)	0.174 (0.078)	0.169 (0.078)
Days × Pr(*Female*)	0.054 (0.013)		0.057 (0.012)	
Days × *Female*		0.048 (0.012)		0.054 (0.012)
Observations	1,104,916	1,055,416	1,058,933	1,026,679

Notes: Each column reports the coefficient estimates of an OLS regression. The dependent variable is the four-year difference of a binary variable, which equals to 100 if a registered voter turns out to vote in a general election, and zero otherwise. The rows indicate the explanatory variables in each regression. In all regressions, the sample is limited to individuals living within two miles of a county border. Pr(*Female*) is the probability of an individual being female based on the gender frequency in Social Security administrative data of his/her first name. *Female* is an indicator variable equal to one if Pr(*Female*) $\geq$ 0.95, zero if Pr(*Female*) $\leq$ 0.05, and missing otherwise. In the left panel, Pr(*Female*) is inferred from the national frequency of females based on national birth records in the same birth year of the individual; in the right panel, Pr(*Female*) is inferred from the national frequency of females based on Ohio birth records in the same birth year of the individual. All regressions include a set of 2 × 2 mile geographic fixed effects. Individual fixed effects for turnout propensity are differenced out. Standard errors are clustered two-way by county and by county border.

estimate treatment effects using a two-way county-time fixed effects model with days interacted with age group. We estimate for men and women separately. Our results are in online Appendix Figure A.4.

In general, we do not see differential patterns by age across males and females. We see low estimates for the youngest age group followed by large and declining estimates. Both for men and women, estimates are highest for those in their late 20s and 30s. These estimates are not as well identified as the prior estimates by gender alone and by age alone. However, they are suggestive that life cycle effects are strong and that they are present both for men and women.

G. *Effects by Type of Early Voting Day*

Having shown heterogeneity of effects across different types of voters (by partisanship, by gender, and by age), we now look at the differential impact by type of early voting day. Online Appendix Figure A.1 shows the changes in total hours of early voting, number of weekend days, number of Sundays, number of days of same-day registration, number of weekend days with same-day registration, and number of days for presidential elections.[12] Most counties saw expansions in number of weekend days as well as number of Sundays. The counties that saw declines in weekend or Sunday early voting were the large urban counties. Same-day registration days are early voting days more than 28 days before the election when people

---

[12] Cantoni and Pons (2019) finds sizable effects of same-day registration using a cross-state research design.

could still register to vote and then actually vote at the same early voting polling station. Only two counties saw increases in same-day registration between 2008 and 2012. All other counties saw reductions in early voting. The larger declines occurred in the more urban areas. Since we have the exact hours that polling stations were open on each day, we also computed the number of days that polling stations were open until 7 PM or later (which we term days open late). Most counties saw an increase in days open late. The only exceptions were four counties with large, urban populations.

We estimate heterogeneous effects by type of treatment in Table 8. The estimates are noisy and thus not statistically significant even when sizable. Our point estimates, however, point to large impacts of Sunday early voting and same-day registration days.[13]

H. *Nonlinear Treatment Effects*

The average impact of an additional day of early voting is 0.218 percentage points of additional turnout. However, some counties saw large contractions in early voting, others large expansions, and yet others very modest changes or even no change at all. In this section, we test whether turnout is linear in the change of early voting days. We first difference the data at the individual level between 2008 and 2012 and include the quadratic term of the change in early voting days. The estimation is given by equation (5). We show these results in column 1 of Table 9. The quadratic term is relatively small and statistically insignificant. The estimated marginal effect of one more day of early voting when there is a 10-day increase is only 15 percent larger than the estimated marginal effect when there is a 10-day decrease. We also estimate the potential nonlinearity in the number of early voting days by including the difference of squared early voting days in 2012 and in 2008. Again, we do not find statistically significant nonlinear effects of early voting days on turnout.

We also consider the possibility that changing the number of early voting days may have a different long-run from short-run impact due to short-run confusion. In particular, expansions may have a smaller impact in the short run than contractions. People may not realize that early voting changes have occurred. Thus, people may plan to vote early only to realize that they are too late because the timing of early voting availability has changed. However, people who would otherwise not have voted may not be aware of expansions and may, thus, underutilize them. To deal with these concerns, we estimate the effects of expansions and contractions separately. We then re-estimate, restricting to cases where one of the counties had no change in early voting days and the other side had either an expansion or a contraction. Column 2 of Table 9 shows the full sample estimates. We do find that the impact of contractions is larger than the impact of expansions. Our point estimates suggest that contractions reduce turnout by 0.244 percentage points per day while expansions only increase turnout by 0.071 percentage points.

---

[13] Since only 22 county pairs contain variation in changes in numbers of Sundays, we also estimated our results using the two-way wild cluster bootstrap. The *p*-value for the only significant coefficient in Table 8 (column 3) drops below 0.01. Other coefficients remain insignificant at even a 10 percent level.

TABLE 8—THE TURNOUT EFFECTS OF EARLY VOTING LAWS: BY TYPE OF DAY

	(1)	(2)	(3)	(4)	(5)
Days	0.071	0.200	0.373	0.073	0.029
	(0.071)	(0.124)	(0.148)	(0.069)	(0.199)
Weekend days	0.173				0.181
	(0.144)				(0.160)
Days with same-day registration		0.084			0.166
		(0.674)			(0.681)
Days open late			−0.067		
			(0.066)		
Saturdays				0.001	
				(0.173)	
Sundays				0.286	
				(0.239)	
Observations	1,141,237	1,141,237	1,141,237	1,141,237	1,141,237

Notes: Each column reports the coefficient estimates of an OLS regression. The dependent variable is the four-year difference of a binary variable, which equals to 100 if a registered voter turns out to vote in a general election, and zero otherwise. The rows indicate the explanatory variables in each regression. In all regressions, the sample is limited to individuals living within two miles of a county border. All regressions include a set of 2×2 mile geographic fixed effects. Individual fixed effects for turnout propensity are differenced out. Standard errors are clustered two-way by county and by county border. In the wild bootstrapping tests with error clustering at the county level, neither the number of Saturdays nor the number of Sundays significantly affect turnout.

TABLE 9—THE TURNOUT EFFECTS OF EARLY VOTING LAWS: NONLINEARITY

	(1)	(2)	(3)
ΔDays	0.179		
	(0.047)		
ΔDays^2	−0.007		
	(0.007)		
min$\{\Delta$Days, 0$\}$		0.244	0.190
		(0.095)	(0.059)
max$\{\Delta$Days, 0$\}$		0.071	0.160
		(0.071)	(0.044)
Individual fixed effects	Y	Y	Y
Geographic fixed effects	Y	Y	Y
Subsample			Y
Observations	1,141,237	1,141,237	753,491

Notes: Each column reports the coefficient estimates of an OLS regression. The dependent variable is the four-year difference of a binary variable, which equals to 100 if a registered voter turns out to vote in a general election, and zero otherwise. The rows indicate the explanatory variables in each regression. In all regressions, the sample is limited to individuals living within two miles of a county border. The last column, i.e., the column with "Subsample" marked as Y, restricts the sample to those counties with at least one county experiencing no change in the number of early voting days between 2008 and 2012. All regressions include a set of 2×2 mile geographic fixed effects. Individual fixed effects for turnout propensity are differenced out. Standard errors are clustered two-way by county and by county border.

In cases where we have two neighboring expanding counties, two neighboring contracting counties, or one of each, it may be difficult to cleanly identify expansion or contraction effects. We thus also show estimates that restrict to comparisons where one side of the border had no change in aggregate days. This shrinks our sample size by 36 percent. Nonetheless, the greater comparability of the treatments

reduces the standard errors for the expansion coefficient. The two coefficients, in this sample, are quite close: the coefficient on contractions is 0.190 and that on expansions is 0.160. Thus, our evidence suggests that contractions have a similarly sized impact to expansions, and thus likely long-run effects and short-run effects are similar.

V. Aggregate Effects

In this section, we use our geographic fixed effect estimates on turnout in presidential elections to simulate the impact of the Kasich reform as well as three different benchmark scenarios for national early voting legislation. In the case of our national simulation, this is made under the maintained assumption that our estimates from Ohio are externally valid.

A. *Ohio Impacts*

We use the estimates by party to estimate the impact on voter turnout and on the Democratic vote share of the Kasich reform for 2012. For turnout effects, we multiply the estimated effect by the number of registered voters in each county and then multiply by the change in the number of days. We then add up across counties to get the total turnout effect. We express this in the equation below:

$$(6) \qquad\qquad T = \sum_c \beta \mu_c R_c,$$

where β is our estimated effect per day of early voting on turnout, μ_c is the change in the number of days of early voting available, and R_c is the number of registered voters in a county. We find that though many counties increased early voting days between 2008 and 2012, large reductions in dense urban counties like Cuyahoga, Franklin, and Summit more than outweighed the early voting expansions. The net effect was to reduce total voting by 45,225 votes in the 2012 election.

We now look at the impact on the Democratic vote share. In order to do this, though we have estimated the impact of early voting expansion by party, we face two main problems. First, we do not know that all Democrats vote Democratic and all Republicans vote Republican. Second, we don't know who independents vote for. We proxy the probability of voting for the Democrats using the precinct-level correlation between a partisan group's registration share and the aggregate Democratic vote share. We show these correlations by year and party in online Appendix Table A.4. The correlation coefficients are decently stable across elections. The Republican registration share correlation with the Democratic vote share ranges is -0.769 in 2012 and -0.822 in 2008. The Democratic registration share is significantly lower mainly because independents lean heavily Democrat. The correlation is 0.548 in 2012 and 0.571 in 2008. The independent share is positively correlated with the Democratic vote share. It is also unsurprisingly more unstable. The correlation for the independent share is 0.380 in 2012 and 0.297 in 2008.

We then compute the net vote change for Democrats by Democrats, Republicans, and independents. We start by computing the expected increase in votes for Democrats per registered Democratic primary voter. This is obtained by multiplying the effect of an additional early voting day on a registrant of party p by the probability that a registrant of party p votes for the Democrats. We denote by β_p the turnout effect for registrants with party p and by ρ_p, the correlation between registration shares in a precinct and the Democratic vote share in the precinct. We then multiply this by the number of registered party p voters in county c in election e: ω_{pc}. Altogether, this gives us the expected net change in Democratic votes from a one-day increase in early voting in county c. Finally, we multiply this by the net change in days of early voting in the county that we denote by μ_c. The expected increase in votes for Democrats in county c is thus $\beta_p \rho_p \mu_c \omega_{pc}$. Our equation for the net change in Democratic votes, T_p, is given by summing over all counties:

$$T_p = \sum_c \beta_p \rho_p \mu_c \omega_{pc}.$$

We compute the total effect on Democratic votes by adding the effect on Democrats to that on independents as well as the effect on Republicans. We then divide by total votes in the election to get the impact of the Democratic vote share:

$$\Delta V_p = \frac{T_D + T_R + T_I}{Turn},$$

where D denotes Democrat, I denotes independent, R denotes Republican, and $Turn$ is the actual total election turnout.

On net, our estimates imply an increase in the Republican vote share of 0.36 percentage points in the 2012 presidential election. This may seem small given the magnitude of the contractions in Democratic counties combined with the fact that some Republican counties actually saw increases in days. However, a few things must be kept in mind. First, the change in the number of days matters more than the distribution of changes over counties. The reason for this is twofold. First, the effects on Republicans and Democrats are small. Therefore, the effects upon the Democratic vote share largely rely upon the magnitude of changes to independents. In addition, the differences across counties in partisanship are modest. Going from the twenty-fifth to seventy-fifth percentile in Democratic share of registrants only increases the Democratic registrant share by 10 percentage points. Moreover, much of the overall effect is concentrated in the very large, urban counties that lean Democrat less heavily than the very rural areas lean Republican. Overall, the changes to early voting in Ohio had a positive though modest impact on the Republican vote share.

B. *Impacts on Federal Election Outcomes*

We now simulate the effect of three potential national early voting laws. The first scenario is a national ban. The second scenario is a national mandate at 23 days of early voting. This is what the state of Ohio currently provides. For comparison, on average, 19 days of early voting are available for the 37 states and DC that currently

allow early voting. Finally, we consider a third scenario with double Ohio's provision of early voting: 46 days. Minnesota has the most generous early voting in the country and it has 46 days of in-person early voting. One caveat to our results is that the ban and the 46-day expansion both extrapolate linearly out of sample. Though our estimates show evidence of linearity, it is possible that increasing the amount of early voting to 46 days may reduce the marginal effects of early voting expansions, and even more so, it is possible that a shift on the extensive as opposed to intensive margin may have an additional impact.

To simulate the impact of these three scenarios, we first compute the impact per additional day of early voting on the Democratic vote share. For each party, we multiply the effect of an extra day of early voting on each group (Democrats, Republicans, and independents) by the probability for each group of voting Democrat; we then multiply this product by the share of each group in the registered population.[14] We obtain slightly more than a percentage point increase in the Democratic vote share if 23 days of early voting were nationally available.

We now move from computing the impact upon the two-party Democratic vote share of an additional day of early voting to the impact on the outcome of national elections of our three different national early voting scenarios. We can compute the change in election outcome for chamber c, under scenario r, and during year y. We express the outcome change as ΔO_{cry}. An outcome is the number of seats for House and Senate elections and number of electoral votes for presidential elections. We also denote by α_{csy} the change in early voting days for scenario c, state s, and year y.[15] The function $F(\alpha_{csy})$ takes on $+1$ if plurality in a state changes toward the Republicans, -1 if plurality in a state changes toward the Democrats, and 0 otherwise. Finally, we denote by E_{cs} the electoral votes in state s and chamber c. For House and Senate elections, the value of E_{cs} is 1. For presidential elections, the value of E_{cs} is equal to the electoral votes in the state.[16] The formula we use for computing outcome changes for national elections is thus given by

$$\Delta O_{cry} = \sum_s \Theta_e F(\alpha_{csy}) E_{cs}.$$

We present the results of our predictions in Table 10. Impacts upon independents are large in presidential elections where they are the marginal voters and they swing Democrat in their voting patterns. In 2012, we predict that in the Senate, one state, North Dakota, would have swung toward the Republicans with the elimination of early voting and one, Nevada, toward the Democrats with 23 days of early voting. Nevada electing a Democrat in 2012 would have given the Democrats the majority in the Senate in 2012. In 2016, we find no impact of getting rid of early voting in the Senate; however, we find that one seat (Pennsylvania) would have switched to the Democrats with 23 days of early voting and 2 additional seats would have switched

[14] We use the average across the 2008 and 2012 elections to compute the correlation coefficients and the group shares that we use in this equation.

[15] s denotes House district in the case that the chamber, c, is the House of Representatives.

[16] In the case of Maine and Nebraska, each electoral district decides its own elector and the remaining two electors are decided by the plurality outcome in the state. For these two states, s indexes each of the electoral districts as well as the state.

TABLE 10—CHANGES OF REPUBLICAN SEATS AND ELECTORAL COLLEGE VOTES UNDER HYPOTHETICAL
STANDARDIZED EARLY VOTING

Election type	Year	Observed Republican seats/electoral votes	Standardized early voting		
			0 days	23 days	46 days
President	2012	206/538	47	0	−15
President	2016	304/538	10	−65*	−75*
Senate	2012	51/100	1	−1*	−1*
Senate	2016	52/100	0	−1	−3*
House	2012	234/435	7	−6	−11
House	2016	241/435	6	2	−6

Notes: Each element of the columns under the heading "Standardized early voting" reports the simulated impact of national legislation requiring 0, 23, and 46 days of early voting, respectively. Impacts are computed using estimates by party and election type from Ohio but are applied nationally. For "Senate" and "House" rows, numbers reflect the change in the number of seats to the Republican Party. For the "President" row, numbers reflect the change in the number of electoral votes to the Republican Party. Positive numbers indicate a net shift in favor of the Republican Party, and negative numbers indicate a net shift in favor of the Democratic Party. * indicates a change of majority control in the Congress or majority of electoral votes.

with 46 days (Missouri and Wisconsin). Three additional seats in 2016 would have flipped majority control in the Senate from the Republicans to the Democrats.

For House elections, our simulations suggest no shift in majority control from any of the scenarios that we consider. However, we do find that elimination of early voting would have given the Republicans 7 additional seats in the 2012 election and 6 additional seats in the 2016 election. Nevertheless, a 46-day law would only have induced a movement toward the Democrats of 11 seats in 2012 and 6 in 2016.

Finally, we consider the impact upon the 2012 and 2016 presidential elections. We find that elimination of early voting in 2012 would have pushed Florida and Ohio to Mitt Romney and 46 days would have led to Obama winning in North Carolina. Neither of these changes would have been pivotal for the overall election outcome. However, in the very close 2016 election, we find that both 23 and 46 days of early voting would have yielded Democratic victories in Florida, Michigan, Pennsylvania, and Wisconsin and thus would have altered the 2016 presidential election outcome. Additionally, elimination of early voting would have led to Republican victory in Minnesota.

VI. Conclusion

In this paper, we estimate the impact of early voting upon voter turnout. We compare people within the same square mile block on opposite sides of county borders when Ohio Governor John Kasich passed laws homogenizing early voting across counties. We find that a day extra of early voting increases turnout by 0.218 percentage points. We additionally show evidence that those in child-rearing years and prime working years are particularly impacted by early voting availability. We further find that women react over 30 percent more strongly than men to additional early voting. We do not find strong responses to days where polls are open late. However, we do find a large (though statistically insignificant) differential turnout response to Sunday voting as well as to same-day registration. The methods we use

for this paper are also well suited for looking at heterogeneity by race, which is crucial for electoral law in the United States.

We further find that effects are larger on Democrats than on Republicans and that effects on independents are very large. We use our estimates on partisan impacts of early voting to simulate the impact of national early voting legislation and find that requiring all states to implement 23 days of early voting (as is currently the case in Ohio) would have altered the outcome of the 2016 presidential election as well as majority control of the Senate in the 2012 election. Additionally, we find that 46 days of national early voting (at the current level of Minnesota) would have swung majority control of the Senate to the Democratic Party in the 2016 elections.

Finally, we find that early voting expansion likely has a de-polarizing effect on the electorate in that independents are most impacted. Overall, our evidence demonstrates substantive electoral impacts of early voting on turnout, on partisan outcomes, and on the polarization of the electorate.

REFERENCES

Ansolabehere, Stephen, and Brian F. Schaffner. 2017. "CCES Common Content, 2016." https://doi.org/10.7910/DVN/GDF6Z0.

Biggers, Daniel R., and Michael J. Hanmer. 2015. "Who Makes Voting Convenient? Explaining the Adoption of Early and No-Excuse Absentee Voting in the American States." *State Politics and Policy Quarterly* 15 (2): 192–210.

Braconnier, Céline, Jean-Yves Dormagen, and Vincent Pons. 2017. "Voter Registration Costs and Disenfranchisement: Experimental Evidence from France." *American Political Science Review* 111 (3): 584–604.

Burden, Barry C., David T. Canon, Kenneth R. Mayer, and Donald P. Moynihan. 2014. "Election Laws, Mobilization, and Turnout: The Unanticipated Consequences of Election Reform." *American Journal of Political Science* 58 (1): 95–109.

Cantoni, Enrico, and Vincent Pons. 2019. "Does Context Trump Individual Drivers of Voting Behavior? Evidence from U.S. Movers." Harvard Business School Working Paper 19-025.

Card, David, and Alan B. Krueger. 1994. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review* 84 (4): 772–93.

Card, David, and Enrico Moretti. 2007. "Does Voting Technology Affect Election Outcomes? Touch-Screen Voting and the 2004 Presidential Election." *Review of Economics and Statistics* 89 (4): 660–73.

Dell, Melissa. 2010. "The Persistent Effects of Peru's Mining Mita." *Econometrica* 78 (6): 1863–1903.

DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao. 2017. "Voting to Tell Others." *Review of Economic Studies* 84 (1): 143–81.

Dube, Arindrajit, T. William Lester, and Michael Reich. 2010. "Minimum Wage Effects across State Borders: Estimates Using Contiguous Counties." *Review of Economics and Statistics* 92 (4): 945–64.

Gentzkow, Matthew. 2006. "Television and Voter Turnout." *Quarterly Journal of Economics* 121 (3): 931–72.

Gentzkow, Matthew, Jesse M. Shapiro, and Michael Sinkinson. 2011. "The Effect of Newspaper Entry and Exit on Electoral Politics." *American Economic Review* 101 (7): 2980–3018.

Glynn, Adam N., and Konstantin Kashin. 2017. "Front-Door Difference-in-Differences Estimators." *American Journal of Political Science* 61 (4): 989–1002.

Gronke, Paul, Eva Galanes-Rosenbaum, and Peter A. Miller. 2007. "Early Voting and Turnout." *PS: Political Science and Politics* 40 (4): 639–45.

Gronke, Paul, and Peter Miller. 2012. "Voting by Mail and Turnout in Oregon: Revisiting Southwell and Burchett." *American Politics Research* 40 (6): 976–97.

Herron, Michael C., and Daniel A. Smith. 2012. "Souls to the Polls: Early Voting in Florida in the Shadow of House Bill 1355." *Election Law Journal* 11 (3): 331–47.

Herron, Michael C., and Daniel A. Smith. 2014. "Race, Party, and the Consequences of Restricting Early Voting in Florida in the 2012 General Election." *Political Research Quarterly* 67 (3): 646–65.

Kaltenthaler, Karl. 2010. "A Study of Early Voting in Ohio Elections." https://moritzlaw.osu.edu/electionlaw/litigation/documents/Ohio194.pdf.

Matthews, Tom J., and Brady E. Hamilton. 2009. *Delayed Childbearing: More Women Are Having Their First Child Later in Life*. Hyattsville, MD: National Center for Health Statistics (NCHS).

McDonald, Michael P. 2016. "A Brief History of Early Voting." *Huffington Post*, December 6. https://www.huffpost.com/entry/a-brief-history-of-early_b_12240120.

Naidu, Suresh. 2012. "Suffrage, Schooling, and Sorting in the Post-Bellum U.S. South." NBER Working Paper 18129.

Pons, Vincent. 2018. "Will a Five-Minute Discussion Change Your Mind? A Countrywide Experiment on Voter Choice in France." *American Economic Review* 108 (6): 1322–63.

Snyder, James M., and David Strömberg. 2010. "Press Coverage and Political Accountability." *Journal of Political Economy* 118 (2): 355–408.

Spenkuch, Jörg L., and David Toniatti. 2018. "Political Advertising and Election Outcomes." *Quarterly Journal of Economics* 133 (4): 1981–2036.

Strömberg, David. 2004. "Radio's Impact on Public Spending." *Quarterly Journal of Economics* 119 (1): 189–221.

This article has been cited by:

1. Enrico Cantoni. 2020. A Precinct Too Far: Turnout and Voting Costs. *American Economic Journal: Applied Economics* **12**:1, 61-85. [Abstract] [View PDF article] [PDF with links]