What If:

Estimating the Effects of Restrictive Voting Policies on Turnout in Ohio, Kansas, and Pennsylvania Using Synthetic Controls

Jackson Rudoff[1]

for Fulfillment of the M.A. in Quantitative Methods in the Social Sciences

Advised by Mike Z. He[2]

Spring 2023

[1] Graduate School of Arts and Sciences, Columbia University
[2] Department of Sociology, Columbia University

Table of Contents

Abstract

The study of American voter behavior has long centered on the factors affecting the decision to vote. This includes demographic factors like education, race, or partisanship. More recently, though, the discussion has revolved around *structural* factors, due to an uptick in restrictive voting policies in numerous states. These policies, in theory, raise the "cost" of voting, which could negatively impact turnout and in turn affect electoral competitiveness. However, findings on this matter remain mixed, and a key question remains unanswered: what has been the *exact* effect of restrictive voting policies on electoral behavior in the United States? This study employs the Synthetic Control Method (SCM) to measure the treatment effect of restrictive voting policies enacted from 2011-2012 on aggregate voter turnout in Ohio, Pennsylvania, and Kansas. Synthetic counterfactuals were constructed for each state using several relevant covariates for turnout, such as education, income, and non-white population, with a donor pool of 32 states. Findings from this study suggest a non-uniform treatment effect of restrictive voting policies, with Ohio being the only observed to have experienced a post-treatment turnout decline. Broadly, the results point toward the 2011-2012 *policy wave* as change point in American voter turnout across treated states. The completeness and precision of these findings additionally demonstrate the SCM framework as a suitable option for causally oriented research in comparative political science.

Introduction

Since the 1950s, American voting behavior has been a core topic within political science. As analytical frameworks and field survey methods have developed, so has a complex web of theory attempting to model people's decision to vote. A couple of patterns have been firmly established: first, that even amidst an apparent decline from the 1970s onward, voter turnout has remained relatively stable and general elections remain competitive; and second, that the personal "cost" of voting has been rising for the last decade. This latter element is the result of new voting laws being passed by state legislatures across the country; these policies have targeted many aspects of the voting process, ranging from changes to forms of acceptable ID at polling places to reductions in early and absentee voting. Since an initial "wave" of such policies implemented in 2011-2012, their adoption has only increased as unfounded fears of voter fraud have precipitated to the public and transformed these policies into a new partisan battle (Gronke et al. 2019).

As a result of these changes to the electoral environment, concepts like voter turnout, participation, and electoral competitiveness have returned to the center of academic discussion. Some authors, from basic models of post-reform elections, have claimed that new restrictions on voter identification have depressed turnout among minorities (Hajnal et al. 2017). Others, however, have questioned these findings due to measurement error and imprecise computational methods, positing that turnout or partisan dynamics may just be a product of local conditions (Grimmer et al. 2018; Walker, Herron, and Smith 2019). And perhaps most significantly, prior results from turnout research continues to be undermined by the error incurred from overreliance on surveys, such as the American National Election Studies (McDonald and Popkin 2001; Jackman and Spahn 2019). This lack of consensus bodes poorly for making sense of clear, potentially related trends, such as the conversion of Ohio from a coin-flip in general elections to a consistent Republican stronghold (Kaplan and Yuan 2020).

The state of this research area, then, is one of uncertainty, both methodological and substantively. Grimmer et al. (2018) and Highton (2017) have both attested that conventional approaches to studying voter ID laws, such as DiD regression using survey data, are plagued by specification and identification problems. For there to be advances in our understanding of these policies, there needs to be a modernization of both the data that scholars employ and the model architectures they consider. Otherwise, the cycle of erroneous results and corrective responses will continue, while our precise estimates of these policy effects only become more uncertain.

The purpose of this study is to address this split-task directly using the *synthetic control method* (Abadie et al. 2010; Abadie, Diamond, and Hainmueller 2015), an established method in causal inference. Aside from either reinforcing or invalidating the effects uncovered in previous treatments, the following study will function as a prototype for future designs addressing electoral participation. Focusing in on three states affected by post-2012 voting restrictions – Ohio, Pennsylvania, and Kansas – two key questions will have more precise answers. First, are *synthetic controls* suitable for the estimate of effects related to voter behavior? And second, what are the effects of restrictive voting policies on turnout, *and* are these effects variable between state-specific political contexts?

Literature Review

Voter Turnout: Why Does Anyone in the U.S. Vote?

At a basic level, the choice to vote in the U.S. can be considered a cost-analysis problem: do the costs of voting outweigh the cost of not voting? This simple calculation, however, has driven decades of debate because it superficially represents a paradox in the *rational choice* framework (Feddersen 2004; Aldrich 1993; Ferejohn and Fiorina 1974; Downs 1957). The actual utility of a single vote is extremely low, and voters will often incur far greater costs when turning out to vote (e.g., missing work, waiting in line, etc.). There is also strong empirical evidence that voters are strategic in both their decision to vote and candidate choice, despite the minimal effects from a strategy change (Feddersen 2004). Though there are numerous proposed solutions to this paradox, models supportive of "rational" voting behavior generally revolve around the notion that voters can be mobilized by factors external to their own self-interest. Edlin et al. (2007) find support that social (selfless) preferences take precedence for any given voter, and the decision to vote is thus "rational" for them because it benefits the group. Context for the election itself can affect how voters perceive the act as well. If the U.S. has been experiencing political unrest, voters may have less confidence in themselves, and therefore the value of their vote (Matsusaka 1995). Other have made note of the role of collective action, and the role that learning from past electoral outcomes plays in a voter's more immediate behavior (Kanazawa 1998). Generally, the theoretical side of "cost" estimation points toward voters as not necessarily irrational; instead, they can recognize sources of utility or value that are independent of themselves or derived from retroactive negative outcomes.

As the theoretical side of voter turnout has developed, empirical findings have unlocked a trove of variables that correlate with participation and underscore the socio-behavioral quality of cost estimation. Abramson and Aldrich (1982) put forth one of the earliest comprehensive studies of the decline in American voter participation, which had reached an apex in the late 1970s. Their analysis of SRC/CPS surveys from 1952-1980 found declines in strength of partisan identification and "external" political efficacy (belief in government) to be at the center of this trend. Timpone (1998) suggested that the trend may be more generally a consequence of the voting process in the United States. By pooling data from the 1980-1988 National Election Studies, he showed that dividing voters into three categories – non-registrants, non-voters, and voters – made previously uninteresting associations between turnout and race, gender, and social variables like relocation or marriage more salient.

An additional variable – education – has also been closely monitored in several studies since the beginning of the American turnout decline (see Wolfinger and Rosenstone 1980; Miller 1992; Sunshine Hillygus 2005; Tenn 2007). Burden (2009) contextualized this variable within "Brody's puzzle," another voter behavior paradox describing the increase in education not leading to an increase in turnout. Burden's analysis again deploys survey data from the NES, from 1950-2004, and builds several logit models comparing different levels of educational attainment (high school, some college, college degree) and U.S. turnout data. His findings showed that the positive relationship between education and turnout shifted dramatically in the 1980s, and that college education developed into a distinct predictor of voter behavior.

The noisiness and complexity of empirical turnout data clouds the external validity of these findings and points toward the salience of structural political variables. More recent research has thus explored turnout in the context of policies that reduce the net cost to voters. In a 2005 synthesis of several cross-sectional studies of mail/internet-voting, early voting, and absentee voting, Berinsky found that such electoral reforms alone fail to overcome the turnout deficit observed in lower SES communities. A similar study of aggregate and individual-level turnout data from the 2004 and 2008 presidential elections uncovered a similar effect (Burden et al. 2014). Regressing early voting and same-day registration against other core covariates of turnout showed same-day registration to be positively correlated with turnout, while early voting dampened the known positive effect of electoral "significance." Though these findings are difficult to reconcile, a general pattern has emerged, one indicative of the importance of electoral structures and policy, both in their direct effects on voting (reductions in vote cost) and indirect effects (reinforcement of any potential educational or SES biases).

One of the questions to be addressed here, then, is how exactly to capture the evident complexity of the decision to vote and how shifts in the political environment affect the apparent cost to voters. As such a nebulous and difficult to quantify variable, the correct empirical approach is less oriented around quantifying the exact *cost*, and instead measuring exactly how voters respond to the specific factors affecting their calculations. Moving forward, the central goal will be to capture the structural side of this vote cost equation.

Restrictive Voting Laws: Identifying the "Problem"

Developments in the voter turnout literature have increasingly focused on policy effects, particularly as states across the U.S. have adopted measures intended to both expand and restrict when and how people can vote. Studies of restrictive voter ID laws have dominated the literature, due mainly to their possible suppressive effects on turnout, particularly for ethnic minorities. The initial literature on the subject proved skeptical; in an early study on the potential effects of voter ID laws, Mycoff, Wagner, and Wilson (2009) implemented both bivariate and hierarchical models of voter ID laws and turnout for elections from 2000 to 2006, finding no significant positive or negative effect on turnout. But in the decade that followed, these laws became increasingly common and comprehensive, providing more robust and accurate data on their effects. Hajnal et al. (2017) catalyzed a renewed interest in the topic, using panel data from the Cooperative Congressional Election Study to model turnout as a product of voter ID strictness, controlling for other key demographic variables (e.g., education, class, etc.). They found that strict ID laws induced a significant reduction in both general election and primary election turnout for minority voters as compared to white voters. A follow-up analysis that combined extracts from 2008-2014 state and national surveys of voter populations provided further validity to these findings (Barreto et al. 2019). The authors found that white voters were significantly more likely to possess 'valid' ID, and that a disparate impact of voter ID was present in both the simplified and robust multivariate versions of their logit model. Perhaps most concerningly, there is evidence that enforcement of these laws is disproportionate: using a 2008 exit poll of 2,399 voters, Cobb (2012) found that the posterior probability of being asked for ID was significantly higher for non-White voters, even when controlling for precinct-level variables. A later field experiment reinforced these bureaucratic deficiencies; emails requesting information from election officials with Latino-sounding aliases received significantly fewer responses than those sent with White-sounding names (White, Nathan, and Faller 2015).

There is also evidence that the suppressive effects of state-level voting restrictions dampen partisan competition and tilt the balance in favor of Republicans. It has been established that Democrats benefit from higher voter turnout, with cross-sectional election data from 1948-2000 demonstrating how the party benefits from the volatility and incumbency effects of high-turnout elections (Hansford and Gomez 2010). With increased turnout generally comes increased competitiveness, a trend noted by Hicks et al. (2015) in their analysis of voter ID laws in Republican-led states. Through a pooled cross-sectional model of ID bills in state legislatures, they found a score of context-driven political factors affecting the implantation of voter ID laws, but most significantly that states with "competitive" elections were increasingly likely to adopt ID restrictions as the number of Republican legislators increased. These partisan-boosting effects even extend to restrictions on early voting. A panel analysis aggregating voter files from the 2008 and 2012 Florida general elections found that a disproportionate drop in early voting by ethnic minorities, workers, and Democrats was associated with a reduction in the length of the early voting period (Herron and Smith 2014). The sum of these findings suggests, broadly, that competitive electoral environments may decline in both in sheer turnout and overall competitiveness when conditioned by voting restrictions.

Methods in Case Studies of Elections

Most empirical elections research is based upon case studies, which themselves are typically constructed from panel surveys or samples of larger field surveys. The foundational studies in this area exemplify this, with their authors primarily addressing their questions through isolated regression models of particular election years (Durden and Gaynor 1987; Powell 1986; Shaffer 1981; Abramson and Aldrich 1982). In recent years, this approach has encountered greater scrutiny, largely because of internal disagreement over analytical robustness. Hajnal et al. (2017), one of the earliest papers to claim a negative effect of voter identification policies, has seen its data re-evaluated by other authors because of possible measurement and sampling errors (Grimmer et al. 2018). Though they do not entirely discount the possibility of a negative effect, Grimmer and his co-authors specifically identify several problems with the kinds of data used in recent studies of electoral policy effects. They make specific note, for example, that use of a national survey in state-level research is prone to measurement errors. It is for this reason that they specifically advocated that future research into American elections be grounded in true voter population data, rather than small samples of surveys with limited generalizability.

Beyond the survey-related issues identified in Grimmer et al. (2018), political science has generally struggled with a transition toward causal inference. This is not to say that a broader movement toward causal inference has not occurred, in fact quite the contrary. Grimmer (2012) identifies quite plainly that political scientists have become perhaps too focused on research with causality at the core, with the value of descriptive or validatory measurements being forgotten. This comes in combination with general flaws in the implementation and interpretation of causal models in political science. As more authors orient themselves around the measurement of an observed treatment effect, problems of poor identification or incomplete design arise, where authors fail to conduct the necessary groundwork in isolating the effect of interest (Keele 2015). Hajnal et al. (2017) serves as an illustrative example of this problem, as their observed effects were ultimately invalid due to errors in the coding of when certain states received their policy "treatment." This amounts to poor identification and is particularly problematic considering their specification of a differences-in-differences (DiD) design. Keele and Minozzi (2013) demonstrated in a study of election-day registration in Minnesota and Wisconsin that simply changing identification assumptions during the specification of a DiD yields varying results. What becomes evident is that, aside from its immense difficulty, causal identification and counterfactual estimation in political science are simultaneously highly valued and poorly implemented.

Consequently, newer research designs like the *synthetic control method* present as an attractive option in the research of elections and voting behavior (Abadie et al. 2010; Abadie, Diamond, and Hainmueller 2015; Xu 2017; Roesel 2017). Though other designs like DiD, regression discontinuity (RDD), and instrumental variables (IV) have attained greater popularity, the synthetic control method carries several advantages for executing causal research where randomization is impossible. The method's core design is relatively straightforward; using a region that received a specific policy as the "treated" unit, regions that did not receive the intervention are used to construct a "synthetic counterfactual" for the treated unit for all pre- and post-intervention observations. Synthetic control models are also capable of handling several pre-treatment covariates, an especially crucial feature in the context of political science research where selection bias often interferes with causal inference. Political scientists' interest in synthetic controls has certainly increased in conjunction with developments in causal estimation, best demonstrated by modernized architectures like *gsynth*, which incorporates interactive fixed-effects and a built-in cross-validation procedure (Xu 2017). As a result, several authors have both promoted and employed this design in explorations of the effects of changes to policy, political structures, and other contextual variables on elections and voter behavior (Riera 2020; Roesel 2017; Heersink, Peterson, and Jenkins 2017).

Methods and Data

Quantifying Voter Behavior

The outcomes of voter behavior take several forms. The act of voting involves several interrelated decisions, including but not limited to registering to vote, the way one votes (early, on election day, by mail, etc.), and whom they vote for. Many of these variables, however, suffer from severe measurement imprecision. Vote choice, both in terms of a specific candidate or party, can only be estimated through self-reporting by voters, which causes immediate problems in terms of non-response or misreporting error (Jackman and Spahn 2019).

To alleviate these concerns, the outcome variable in this study will be turnout, estimated using data from the US Elections Project (McDonald 2022). These turnout estimates were selected for several reasons. First, McDonald's data calculates turnout as a proportion of the voting-eligible population, which has been demonstrated to yield better aggregate estimates of shifts in turnout across election periods (McDonald and Popkin 2001). Second, these estimates are not derived from voters' self-reporting of their behavior, a shortcoming of previous studies (see Grimmer et al. 2018, in response to Hajnal, Lajevardi, and Nielson 2017). Lastly, turnout has been established as compatible with techniques in causal inference broadly and within research incorporating a synthetic control design (Highton 2017; Keele and Minozzi 2013; Xu 2017).

In deviation from prior research, this study will track turnout outcomes across both midterm and presidential elections (as opposed to merely the latter). Doing so may enable quantification of an effect that is exclusive to certain types of elections, perhaps revealing an imbalanced treatment effect. Assuming proper specification of the model, any potential deficiencies in this approach should be revealed prior to inference.

Treatment and Covariate Identification

A difficulty in estimating the effects of these policies is isolating a definitive "treatment," which is foundational to causal identification. This challenge originates from both the varying *types* of policy encountered by states, and the incongruity between receipt of treatment and measurement in the outcome variable. A policy affecting election procedure passed in 2011 will not truly condition the electoral environment until the next major election in 2012. Though this presents certain inferential challenges, there are a few approaches to overcoming this potential identification issue.

This study takes the approach of grouping the policies passed throughout 2011-2012 as a shared phenomenon, whereby the 2012 election represented the beginning of a targeted treatment across several states. Several sources, such as the Brennan Center for Justice, have identified this period as the turning point in modern US voting policy, whereby state legislatures undertook a collective effort to reverse course on the expansion of voting opportunities. Given that the policies passed during this period were broadly implemented with the same stated goal – improving election security – we can comfortably assume a shared class of policy treatment. This is not to imply that this approach is immune to identification issues or potential heterogeneity, but more careful consideration of these potential issues will be made in the **Discussion** section. Generally, case selection will proceed based first on if the state was included in the 2011-2012 *policy wave*, and secondarily on the type of intervention they received.

An equally important step in identification is covariate selection, particularly for a context-inclusive model like the synthetic control method. The lack of consensus regarding the exact covariates of turnout makes this a more challenging task, as much of the foundational research on voter turnout has failed to endure scrutiny (Kam and Palmer 2008; McDonald and Popkin 2001; Highton 2004). Selected covariates also need to be reflective of the patterns related to this policy and potential confounding state-level characteristics. If there were shifts in the population unrelated to the elections or voting policies themselves, failing to capture those shifts would preclude inference.

The set of relevant covariates and their basis in the literature are as follows:

1. *Non-White Population* (Herron and Smith 2014)

2. *Unemployment* (Burden and Wichowsky 2014)
3. *Income* (Filer, Kenny, and Morton 1993)
4. *College education* (Burden 2009; Sondheimer and Green 2010)
5. *State Partisan Control* (Hicks et al. 2015)

Aside from controlling for confounding, these inclusions will also yield evidence as to the association of these covariates with turnout in the context of these policy treatments.

Case Selection

Three states were selected from the 2011-2012 *policy wave* pool: Ohio, Pennsylvania, and Virginia. Each of these states represents a different *mode* of restriction, enabling comparisons between the effects of different policy and specific contexts.

Table 1: Restrictive Policies in Ohio, Kansas, and Pennsylvania

State	Early Voting	Non-Strict ID	Strict ID
Ohio	Yes	No	No
Pennsylvania	No	Yes*	Yes*
Virginia	No	No	Yes

Note: Pennsylvania's treatment was intended solely for first-time voters after 2012

The purpose of this multi-class comparison is to tease out structural, contextual, and political explanations for the differences in observed effects on turnout. Study of electoral turnout has frequently been clouded by uncertain and competing findings on relevant covariates. The surge of initial quantitative research on the broader American electorate in the 1980s and 1990s found few explanatory factors that were consistently associated with increased turnout. At that time, however, these factors were analyzed with a broader assumption of an otherwise unchanged participatory environment. But as authors began exploring the effects of the voting laws being implemented after 2008 – and especially from 2011-2012 - the context of this work became largely state-specific. This, in combination with issues in model design, pointed toward contradictory, and at times counterintuitive effects on turnout, wherein restrictive policies could suppress or promote turnout in elections.

These selected cases will enable a broader view of these effects with the ability to narrow in on *where* differences occur, and *why* this might be the case. By virtue of these states being prime swing states during the initial treatment period, we can capture an additional latent factor of inherent competitiveness and attention. Candidates, stakeholders, and media outlets present clear incentive to voters in swing states, pointing toward the potential for greater sensitivity to structural changes in participation. In general, these cases provide a framework for taking the variation in turnout responses as a given, with room to explore the implications of this non-uniformity more thoroughly.

Kansas: A Strict Voter ID Implementation

While many states subject to voting restrictions experience arduous paths to implementation, this was hardly the case in Kansas. In 2011, Governor Sam Brownback signed a bill that required all citizens to show photo identification in order to vote. The policy was a part of the state's larger effort to reinforce election "safety," a charge based largely on spurious claims of rampant voter fraud in recent elections. Unlike other states subjected to such policies (like Pennsylvania), there were few serious legal challenges to this provision, with the only major lawsuit being dismissed in 2014. Certain protections were provided for citizens lacking an acceptable form of ID; residents who lacked any acceptable form ID could sign an affidavit that rendered them eligible for a free ID from the state, and certain exemptions were made for the disabled, military veterans, and the elderly (Wendy R. Weiser and Lawrence Norden 2012).

Kansas's selection is the most straightforward of the three treated states. It stands as a "charter member" of the 2011-2012 restrictive policy wave, with one of the least embattled and most clearly defined strict voter ID laws. Prior research on Kansas' voter ID law has, however, revealed that communication of the policy varied between counties. Consequently, it may be incorrect to hypothesize a suppressive effect, as within the state itself there may have been differing levels of voter awareness (Bright and Lynch 2017). As a result, the aggregate effect being captured here needs to be evaluated with some caution, as there may be missing within-state granularity. However, Kansas still represents an excellent case in terms of exploring turnout outcomes at the state-level and is highly suitable for this comparative design.

Ohio: Reductions in Early Voting Accessibility

Ohio's history with voting restrictions is complex, as it was one of the earliest recipients of the "Voter ID" debate but a simultaneous adopter of early voting expansion. In 2005, Ohio implemented the "Golden Week," which allowed voters to cast absentee ballots up to 35 days before an election and to register *and* vote within a 7-day period. This came in response to a disastrous set of presidential elections, with Election Day voters experiencing long wait times, logistical mishaps, and other procedural roadblocks (Liptak 2004). The expectation was that expansions in both the early voting period and registration options would mitigate Election Day chaos and assure voters that participation would not yield frustration. In general, the Golden Week was well-received by the public and policy experts, the majority of whom readily welcomed improvements to voter accessibility and quality-of-life (Clark 2013).

These changes were, however, short-lived, as 2011 brought efforts to undo them. Aside from just repealing the Golden Week, Ohio Republicans additionally passed measures to reduce early voting opportunities overall. The first changes also seemed relatively minor, merely removing 3 days from the early voting period; a successful legal challenge eventually blunted these restrictions, though an overall contraction in the early voting period remained (Kaplan and Yuan 2020). Within just two years, Ohio's Secretary of State Jon Husted led a legislative movement to enact other restrictions, including the total removal of the *Golden Week*. Despite a protracted, back-and-forth legal battle, a 2016 Supreme Court decision allowed these restrictions to remain.

Ohio, as a case archetype, is a perfect example of how policies from the 2011-2012 wave "treat" states when they are implemented. From 2012 onward, Ohio voters were essentially at the mercy of court reporting to understand their options, as decisions on their rights were frequently made directly before an election (Getachew 2014). What can be safely said, however, is that from 2014 onward, Ohio voters were subject to a policy environment largely discouraging of early and absentee balloting. It became more difficult to understand exactly when one could vote or register, and if broad mobilization could be accommodated ahead of election day.

Pennsylvania: Voter ID Treatment with an Asterisk

Pennsylvania is the most "edge" of the cases selected for this study. Pennsylvania represented one of the earliest, and most aggressive adopters of new voter ID requirements. The bill passed by the state legislature required that *all* voters present a government-issued photo ID for their vote to be officially counted (Pennsylvania Department of State 2014). As a "soft" introduction, however, the government first implemented this enforcement *by request*, meaning that only some voters would be asked to present identification (which a voter could then decline). Immediately, this new system encountered issues, the most glaring of which was the "alternative identification" provision. Due to a large swathe of citizens lacking a driver's license or other accepted form of required identification, Pennsylvania claimed it would introduce an official ID through the Department of Elections (Bronner 2012). State officials later acknowledged, though, that this was too difficult for them to accomplish, amidst other critical infrastructure and communication problems preventing easy adoption of the new ID mandate.[3] The law was formally suspended (for known voters), before being entirely struck down by a state judge.

However, Pennsylvania remains a "treated" unit for three important reasons. First, the law remained in-place for first-time and mail-in voters during the post-treatment period. Second, confusion over the electoral code led to partial enforcement in elections even after the provision was suspended, with some poll workers and voters being unaware that the law was not enforceable (Connor 2016). Lastly, the state still serves as an informative archetype of "treatment" recipient, as it provides a window into the effects of a treatment rendered incomplete by successful counter-advocacy. This may enable more exact observation of the "mobilizing anger" described by Valentino et al. (2017).

Synthetic Controls: Outline, Rationale, Assumptions

This study employs a *synthetic control* design to estimate the effects of these policies on voter behavior (Abadie et al. 2010). For the purposes of this study, the policies introduced prior to the 2014 general election will be studied collectively, as they shared a common goal of

[3] These include, but are not limited to, a state policymaker admitting the law was part of an attempt to get a Republican governor elected, unclear requirements for acquiring the alternative form of ID, and an insufficient number of buildings to provide IDs to voters.

sharpening the security of elections (Herron and Smith 2014; Barreto et al. 2019). This logic is adapted from Abadie et al.'s (2010) estimation of the effect of tobacco control policies on cigarette sales in California.

The synthetic control method (SCM) purports to extract policy effects by imputing a "synthetic counterfactual" for a selected case, which is identified as the "treated" unit. Rather than exclude other regions from the study, the synthetic control method leverages demographically similar regions to reconstruct the treated state in the aggregate for each pre-treatment observation. To do this, available states are first split into two groups, a treated and untreated "donor" pool. These donors are then used to compute a set of weights that minimize the distance between a *synthetic control* vector and the vector of actual observations. Estimation of the treatment effect proceeds through direct comparison of the treatment unit and its synthetic counterfactual. Proper execution of the model additionally requires "placebo runs" with all available donors, which enables statistical inference by providing a baseline for the paths in the outcome variable for states that supposedly were untreated.

There are several advantages to this approach, many of which pertain to its extensions of conventional fixed-effect designs and built-in validation measures. Several authors have called for more robust models of restrictive voting policies, with comparative cross-state approaches being a particular deficiency (Grimmer et al. 2018). Differences-in-differences approaches are common, but with mixed results regarding both their identification procedures and robustness of model specification (Keel 2015; Erikson and Minnite 2009). SCM directly accommodates these shortcomings, as it extends the DiD framework to be inclusive of time-varying unobserved confoundings. This enables it to capture effects that may go beyond the usual temporal scope of a DiD model and provide more information about the influence of the included covariates. Using an SCM therefore accomplishes several goals at once; aside from providing an additional estimate of effects, it will serve as a validation check of prior claimed effects.

The model makes a few key assumptions that are important to address. First and foremost, SCM assumes “parallel paths” for the intervention and synthetic control in the absence of any treatment. This assumption is critical to inferences, as it makes the differences between the treated and non-treated synthetic unit meaningful and forms the basis of the statistical output. Secondly, the method carries the Stable Unit Treatment Value Assumption (SUTVA), which states that the treatment of the treated unit did not in some way affect the neighboring control units (Rubin 2005). Lastly, there is an assumption of appropriate covariate specification, meaning that an adequate amount of pre-treatment confounding variables has been specified. The deeper consequences of these assumptions on interpretation will be addressed in the concluding discussion of results.

Model Specification and Data

A standard synthetic control model takes the form:

$$Y_{it} = \alpha_t + \delta_t * ß_i + \lambda_t * X_i + \varepsilon it$$

where Yit is the intervention outcome for unit i at time t, α_t is a fixed effect at t describing shared trends across units, δ_t is the causal effect interacted with β_i, the binary indicator of a policy intervention, λ_t is a vector of factor-loadings[4] and X_i is a vector of pre-treatment variables in the treated unit. The weights – as previously described – are obtained through the function:

[4] Factor loadings capture a time-varied effect between the selected pre-treatment covariates and the outcome.

$$||X_1 – X_0 * W||^2$$

where the distance between a vector of pre-treatment variables (X_1) in the treated unit and a matrix of pre-treatment variables (X_0) in the control pool is minimized. This optimization process produces a vector of weights (W) for *j* units in the control pool, which can be used to impute a synthetic control for every observation of *i* at each time *t*.

So, for a given *treated* state in this study, the method converges on the weighted combination of variables from the untreated states with the smallest distance from the observed characteristics of the treated stated at each stage of the pre-treatment time series.

The importance of this optimization process is best demonstrated through an imputation of a heavily biases synthetic control, where only values for the outcome variable are supplied. If we run a de-contextualized simulation using Ohio, we obtain the following results:

Table 2: Loss in Simulated Model with Uncontrolled Bias

Unit	Treatment	Variable MSPE
Ohio	1	5.267

Note: Only mean turnout was specified for each unit

In this table, the *Variable MSPE* (mean squared prediction error) tells us the magnitude of error across the pre-intervention observations. MSPE is core to application of the synthetic control method, as it reflects the optimizer’s convergence on suitable predictions in a synthetic unit; higher prediction error in the outcome variable would thus indicate poor specification and fundamentally uninterpretable results. The high MPSE observed in **Table 2**, however, is not necessarily disqualifying, as there are routes for improving specification and overall precision. If the model is re-specified with covariates and lagged outcomes, there should be an observable improvement in the *Variable MSPE*, and thus the fit of the model. In addition to providing covariates, rather than supplying the outcome (turnout) as the mean of *t* observation periods, it will be specified as a lagged variable for final 8 pre-treatment periods. Upon making these improvements, the model optimizes much more effectively:

Table 3: Loss in Simulated Model with Controlled Bias

Unit	Treatment	Variable MSPE
Ohio	1	1.923

Note: Covariates supplied for education, income and unemployment.

Beyond the minimization of prediction loss, these changes enable better control for bias incurred by unobserved confounding, improving both identification and treatment effect estimation (Abadie et al. 2010).

It is worth noting that other synthetic control models have been developed since Abadie et al. (2010). These include *gsynth*, a generalized synthetic control approach with interactive fixed-effects, and *augsynth*, which debiases any imbalance in an SCM's covariates (Xu 2017; Ben-Michael, Feller, and Rothstein 2018). Both methods introduce certain advantages over the standard weighted-combination architecture, but also present with certain drawbacks. With *gsynth*, for example, the model introduces far greater complexity that makes both its explanation and interpretation more difficult. Additionally, the method can present with optimization issues with outcome variables like *electoral turnout* which fluctuates heavily between observations. *augsynth*, however, can be used effectively as a supplementary tool for statistical inference, as it can provide more precise, de-biased estimates of pre-treatment fit. This can be useful for validating the estimates in a standard SCM model.[5]

In this study, a standard synthetic control estimator was used, with all the necessary computations being carried out in R using *tidysynth* (Eric Dunford 2021). The package was also used to generate many of the plots and tables containing the model results, as it includes several helper functions to output relevant descriptive and inferential statistics. Due to its ease-of-use and readable coding style, *tidysynth* has become a preferred option among researchers employing a synthetic control design (Lang, Esbenshade, and Willer 2022).

The model included the following set of covariates, selected due to their established relevance to political turnout and trends in electoral politics:

Table 4: Covariates included in Synthetic Control Estimation

Variable	Metric	Source
Non-White Population	% of state population	NIH SEER Estimates
Income	Logged household income	Bureau of Economic Analysis
Unemployment	Unemployment rate for full-time workers	Bureau of Economic Analysis
Education	Percent of population with college degree	US Census Bureau
Partisan Control	Democratic control of state house and senate	Carl Klarner
Lagged Turnout	Turnout in the voting-eligible population	US Elections Project

Non-White population was included as a means of controlling for the ethnic makeup of states, which has been shown to factor into both the implementation and intended effect of restrictive voting policies. *Income* and *unemployment* were included as controls for state-level socioeconomic associated with turnout effects, particularly those that occurred both prior and during the economic recession. *Education* is one of the few factors consistently shown to be

[5] Lang, Esbenshade, and Willer (2022) used *augsynth* to verify their calculated unit weights for a study measuring the policy time series in terms of weeks. They included no covariates other than lagged vaccine uptick.

associated with turning out to vote, with a college education generally being correlated with a higher likelihood of voting. Lastly, *partisan control* provides the model with a measurement of Republican control, a recognized indicator of increased restrictions.

An equally important step in SCM specification is the selection of suitable donors, or units untreated by a policy intervention. In the context of this study, the primary distinction between a treated and untreated unit was whether their state legislature passed a restrictive voting policy during the 2011-2012 wave. 15 states were excluded based on this criterion. However, other states were excluded from the donor pool for either data missingness or systemic differences in the electoral environment. This left a donor pool of 32 states, a more than adequate sample of states for constructing a synthetic control. A complete list of donors and excluded *treated* states can be found in **Appendix A**.

Results

Tables 5, **6**, and **7** show the balance tables for the three *treated* states, which illustrate the difference between the real and synthetic values across the included covariates and lags for the outcome variable. Generally, the balance tables demonstrate accurate optimization of the model fit, apart from partisan control and college education. These large margins of error are not necessarily disqualifiers of the model, and instead reflect an incongruence of these values with changes in the outcome variable. Most critically, though, the synthetic controls were very well-matched along the outcome variable, indicating strong fit and useable estimates of the core treatment effect. The highest pre-intervention MSPE was observed for Ohio, at just over 2. This performance exceeded that of the prototypical case study outlined by Abadie et al. (2010) despite the heavy fluctuations in the outcome variable.

Table 5: Balance Table for Ohio

Variable	Ohio	Synthetic Ohio	Donor
College Degree	14.95	22.68	19.11
House Control	42.06	49.86	52.25
Senate Control	35.26	49.49	51.48
Log(Income)	10.36	10.46	10.39
Pop Non-White	13.04	13	15.16
Unemployment	6.13	5.91	5.62

Table 6: Balance Table for Pennsylvania

Variable	Penn.	Synthetic Penn.	Donor
College Degree	17.07	17.11	19.11
House Control	48.36	48.92	52.25
Sen Control	40.89	51.86	51.48
Log(Income)	10.45	10.39	10.39
Pop Non-White	12.51	19.34	15.16
Unemployment	5.7	5.77	5.62

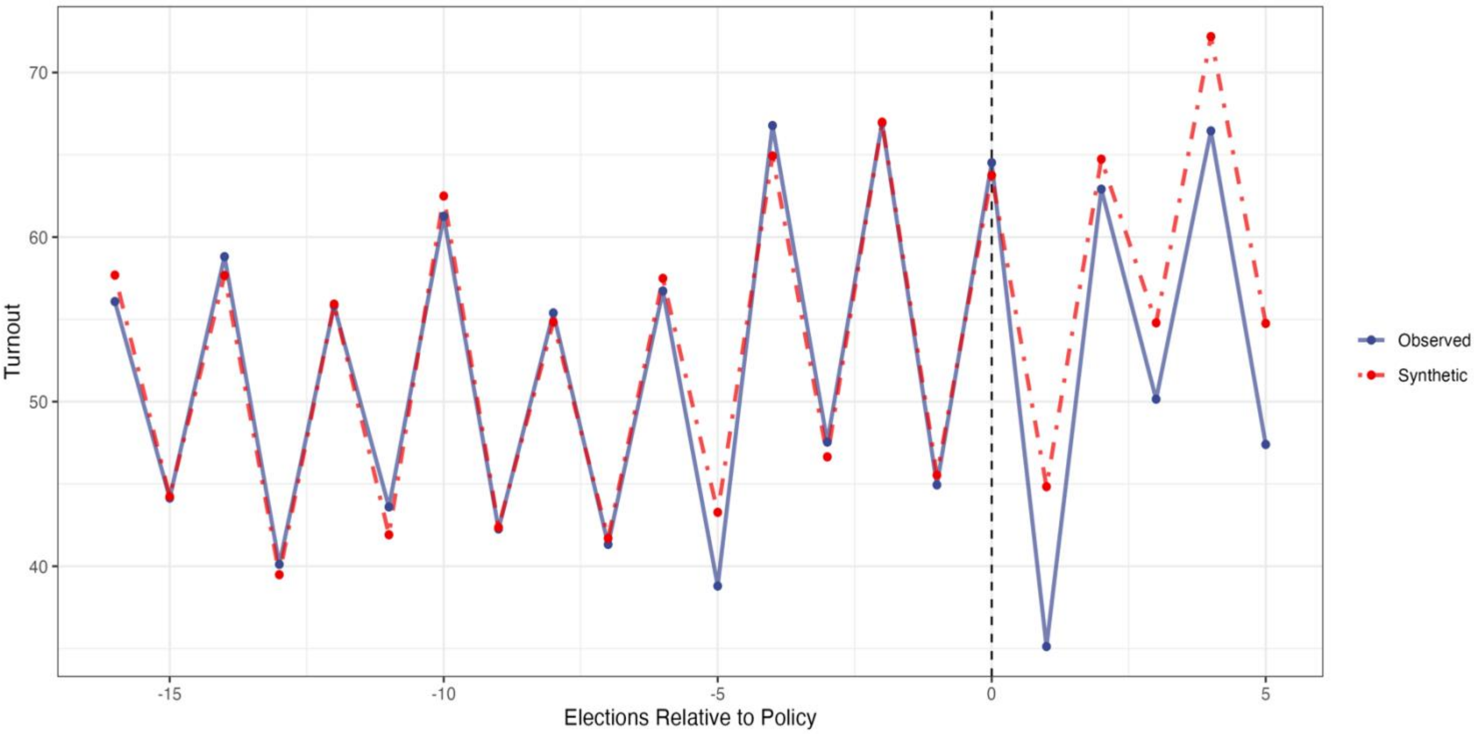
Table 7: Balance Table for Kansas

Variable	Kansas	Synthetic Kansas	Donor
College Degree	18.08	15.9	19.11
House Control	35.73	40.68	52.25
Senate Control	26.67	41.3	51.48
Log(Income)	10.39	10.32	10.39
Pop Non-White	9.07	13.73	15.16
Unemployment	4.9	5.17	5.62

It needs to be noted, however, that these balance tables reflect shortcomings of this synthetic control design, and that there should be necessary caution with interpreting the estimated treatment effect. The *Donor* sample shows that the distribution to be drawn from for covariate optimization was generally incompatible with the three intervention states; for this reason, it is unsurprising that the optimizer failed to arrive at uniformly precise linear combination across the set of relevant variables. Ohio's covariates were well-matched for *Income, Non-White population,* and *Unemployment*; synthetic Ohio, however, exceeded the means of its observed counterpart by at least 7 percentage points in both *Partisan Control* variables and in *Education*. In Pennsylvania, all covariates matched well, aside from *Senate Control* and *Non-White population*; Synthetic Pennsylvania was imputed with means higher by ~10 and ~7 percentage points respectively for these two variables. Lastly, synthetic Kansas was constructed to have covariate means that were all within 5 percentage points of the observed means, with the exception of *Senate Control*, which was poorly matched at a difference of ~13 percentage points. Aside from accentuating the specificity of these characteristics to each state, these differences indicate the potential for further model experimentation.

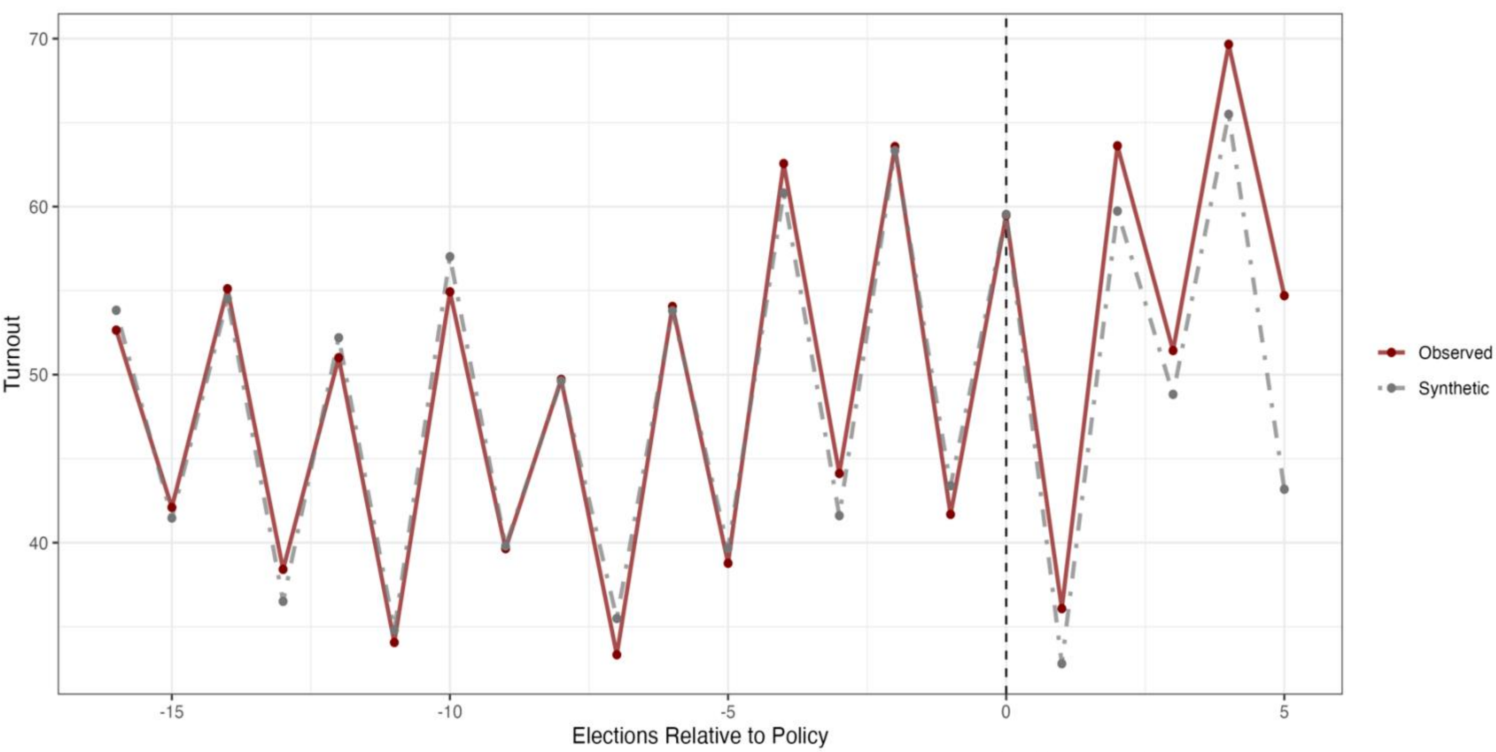
Figures 1, **2**, and **3** show a time series of turnout for both the observed and synthetic units of Ohio, Pennsylvania, and Kansas. In these plots, the "treatment" point is indicated by the dashed line through 2012, the designated cutoff for receiving an intervention. Each shows a high accuracy of the estimates for the synthetic counterfactual, despite the fluctuating path of turnout between midterm and presidential elections. Interpretation of these plots is intuitive, with the synthetic line representing the model's estimation of the turnout rate in each state in the absence of an intervention. Ohio's synthetic counterfactual, for example, is estimated to have a higher turnout rate for all post-treatment observations, reflective of a negative treatment effect induced by its early voting restrictions.

Figure 1: Synthetic Time Series of Turnout in Ohio



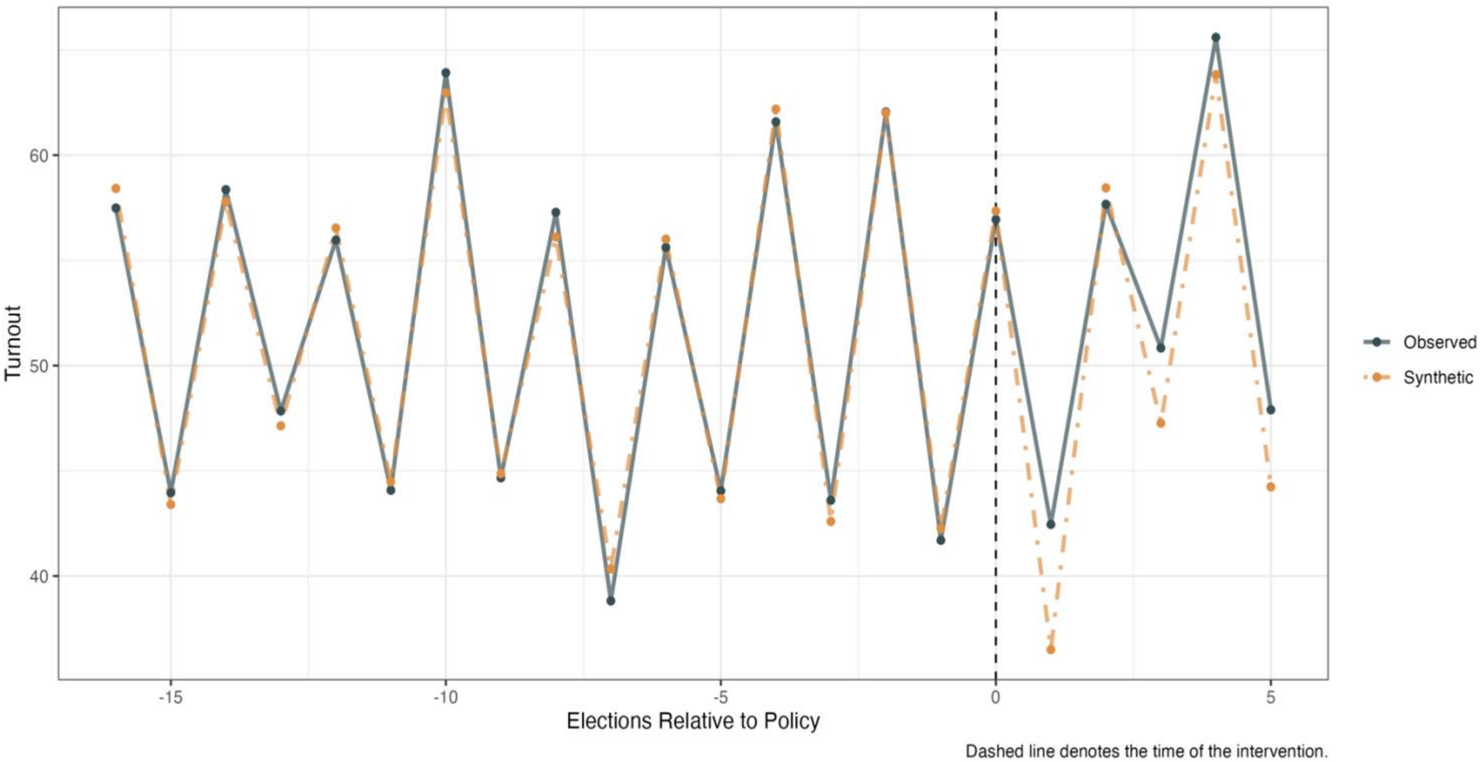
Dashed line denotes the time of the intervention.

Figure 2: Synthetic Time Series of Turnout in Pennsylvania



Dashed line denotes the time of the intervention.

Figure 3: Synthetic Time Series of Turnout in Kansas



There are multiple routes to statistical inference within an SCM. At the basic level, a treatment effect can be estimated using the difference in the outcome variable between the observed and synthetic unit after the policy intervention. Turnout is a slightly unusual variable in this regard, as it shifts levels significant between observations; however, it is still possible to quantify where these differences stand at the end of the post-intervention period. For the three treated units of Ohio, Pennsylvania, and Kansas, the final post-treatment differences are -7.4, 9.8, 3.5 in the states' respective rates of turnout. **Figures 4**, **5**, and **6** show these differences (located in **Appendix B**).

Evaluating the scale of these effects can be done through comparison with "placebo runs," wherein a synthetic counterfactual is imputed for all members of the donor pool. The pre/post MSPE ratio provides information of how large the post-intervention effect for the treated unit is, which can then be directly compared and standardized along with the placebo units. **Table 8** displays the relevant inferential statistics for the three treated states, including the MSPE ratio, rank amongst the placebo runs, and the p-value associated with the MSPE ratio.

Table 8: Inferential Statistics for Three Treated Units

Unit	Type	Pre-MSPE	Post-MSPE	Ratio	Rank	*p* value
Kansas	Treated	0.55	13.07	23.93	1	0.03
Ohio	Treated	2.05	41.23	20.08	1	0.03
Pennsylvania	Treated	1.78	35.06	19.7	1	0.03

Put simply, each of the post-treatment MSPEs were found to be significantly large relative to their placebo runs. Ohio's post-treatment MSPE ratio of 20.7 ($p = 0.03$) indicates a significant, negative treatment effect on voter turnout from its restrictions on early voting. Pennsylvania's post-treatment MSPE ratio was similarly large at 19.1 ($p = 0.03$), indicative of a significant positive treatment effect from its uncertain ID implementation. Finally, the post-intervention ratio for Kansas optimized at 23.9 ($p = 0.03$), yet again indicative of a significant positive treatment effect of voter ID requirements on turnout. Note that each treated unit also ranked 1st among the placebo runs. **Figures 7, 8,** and **9** plot the placebos for each computed synthetic counterfactual (located in the **Appendix C**).

Lastly, each synthetic counterfactual is made up of several unit weights, which reflect the importance of each donor to its construction. These weights are the initial output of the optimization process, indicating the combination of partial Donor state characteristics that best approximates the observed intervention unit. Evaluating these weights allows one to understand both which state and which covariates are most critical to the model; more thorough analysis of a single-state case may use variation of these weights, for example, as a validity check on the linear optimization procedure. Here, these values are used to convey how much each donor state contributed to the composition of each synthetic counterfactual, with **Tables 9, 10,** and **11** containing the respective unit weights for Ohio, Pennsylvania, and Kansas.

Table 9: Unit Weights for Synthetic Ohio

Unit	Weight	Unit	Weight
Alabama	0	Massachusetts	0
Arizona	0.1	Michigan	0.13
Arkansas	0	Minnesota	0
California	0	Mississippi	0
Colorado	0.25	Missouri	0
Connecticut	0	Montana	0
Delaware	0	Nevada	0
Idaho	0	New Jersey	0.24
Indiana	0	New Mexico	0.08
Kentucky	0	New York	0
Louisiana	0	North Carolina	0
Maine	0.1	North Dakota	0
Maryland	0	Oklahoma	0
Oregon	0.02	South Carolina	0
Utah	0	Vermont	0
Washington	0.07	Wyoming	0

Table 10: Unit Weights for Synthetic Penn.

Unit	Weight	Unit	Weight
Alabama	0	Minnesota	0
Arizona	0.09	Mississippi	0.12
Arkansas	0	Missouri	0.27
California	0	Montana	0
Colorado	0	Nevada	0.16
Connecticut	0	New Jersey	0
Delaware	0.35	New Mexico	0
Idaho	0	New York	0
Indiana	0.02	North Carolina	0
Kentucky	0	North Dakota	0
Louisiana	0	Oklahoma	0
Maine	0	Oregon	0
Maryland	0	South Carolina	0
Massachusetts	0	Utah	0
Michigan	0	Vermont	0
Washington	0	Wyoming	0

Table 11: Unit Weights for Synthetic Kansas

Unit	Weight	Unit	Weight
Alabama	0	Minnesota	0
Arizona	0.04	Mississippi	0
Arkansas	0	Missouri	0.19
California	0	Montana	0
Colorado	0	Nevada	0
Connecticut	0	New Jersey	0
Delaware	0	New Mexico	0
Idaho	0.22	New York	0
Indiana	0	North Carolina	0
Kentucky	0	North Dakota	0.04
Louisiana	0.18	Oklahoma	0.2
Maine	0	Oregon	0
Maryland	0	South Carolina	0
Massachusetts	0	Utah	0
Michigan	0	Vermont	0
Washington	0	Wyoming	0.12

Discussion

These models reinforce the inconsistent effects of electoral policy – and other structural factors - on voter behavior. Rather than a uniform directional effect, the consequences for turnout can be either negative or positive across post-intervention elections, depending on both the state and type of policy intervention it received. But the most immediate, general finding indicated by the significant results across all three models is that the *2011-2012 policy wave* represents a clear turning point in aggregate voter behavior from a pure turnout standpoint.

Both Kansas and Pennsylvania serve as illustrative case studies for the potential of a "restrictive" voting policy to induce a higher level of turnout. Given that both states were recipients of a voter ID policy, however, this finding was not unexpected. Prior research has indicated that certain policies can trigger greater participation as an emotive response to a perceived injustice (Valentino and Neuner 2017). The estimated treatment effects of voter ID policies in this study were both positive, and significant in relation to the estimated effects in the placebo units. While this certainly provides support for a mobilization response, it additionally serves as reinforcement of the difficulty in producing robust estimates of electoral behavior. Counterintuitive effects are common, making the larger goal of theory-building difficult and often superficially self-contradictory.

Especially in the case of Kansas – which was a clear-cut and uninterrupted policy implementation – it remains unclear if voters were simply more prepared for such a policy or experienced a renewed interest in electoral politics in response to unpopular policymaking. Prior research on "advertising effects" has shown that an abundance communication of policy changes can better prepare voters and catalyze increased turnout in a subsequent election, and Kansas voters were advertised to (with varying degrees) about the new laws (Hopkins et al. 2017; Bright and Lynch 2017). What may be captured in this estimated positive effect is an unintentional increase in overall voter preparedness and encouragement via communication of the new laws. Essentially, by introducing and enforcing new voter identification policies, Kansas voters broadly became more equipped and interested in their elections, depending on their given region's level of advertisement. It could also be that because the policies were generally less impeded there were greater efforts on the part of local and national advocacy groups to bring attention to their rights and the upcoming elections. Further research in this area would include greater controls for these communicative differences, or better isolate the presence of advocacy groups in each region. This would improve identification of how changes such factors may be masking or entirely counteracting any negative effects from policies targeting electoral accessibility.

In addition to communicative differences, the observed effects in Kansas and Pennsylvania reinforce a "mobilizing anger" effect of restrictive voting policies. Prior research has demonstrated that partisanship factors into voters' perception of new laws (Gronke et al. 2019). This suggests that voters will respond differently to these policies depending on their own partisanship. Combined with the findings of this study, as well as those of Valentino and Neuner (2017), there is strong evidence that these positive treatment effects are, at least partially, a product of voters turning out in response to a perceived injustice. Instead of reducing voters' willingness, the new policies catalyzed a renewed interest in elections, as the restrictions represented a barrier that could be removed through changes in the elites. The results of succeeding elections in these states lends further support to this notion; though the state legislatures in both states remain majority Republican, both states have experienced Democratic gains since the 2011-2012 treatment period. More definitive evidence of this effect could be accomplished through an evaluation of changes to partisan vote shares. Combined with the theorized role of communications, these positive treatment effects broadly support a level of resilience in American voters, wherein an enacted increase in the requirements or barriers to vote can be overridden by collective interest.

These enacted costs, however, may not be equal in their effects, as indicated by the results from Ohio. Unlike the other two cases, the core restrictive policy in Ohio affected early voting opportunities. Early voting – particularly the Golden Week – were significant reductions in the personal cost of turning out to vote, as they mitigated the severe wait-time and logistical issues hampering election day. In turn, reductions to these expanded opportunities netted a negative response from voters, generally depressing the rate of turnout in all post-intervention elections. Unlike voter ID requirements – which affect whether someone's vote will be counted – changes to the voting period reduce the actual *opportunities* to vote. The increased level of turnout in synthetic Ohio suggests that the elimination of these opportunities outweighs or overrides any potential counter-mobilization effect. It also speaks again to voting models being inclusive of personal costs and circumstances, where the elimination of certain conveniences dissuades citizens from turning out because it is no longer worth the stress or has become incompatible with their schedule (Brady and Mcnulty 2011; Hassell and Settle 2017; Citrin, Green, and Levy 2014). Altogether, this effect serves as new and interesting evidence of the importance of expanding voting options to encourage voter participation, and the complexity underlying the costs that voters experience.

More broadly, these results accomplished the goal of reinforcing the potential of SCM in comparative political science research. Several scholars have identified a need for political science to move beyond conventional, pure-regression research designs, particularly in the context of complex constructs like elections. Though not the end-all-be-all, SCM presents as a plausible causal approach that produces valid estimates of a treatment effect for a particular policy, rather than the estimated association between changes in particular variables and voter turnout. All three models employed here successfully optimized and produced results in-line with prior theories and observations of voter turnout. However, their outputs have quantified with greater specificity the exact post-intervention effects of these policies and validate previous correlational findings, while also accounting for greater socio-political and temporal context. In sum, the successes observed here affirm SCM as a potential means of moving beyond the standard hypothesis-testing framework. With how difficult identification has proven in the context of voter turnout – despite several decades of attempts – the feasibility of this research design is highly promising.

However, there must still be caution in the implementation of SCM, as the potential for biased estimates and uncontrolled confounding remains. In the case of these models, the optimizer failed to converge for a few of the predictors, such as college-educated population and partisan control. Part of this can be attributed to limitations in the data, as these were rates that varied widely, and perhaps generally carried a minimal association with voter turnout. But as has been discussed, there may be nuances to each state context that make this combined-comparative implementation somewhat incomplete in terms of effect identification. Kansas, a state that flipped governor's parties during the post-treatment period, may be lacking a covariate that isolates the state's specific reaction to candidates that represented a stark difference from pre-treatment candidates. Essentially, this model may be capturing a turnout increase inspired by Kansas voters' renewed interest in specific gubernatorial or presidential candidates, rather than an interest catalyzed by a policy perceived as injustice. Changes to Kansas' major candidates may also affect the intervention's implementation itself; Kansas Democratic candidates in 2018, for example, made suppression a core aspect of their campaigning in the final weeks of the election (Smith 2018). Separate and more specific analysis would be required to determine whether this confounds results, however, and falls outside the scope of this study.

In all, these shortcomings are indicative of the design's fundamental limitations, wherein a strong overall fit is obtained despite flaws in covariate specification. A core assumption of SCM – shared with its cousin the DiD regression – is "parallel trends," which can be demonstrated through appropriate matching along the included covariates. In the case of this study, there was inconsistency in how well the included covariates matched between synthetic imputations. So, despite a strong fit in the outcome, the characteristics of the counterfactual were not as precisely controlled for, which adds uncertainty and the potential for uncontained bias in the outcome. It is possible that these errors are also reflective of algorithmic limitations, as other similar implementations have generally found success across a broad suite of demographic covariates (Abadie et al. 2010). In any case, these control variable challenges present an interesting theoretical quandary; does one leave out a covariate that fails to precisely match and assume it will be controlled as unobserved confounding, or include it with several cautions during statistical inference? A more thorough implementation of the methods used in this study would experiment with either different covariates or evaluate the consistency of the estimated treatment effects across shifts in the included covariates.

Additionally, there remains the overarching limitation of the definitive "starting point" for treatment in the context of these policies. As elections only occur every two years, it can be difficult to pinpoint the exact moment where a state is considered "treated." Furthermore, the policy environment did experience shifts in states beyond the initial 2011-2012 wave, which may have caused a certain level of interference.[6] While neither of these issues appear to have impacted the results (given the consistency across placebo runs), future SCM designs incorporating elections and fluctuating policy implementations may need to weigh these considerations further.

Conclusion

The findings presented here are important for two distinct reasons: one, they refine our understanding of the effect of restrictive voting policies on voter behavior; and two, they demonstrate the feasibility of *synthetic controls* in elections research.

Three case studies – Ohio, Kansas, and Pennsylvania - were each evaluated as having observed effects aligning with prior research, yet with more robust estimates of what those effects are. Though most scholars have theorized and found (with mixed reception) that voting restrictions inhibit turnout, this does not always appear to be the case. Both Kansas and Pennsylvania were found to have experienced significant *positive* treatment effects from their respective voter ID laws, although their level and form of treatment varied. Ohio contrarily experienced a *negative* treatment effect from reductions to its early voting period. These findings are not conflicting, in fact quite the opposite. Ohio's restrictions to early voting diminished the available opportunities to vote, which would explain why subsequent elections saw declines in participation to pre-Golden Week levels. What is important to consider in future research, then, is not just whether a policy incurs a greater personal "cost" for voters, but also *how* that cost has been increased. This would likely yield more productive results regarding which policies pose a

---

[6] See Keele 2015 for a more thorough explanation of competing or inconsistent interventions and their implications for causal identification.

greater suppressive threat to electoral participation and those which may be irrelevant in the face of counter-mobilization.

The feasibility of the synthetic control method has also been demonstrated. Though not entirely unknown, it remains an under-utilized option in attempts at causal inference in political science. The findings presented here show that, beyond better controls for unobserved confounding and biased covariates, the synthetic control method converges on valid effect estimates amidst the complexity of political systems. There remain questions related to identification, given the relative imprecision and inconsistency of how restrictive voting policies have been implemented. But these considerations are no more disqualifying than the identification problems that other causal designs have encountered, perhaps speaking instead to the necessity of further experimentation with this method for political questions. Calls have and are still being made for new approaches to the evaluation of voting restrictions and their measurable consequences for political participation. We may be able to answer them with further use and consequent refinement of synthetic controls in comparative political science.

Bibliography

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105 (490): 493–505. https://doi.org/10.1198/jasa.2009.ap08746.

———. 2015. "Comparative Politics and the Synthetic Control Method: Comparative Politics and the Synthetic Control Method." *American Journal of Political Science* 59 (2): 495–510. https://doi.org/10.1111/ajps.12116.

Abramson, Paul R., and John H. Aldrich. 1982. "The Decline of Electoral Participation in America." *American Political Science Review* 76 (3): 502–21. https://doi.org/10.2307/1963728.

Alboukadel Kassambara. 2023. "Ggpubr: 'ggplot2' Based Publication Ready Plots."

Aldrich, John H. 1993. "Rational Choice and Turnout." *American Journal of Political Science* 37 (1): 246. https://doi.org/10.2307/2111531.

Barber, Michael, and Adam M. Dynes. 2023. "City-State Ideological Incongruence and Municipal Preemption." *American Journal of Political Science* 67 (1): 119–36. https://doi.org/10.1111/ajps.12655.

Barreto, Matt A., Stephen Nuño, Gabriel R. Sanchez, and Hannah L. Walker. 2019. "The Racial Implications of Voter Identification Laws in America." *American Politics Research* 47 (2): 238–49. https://doi.org/10.1177/1532673X18810012.

Ben-Michael, Eli, Avi Feller, and Jesse Rothstein. 2018. "The Augmented Synthetic Control Method." https://doi.org/10.48550/ARXIV.1811.04170.

Berinsky, Adam J. 2005. "The Perverse Consequences of Electoral Reform in the United States." *American Politics Research* 33 (4): 471–91. https://doi.org/10.1177/1532673X04269419.

Brady, Henry E., and John E. Mcnulty. 2011. "Turning Out to Vote: The Costs of Finding and Getting to the Polling Place." *American Political Science Review* 105 (1): 115–34. https://doi.org/10.1017/S0003055410000596.

Bright, Chelsie L. M., and Michael S. Lynch. 2017. “Kansas Voter ID Laws: Advertising and Its Effects on Turnout.” *Political Research Quarterly* 70 (2): 340–47. https://doi.org/10.1177/1065912917691638.

Bronner, Ethan. 2012. “Voter ID Rules Fail Court Tests Across Country.” *The New York Times*, October 2, 2012, sec. U.S. https://www.nytimes.com/2012/10/03/us/pennsylvania-judge-delays-implementation-of-voter-id-law.html.

Burden, Barry C. 2009. “The Dynamic Effects of Education on Voter Turnout.” *Electoral Studies* 28 (4): 540–49. https://doi.org/10.1016/j.electstud.2009.05.027.

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: ELECTION LAWS, MOBILIZATION, AND TURNOUT.” *American Journal of Political Science* 58 (1): 95–109. https://doi.org/10.1111/ajps.12063.

Burden, Barry C., and Amber Wichowsky. 2014. “Economic Discontent as a Mobilizer: Unemployment and Voter Turnout.” *The Journal of Politics* 76 (4): 887–98. https://doi.org/10.1017/S0022381614000437.

Citrin, Jack, Donald P. Green, and Morris Levy. 2014. “The Effects of Voter ID Notification on Voter Turnout: Results from a Large-Scale Field Experiment.” *Election Law Journal: Rules, Politics, and Policy* 13 (2): 228–42. https://doi.org/10.1089/elj.2013.0209.

Cobb, Rachael V. 2012. “Can Voter ID Laws Be Administered in a Race-Neutral Manner? Evidence from the City of Boston in 2008.” *Quarterly Journal of Political Science* 7 (1): 1–33. https://doi.org/10.1561/100.00010098.

Connor, Tracy. 2016. “Voters Wrongly Asked for ID at the Polls in Pennsylvania,” November 8, 2016. https://www.nbcnews.com/card/[object Object].

Denora Getachew. 2014. “Voting 2014: Stories from Ohio.” Brennan Center for Justice. https://www.brennancenter.org/our-work/research-reports/voting-2014-stories-ohio.

Downs, Anthony. 1957. "An Economic Theory of Political Action in a Democracy." *Journal of Political Economy* 65 (2): 135–50. https://doi.org/10.1086/257897.

Durden, Garey C., and Patricia Gaynor. 1987. "The Rational Behavior Theory of Voting Participation: Evidence from the 1970 and 1982 Elections." *Public Choice* 53 (3): 231–42. https://doi.org/10.1007/BF00127348.

Edlin, Aaron, Andrew Gelman, and Noah Kaplan. 2007. "Voting as a Rational Choice: Why and How People Vote To Improve the Well-Being of Others." *Rationality and Society* 19 (3): 293–314. https://doi.org/10.1177/1043463107077384.

Eric Dunford. 2021. "Tidysynth."

Erikson, Robert S., and Lorraine C. Minnite. 2009. "Modeling Problems in the Voter Identification—Voter Turnout Debate." *Election Law Journal: Rules, Politics, and Policy* 8 (2): 85–101. https://doi.org/10.1089/elj.2008.0017.

Feddersen, Timothy J. 2004. "Rational Choice Theory and the Paradox of Not Voting." *Journal of Economic Perspectives* 18 (1): 99–112. https://doi.org/10.1257/089533004773563458.

Ferejohn, John A., and Morris P. Fiorina. 1974. "The Paradox of Not Voting: A Decision Theoretic Analysis." *American Political Science Review* 68 (2): 525–36. https://doi.org/10.2307/1959502.

Filer, John E., Lawrence W. Kenny, and Rebecca B. Morton. 1993. "Redistribution, Income, and Voting." *American Journal of Political Science* 37 (1): 63. https://doi.org/10.2307/2111524.

Franko, William W., Nathan J. Kelly, and Christopher Witko. 2016. "Class Bias in Voter Turnout, Representation, and Income Inequality." *Perspectives on Politics* 14 (2): 351–68. https://doi.org/10.1017/S1537592716000062.

Grimmer, Justin, Eitan Hersh, Marc Meredith, Jonathan Mummolo, and Clayton Nall. 2018. "Obstacles to Estimating Voter ID Laws' Effect on Turnout." *The Journal of Politics* 80 (3): 1045–51. https://doi.org/10.1086/696618.

Gronke, Paul, William D. Hicks, Seth C. McKee, Charles Stewart, and James Dunham. 2019. "Voter ID Laws: A View from the Public*: Voter ID Laws: A View from the Public." *Social Science Quarterly* 100 (1): 215–32. https://doi.org/10.1111/ssqu.12541.

Hajnal, Zoltan, Nazita Lajevardi, and Lindsay Nielson. 2017. "Voter Identification Laws and the Suppression of Minority Votes." *The Journal of Politics* 79 (2): 363–79. https://doi.org/10.1086/688343.

Hansford, Thomas G., and Brad T. Gomez. 2010. "Estimating the Electoral Effects of Voter Turnout." *American Political Science Review* 104 (2): 268–88. https://doi.org/10.1017/S0003055410000109.

Hassell, Hans J. G., and Jaime E. Settle. 2017. "The Differential Effects of Stress on Voter Turnout: Differential Effects of Stress on Voter Turnout." *Political Psychology* 38 (3): 533–50. https://doi.org/10.1111/pops.12344.

Heersink, Boris, Brenton D. Peterson, and Jeffery A. Jenkins. 2017. "Disasters and Elections: Estimating the Net Effect of Damage and Relief in Historical Perspective." *Political Analysis* 25 (2): 260–68. https://doi.org/10.1017/pan.2017.7.

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. https://doi.org/10.1177/1065912914524831.

Hicks, William D., Seth C. McKee, Mitchell D. Sellers, and Daniel A. Smith. 2015. "A Principle or a Strategy? Voter Identification Laws and Partisan Competition in the American States." *Political Research Quarterly* 68 (1): 18–33. https://doi.org/10.1177/1065912914554039.

Highton, Benjamin. 2017. "Voter Identification Laws and Turnout in the United States." *Annual Review of Political Science* 20 (1): 149–67. https://doi.org/10.1146/annurev-polisci-051215-022822.

Hill, Seth J. 2017. "Learning Together Slowly: Bayesian Learning about Political Facts." *The Journal of Politics* 79 (4): 1403–18. https://doi.org/10.1086/692739.

Hopkins, Daniel J., Marc Meredith, Michael Morse, Sarah Smith, and Jesse Yoder. 2017. "Voting But for the Law: Evidence from Virginia on Photo Identification Requirements." *Journal of Empirical Legal Studies* 14 (1): 79–128. https://doi.org/10.1111/jels.12142.

Jackman, Simon, and Bradley Spahn. 2019. "Why Does the American National Election Study Overestimate Voter Turnout?" *Political Analysis* 27 (2): 193–207. https://doi.org/10.1017/pan.2018.36.

Jennifer C. Clark. 2013. "Early Voting in Ohio: If It Ain't Broke, Don't Fix It." *Brennan Center for Justice*, December 3, 2013. https://www.brennancenter.org/our-work/analysis-opinion/early-voting-ohio-if-it-aint-broke-dont-fix-it.

Kam, Cindy D., and Carl L. Palmer. 2008. "Reconsidering the Effects of Education on Political Participation." *The Journal of Politics* 70 (3): 612–31. https://doi.org/10.1017/S0022381608080651.

Kanazawa, Satoshi. 1998. "A Possible Solution to the Paradox of Voter Turnout." *The Journal of Politics* 60 (4): 974–95. https://doi.org/10.2307/2647727.

Kaplan, Ethan, and Haishan Yuan. 2020. "Early Voting Laws, Voter Turnout, and Partisan Vote Composition: Evidence from Ohio." *American Economic Journal: Applied Economics* 12 (1): 32–60. https://doi.org/10.1257/app.20180192.

Keele, Luke. 2015. "The Statistics of Causal Inference: A View from Political Methodology." *Political Analysis* 23 (3): 313–35. https://doi.org/10.1093/pan/mpv007.

Keele, Luke, and William Minozzi. 2013. "How Much Is Minnesota Like Wisconsin? Assumptions and Counterfactuals in Causal Inference with Observational Data." *Political Analysis* 21 (2): 193–216. https://doi.org/10.1093/pan/mps041.

Keele, Luke, Randolph T. Stevenson, and Felix Elwert. 2020. "The Causal Interpretation of Estimated Associations in Regression Models." *Political Science Research and Methods* 8 (1): 1–13. https://doi.org/10.1017/psrm.2019.31.

Klarner, Carl. 2013. “State Partisan Balance Data, 1937 - 2011.” Harvard Dataverse. https://doi.org/10.7910/DVN/LZHMG3.

Lang, David, Lief Esbenshade, and Robb Willer. 2022. “Did Ohio’s Vaccine Lottery Increase Vaccination Rates? A Pre-Registered, Synthetic Control Study.” *Journal of Experimental Political Science*, February, 1–19. https://doi.org/10.1017/XPS.2021.32.

Leighley, Jan E., and Arnold Vedlitz. 1999. “Race, Ethnicity, and Political Participation: Competing Models and Contrasting Explanations.” *The Journal of Politics* 61 (4): 1092–1114. https://doi.org/10.2307/2647555.

Levine, David K., and Thomas R. Palfrey. 2007. “The Paradox of Voter Participation? A Laboratory Study.” *American Political Science Review* 101 (1): 143–58. https://doi.org/10.1017/S0003055407070013.

Liptak, Adam. 2004. “Voting Problems in Ohio Set Off an Alarm.” *The New York Times*, November 7, 2004, sec. U.S. https://www.nytimes.com/2004/11/07/politics/campaign/voting-problems-in-ohio-set-off-an-alarm.html.

Matsusaka, John G. 1995. “Explaining Voter Turnout Patterns: An Information Theory.” *Public Choice* 84 (1–2): 91–117. https://doi.org/10.1007/BF01047803.

McDonald, Michael. 2022. “Voter Turnout Data 1996-2020.” https://www.electproject.org/election-data/voter-turnout-data.

McDonald, Michael P., and Samuel L. Popkin. 2001. “The Myth of the Vanishing Voter.” *American Political Science Review* 95 (4): 963–74. https://doi.org/10.1017/S0003055400400134.

Miller, Warren E. 1992. “The Puzzle Transformed: Explaining Declining Turnout.” *Political Behavior* 14 (1): 1–43. https://doi.org/10.1007/BF00993507.

Mycoff, Jason D., Michael W. Wagner, and David C. Wilson. 2009. “The Empirical Effects of Voter-ID Laws: Present or Absent?” *PS: Political Science & Politics* 42 (01): 121–26. https://doi.org/10.1017/S1049096509090301.

Norris, Pippa. 2019. “Do Perceptions of Electoral Malpractice Undermine Democratic Satisfaction? The US in Comparative Perspective.” *International Political Science Review* 40 (1): 5–22. https://doi.org/10.1177/0192512118806783.

Palfrey, Thomas R., and Howard Rosenthal. 1985. “Voter Participation and Strategic Uncertainty.” *American Political Science Review* 79 (1): 62–78. https://doi.org/10.2307/1956119.

Pennsylvania Department of State. 2014. “Voter Identification Requirements for Voting.”

Powell, G. Bingham. 1986. “American Voter Turnout in Comparative Perspective.” *American Political Science Review* 80 (1): 17–43. https://doi.org/10.2307/1957082.

Riera, Pedro. 2020. “Voting after the Change: A Natural Experiment on the Effect of Electoral Reform on Party System Fragmentation.” *International Political Science Review* 41 (2): 271–86. https://doi.org/10.1177/0192512118822891.

Roesel, Felix. 2017. “Do Mergers of Large Local Governments Reduce Expenditures? – Evidence from Germany Using the Synthetic Control Method.” *European Journal of Political Economy* 50 (December): 22–36. https://doi.org/10.1016/j.ejpoleco.2017.10.002.

Rubin, Donald B. 2005. “Causal Inference Using Potential Outcomes: Design, Modeling, Decisions.” *Journal of the American Statistical Association* 100 (469): 322–31. https://doi.org/10.1198/016214504000001880.

Smith, Mitch. 2018. “To Cast Their Ballots, These Voters Will Have to Get Out of Dodge.” *The New York Times*, October 27, 2018, sec. U.S. https://www.nytimes.com/2018/10/26/us/dodge-city-kansas-voting.html.

Sondheimer, Rachel Milstein, and Donald P. Green. 2010. “Using Experiments to Estimate the Effects of Education on Voter Turnout.” *American Journal of Political Science* 54 (1): 174–89. https://doi.org/10.1111/j.1540-5907.2009.00425.x.

Sunshine Hillygus, D. 2005. “The MISSING LINK: Exploring the Relationship Between Higher Education and Political Engagement.” *Political Behavior* 27 (1): 25–47. https://doi.org/10.1007/s11109-005-3075-8.

Tenn, Steven. 2007. "The Effect of Education on Voter Turnout." *Political Analysis* 15 (4): 446–64. https://doi.org/10.1093/pan/mpm012.

Timpone, Richard J. 1998. "Structure, Behavior, and Voter Turnout in the United States." *American Political Science Review* 92 (1): 145–58. https://doi.org/10.2307/2585934.

U.S. Bureau of Economic Analysis. 2023. "Interactive Data Tables." 2023. https://www.bea.gov/data.

Valentino, Nicholas A., and Fabian G. Neuner. 2017. "Why the Sky Didn't Fall: Mobilizing Anger in Reaction to Voter ID Laws: Mobilizing Anger in Reaction to Voter ID Laws." *Political Psychology* 38 (2): 331–50. https://doi.org/10.1111/pops.12332.

Walker, Hannah L., Michael C. Herron, and Daniel A. Smith. 2019. "Early Voting Changes and Voter Turnout: North Carolina in the 2016 General Election." *Political Behavior* 41 (4): 841–69. https://doi.org/10.1007/s11109-018-9473-5.

Wendy R. Weiser and Lawrence Norden. 2012. "Voting Law Changes in 2012." Brennan Center for Justice.

White, Ariel R., Noah L. Nathan, and Julie K. Faller. 2015. "What Do I Need to Vote? Bureaucratic Discretion and Discrimination by Local Election Officials." *American Political Science Review* 109 (1): 129–42. https://doi.org/10.1017/S0003055414000562.

Wolfinger, Raymond E., and Steven J. Rosenstone. 1980. *Who Votes?* A Yale Fastback. New Haven: Yale University Press.

Xu, Yiqing. 2017. "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models." *Political Analysis* 25 (1): 57–76. https://doi.org/10.1017/pan.2016.2.

Xu, Yiqing, and Licheng Liu. 2022. "Gsynth: Generalized Synthetic Control Method." https://yiqingxu.org/packages/gsynth/.

Appendix A: Donor States Used for Computation of Synthetic Controls

Table A: Donor and Excluded States

Donors	**Excluded/Treated**
Alabama	Alaska
Arizona	Florida
Arkansas	Georgia
California	Hawaii
Colorado	Illinois
Connecticut	Iowa
Delaware	Kansas
Idaho	Nebraska
Indiana	New Hampshire
Kentucky	Ohio
Louisiana	Pennsylvania
Maine	Rhode Island
Maryland	South Dakota
Massachusetts	Tennessee
Michigan	Texas
Minnesota	Virginia
Mississippi	West Virginia
Missouri	Wisconsin
Montana	
Nevada	
New Jersey	
New Mexico	
New York	
North Carolina	
North Dakota	
Oklahoma	
Oregon	
South Carolina	
Utah	
Vermont	
Washington	
Wyoming	

Appendix B: Difference Plots for Synthetic Control Runs

Figure 4: Outcome Difference Between Synthetic and Observed Ohio

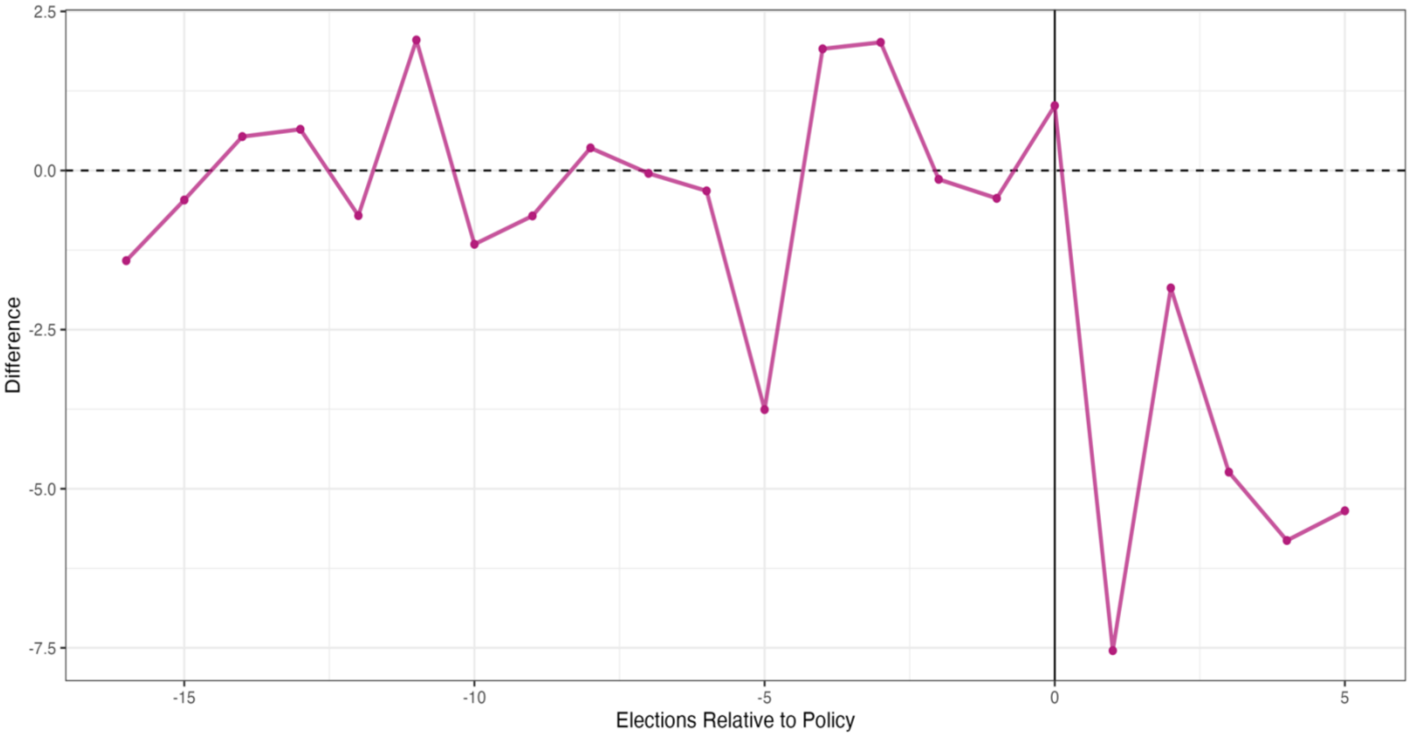


Figure 5: Outcome Difference Between Synthetic and Observed Penn.

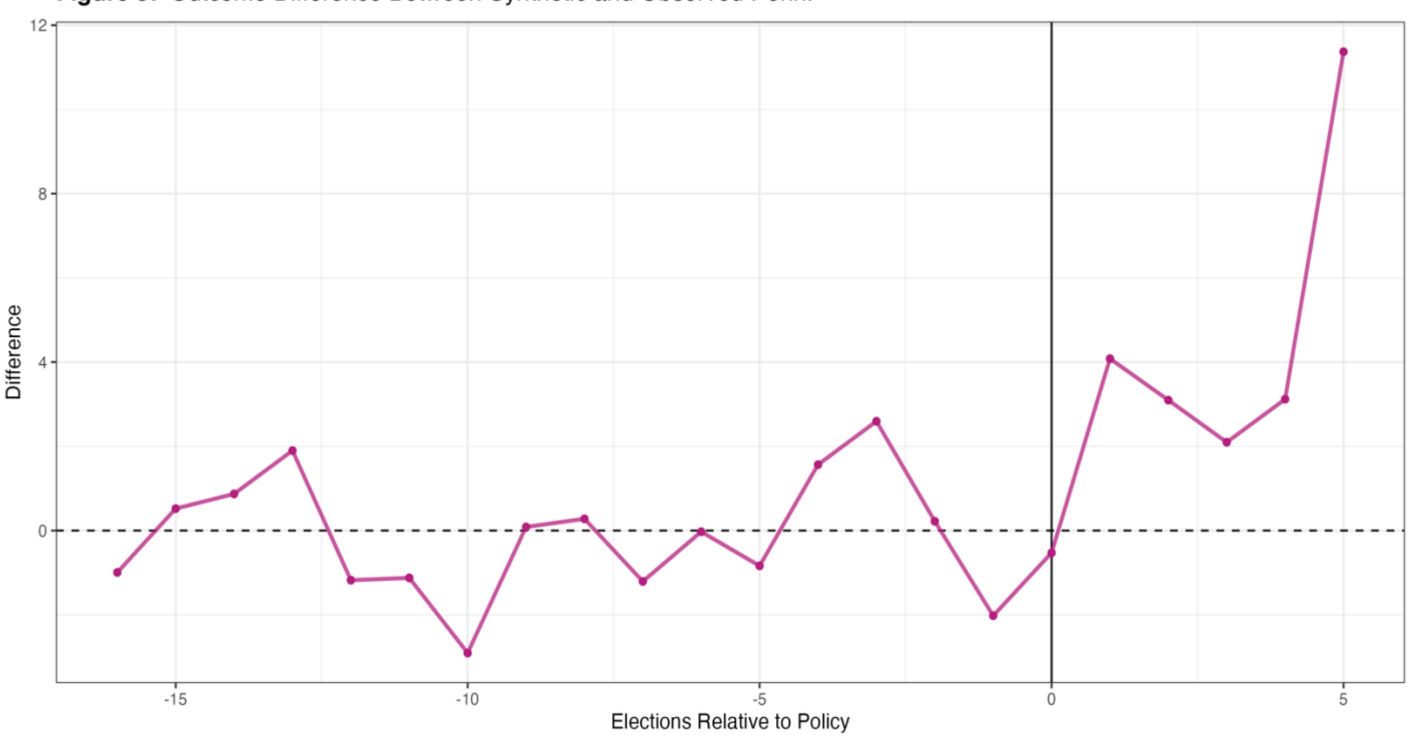
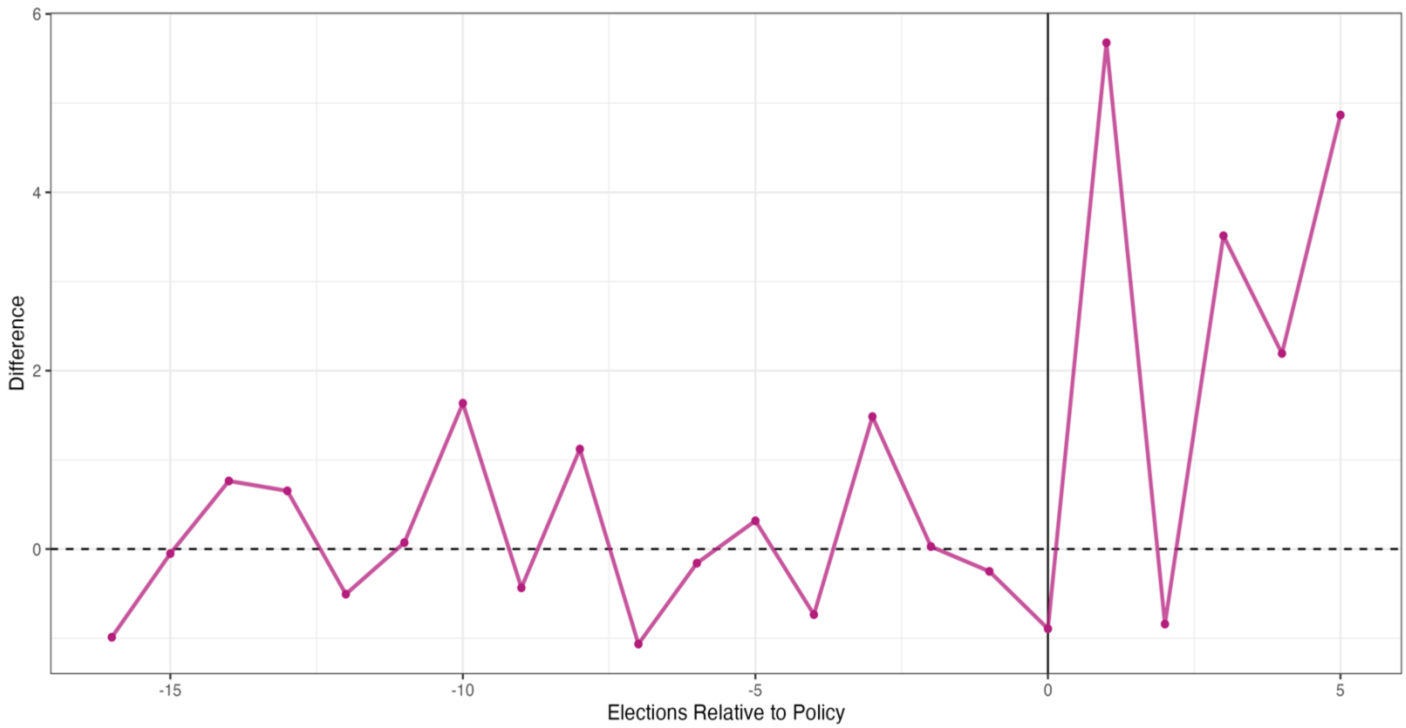
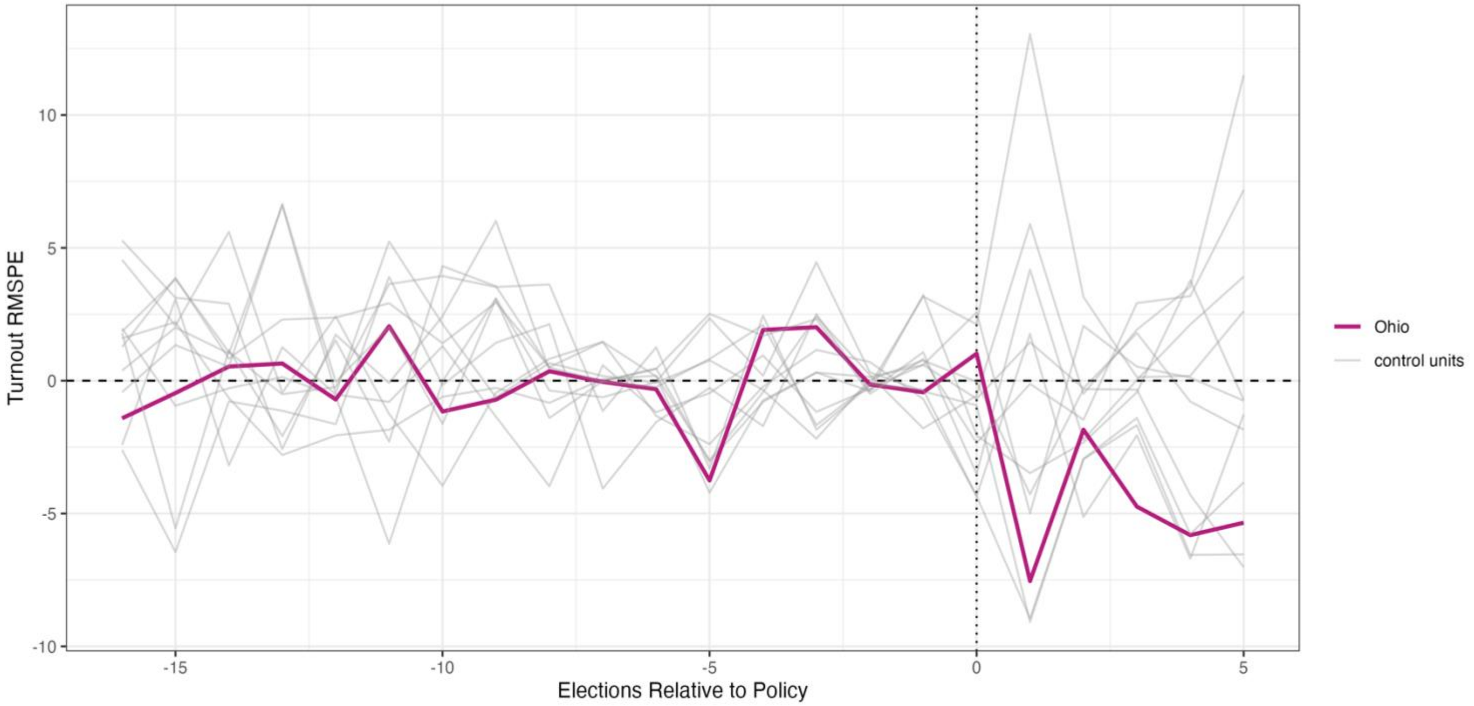


Figure 6: Outcome Difference Between Synthetic and Observed Kansas



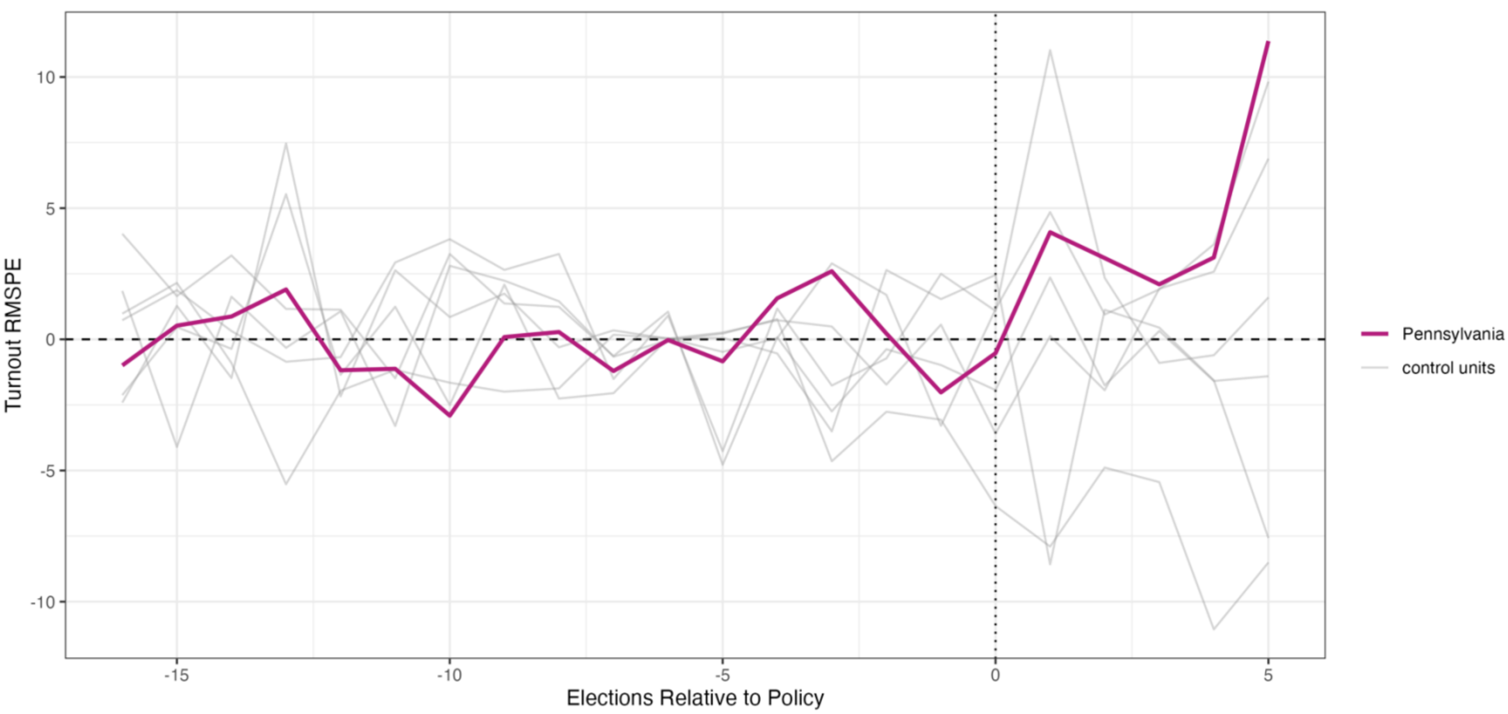
Appendix C: Placebo Plots for Three Synthetic Control Runs

Figure 7: Placebo Runs for Synthetic Ohio



Pruned all placebo cases with a pre-period RMSPE exceeding two times the treated unit's pre-period RMSPE.

Figure 8: Placebo Runs for Synthetic Penn.



Pruned all placebo cases with a pre-period RMSPE exceeding two times the treated unit's pre-period RMSPE.

Figure 9: Placebo Runs for Synthetic Kansas