The Impact of State Voter Registration Policies on State Registration Rates

Presented to the Faculty of the
Program in Quantitative Social Science and
The Nelson A. Rockefeller Center for Public Policy

Dartmouth College

By

Eitan Ziv Darwish
Advised by Professor Ronald G. Shaiko

Hanover, New Hampshire
June 6th, 2021

Dedicated to the democratic (with a small d) process

Acknowledgments

Writing this thesis would not have been possible without the people that shaped my educational experience.

First, I thank Professor Ronald G. Shaiko for ushering me into the world of policy, guiding me through several courses, and taking interest in this thesis. Professor Shaiko taught the first public policy course I ever took–*Introduction to Public Policy Research*, which prepared students for the Policy Research Shop at the Rockefeller Center. Professor Shaiko taught me that public policy has purpose through the PRS and *Leadership in Civil Society*, where our class was able to fund local non-profits through a formal grant-seeking process. Professor Shaiko's mentorship has shown me that I can have an impact on my community, and for that I will forever be grateful. I am delighted that he advised this thesis.

Next, I'd like to thank Professor Jeremy Ferwerda for being my informal advisor on the technical part of my analysis. Professor Ferwerda was incredibly supportive throughout the year, and took the time to walk me through technical concepts such as the basics of machine learning and modeling LASSO regression. Of course, Professor Ferwerda was also a very helpful project manager throughout the year, making sure the entire QSS Honors Thesis cohort was hitting deadlines and on track to finish.

I also want to thank Professor Herschel S. Nachlis for his feedback on drafts of this study. His comments have made this analysis far more robust, and for that I thank him.

Thank you to Professor Michael McDonald at the University of Florida, for all of the work he did compiling voter registration data, and determining each state's voting-eligible population for the past twenty-plus years!

I'd like to briefly acknowledge some professors throughout my Dartmouth tenure that had a meaningful impact on my educational trajectory:

First, Professor Yusaku Horiuchi taught the first QSS course I ever took, *Data Visualization*. After this class, I became so passionate about data visualization that I pursued it in my professional life! The skills I gained in Professor Horiuchi's class helped me land my first internship, snowballing down to the job I have after college.

Next, Professor Michael Herron had a knack for making the theories that support everyday social science fun for me to learn. From learning poisson distributions in *Sports Analytics* to understanding the strategy behind "chicken!" in *Political Game Theory*, Professor Herron helped me feel comfortable with the most fundamental building blocks of statistics and political science. Professor Herron also informally advised me on this thesis. Thank you!

Professor Charlie Wheelan is the next significant professor in my time at Dartmouth. Professor Wheelan taught me that policymaking is as much an art as it is a science.

Finally, of course, I thank my friends, my partner, and my family for the joy they provide me in life. Without that fuel, it is difficult to stay motivated!

Abstract

Unlike former studies investigating the effects of voting policies, this thesis looks strictly at the impact of voter registration policies on *registration rates*. The registration rate in a given state is the proportion of the 'voting eligible population' that is registered to vote prior to a national election. This study observed every US general and midterm election from 2000-2020, and looks at the effects of the following state policies on registration rates:

- Being a part of the Electronic Registration Information Center (ERIC)
- Offering Election Day, or Same Day Registration (SDR)
- Offering Online Voter Registration (OVR)
- Offering Automatic Voter Registration (AVR) through either a state's Department of Motor Vehicles (DMV) or through another state body
- Requiring certain forms of ID in order to register
- Partisan control of state legislatures, state governorships, and the party of the State Secretary

A preliminary OLS multivariable regression found that ERIC membership increases registration rates by 1.82 percentage points, OVR increases registration rates by 2.01 percentage points, and AVR (non-DMV) increases registration rates by 3.01 percentage points at the p=0.05 level.

However, the OLS regression covers few parameters and the dataset is relatively small, making the model prone to high variance. To account for this, thus study employed a LASSO regression to shrink the coefficients of unimportant variables to zero, which yielded slightly different results:

AVR through the DMV, and through other state bodies, are associated with a 1.57, and a 2.05 percentage point increase in registration rates, respectively. ERIC was found to be associated with a 3.47 percentage-point increase in registration rates.

However, after subsetting the data by election type (midterm/general), the LASSO model shifted. In midterm elections, none of the variables survived shrinkage. In general elections, however, AVR (DMV) was associated with a 1.87 percentage point increase in registration rates; AVR (Other) a 9.11 percentage point increase; OVR a 2.30 percentage point increase; and ERIC a 3.72 percentage point increase. Additional regression analyses verified that policy effects were unpronounced, or entirely invisible for ERIC in midterm elections, but significant in general elections.

While other research has demonstrated the efficacy of AVR and OVR, a more robust analysis of ERIC was implemented to investigate causality. A difference-in-differences approach was used to verify that a state's joining ERIC increases registration rates, but only for the general election that follows the state's entry.

Contents

List of Figures & Tables

1 Introduction

Voter turnout was a key topic of conversation throughout the 2020 election cycle. Many variables are baked into how many Americans ended up going out to vote–especially during a global pandemic–such as whether absentee ballots were sent to voters on time and counted, the weather in a given area code, and partisan control of state legislative bodies. **Another key variable in the voter turnout composite is the proportion of citizens in a state who are registered to vote.** While voter turnout is in the headlines during an election cycle, voter registration is top of mind for grassroots activists who want to increase access to the ballot. Over the past few months, however, voter registration has been making national news, as partisan actors look to limit citizens' ability to register to vote in the name of election security following the tumultuous 2020 general election in which the 45th president of the United States alleged, with no evidence, that American elections are insecure and full of fraud. As such, many states have introduced or passed legislation aimed at specific parts of the voter registration process, such as whether voters can be automatically registered through a state's Department of Motor Vehicles, whether citizens can register to vote on election day, and several other specific policies. This thesis asks the question:

Which voter registration policies have greatest impact on the states' registration rate?

As pointed out by Cass Sunstein and Richard Thaler in *Nudge*, nonpartisan and rather mundane policies can have an incredible impact on outcomes of interest. For example, Sunstein and Thaler looked at whether a driver is asked to "opt-in" or "opt-out" of being an organ donor–Sunstein and Thaler found that switching from the former to the latter could save countless lives by shortening waiting lists for organ donations. Like a driver's organ donor status, a citizen's voter registration status is largely determined by whether they filled out a form with certain information. To support the goal of maximizing participation in American democracy, many states have considered elegant ways of making registering to vote easier through leveraging default biases, as described by Nobel Prize-winning behavioral psychologist Daniel Kahneman, as well as Sunstein and Thaler. The unfortunate caveat to the voter registration nudges is that they are political, unlike many of the reforms discussed in *Nudge*. Especially in the wake of the 2020 election, legislators from the Republican party have opposed such reforms as opening the door to fraud. Notwithstanding, this analysis will bring clarity to the question of whether voter registration reforms are effective at increasing enfranchisement, and which reforms specifically are most effective.

It is surely more difficult to register to vote in states where an active effort on the part of the citizen is required in order to register than if the citizen had several opportunities to register by default with an opt-out option. In the United States today, there is a wide range of registration difficulty state-to-state. In New Hampshire, you can register at the polls five minutes before you cast your ballot. In California, a citizen can register online fifteen days before Election Day. In Arkansas, though, registration is only available by mail or in-person at least thirty days before Election Day. In Oregon, however, every citizen who applies for or renews their driver's license is automatically registered to vote. Each state's policies around voter registration are different, and the purpose of this thesis is to isolate the effect of voter registration policies on the proportion of registered voters in that state.

2 Definitions

Understanding what "registration rates" are

The outcome variable used in this analysis, called "registration rate," is the share of the "voting-eligible population" (VEP) that is registered. The voting eligible population of a state removes non-citizens, those under the age of 18, and, in many states, those who have felony convictions. Data for voting-eligible populations of states over time were available through the University of Florida's *www.electproject.org*.

What can impact a state's registration rate?

The federal government set some baseline standards with the National Voter Registration Act of 1993, but some states chose to promote registration more than others. Since 2000, each state has implemented various policies that could potentially impact registration rates. Across the nation, states have implemented a variety of policies, including:

- **Sharing data with The Electronic Registration Information Center (ERIC)**
- **Purging/removing voters from registration rolls (Removals)**
- **Same-Day Registration (SDR), or Election Day Registration (EDR)**
- **Online Voter Registration (OVR)**
- **Automatic Voter Registration (AVR)**
- **ID Requirements for Registration**

For AVR, states can choose to implement automatic registration through their Department of Motor Vehicles (DMV), or through some other state agency. For example, Alaska passed a ballot measure in 2016 that automatically registers all citizens who apply for a Permanent Fund Dividend, the state's version of basic income. Still, beyond policies around how that target voter registration, there are certain characteristics about a state that may influence registration rate, such as:

- **Having a majority Democrat/Republican Legislature**
- **Having a Democrat/Republican Governor**
- **Having a Democrat/Republican Secretary of State**

These variables are important because they can impact registration rates in ways not captured by policies alone, including executive actions. So, there are many choices states can make that have an impact on its registration rate. Is the state a part of ERIC? Does the state have online or automatic voter registration? Across states and throughout the years, these variables will have an impact on registration rates.

3 Literature

A building block of this study is the fact that registration rates are important because they translate to voter turnout in elections. Highton in 2004 made this link by analyzing the effects of the National Voter Registration Act of 1993 and observing sharp increases in turnout along with registration rates. However, Highton notes that there remains a subset of the American population that is uninterested in voting, and that there will be a limit to turnout after a certain threshold has been reached.

Brown and Wedeking (2006-07) reached a similar conlusing when analyzing the National Voter Registration Act (NVRA). They found that, although registration used to translate to turnout nearly one-to-one, the NVRA created "a pool of registered citizens less likely to vote." This is because the NVRA focused on registering disenfranchised Americans, often lower-income, who are less likely to vote because of a litany of circumstances. After the NVRA, registration rates still predict turnout, but not like before the NVRA.

In terms of a direct analysis of voter registration policies, several non-profit organizations, including Demos, Vote.org, and the Brennan Center for Justice at New York University have employed small-scale non-academic analyses of automatic voter registration. A 2015 Demos research report projected that "approximately 27 million eligible persons" would be "added to voter rolls across the country if every state adopted automatic voter registration." The Demos figure was calculated rudimentarily by applying the increase in registration seen after Oregon's AVR reform to the entire country writ large. The Brennan Center for Justice has similar reports that give a high-level overview of why certain voter registration reforms are desireable a based on the evidence available, but the Center has also not performed a larger-scale statistical meta-analysis of registration rates by policy type and state.

Gonzalo Contreras, et al. (2014) examined local election turnout in Chile in 2012, and used an observational approach to estimate the impact of Chile's new automatic voter registratrion policy. Once again, this article was concerned with voter turnout rather than a direct impact on registration rate based on policy types. Due to the federal structure of the United States, it is unlikely a study in any other country could compare to what could be done in the United States, since there are such differences in laws state by state.

One of the most compelling pieces of research on voting reforms and their direct effects on voting was undertaken by Michael J. Hanmer in his book, *Discount Voting: Voter Registration Reforms and their Effects*. Hanmer presents point estimates in the effect of Election Day Registration (EDR, synonymous with SDR) and Motor Voter registration (AVR) on the probability that someone in a state will vote. Hanmer found that the probability of voting for a citizen in a state that newly implements EDR increases by four percentage points. Additionally, citizens in states that implemented EDR long ago are 11 percentage points more likely to vote. This is an important contribution to the literature, though it answers a question that this thesis does not address: voter turnout. This thesis is purely interested in calculating the effects of reforms on the registration rate in a state, and comparing the effects of those reforms. However, the point estimates generated by Hanmer will be a helpful anchor for the work done in this thesis. Hanmer also uses individual-level data from the Current Population Survey to generate his estimates, while this thesis uses aggregate-level data.

Hanmer's contribution to the literature is significant, because he not only provides point estimates for voting reform policies, but also presents new theoretic frameworks for thinking about voting reform policies and their implementation. First, Hanmer argues that policies in states that implemented reforms for the explicit purpose of increasing registration rates have different levels of effectiveness than policies in states that implemented reforms for other reasons, such as data management and other bureaucratic functions. For example, Hanmer compares New Hampshire, a state that implemented EDR in a non-purpose-driven way, with Maine, a state that implemented EDR for the purpose of increasing registration rates. Regardless of the statistical approach, Hanmer finds that the effects of EDR are far more pronounced in Maine (~9 percentage point increase) than in New Hampshire (~1.5 percentage point increase). A similar pattern is shown with Motor Voter laws (or Automatic Voter Registration through a state's department of motor vehicles).

For example, Michigan, which implemented Motor Voter laws in a purposeful way, saw a more pronounced effect than states like Wisconsin and Nevada, which did not. Hanmer's contributions are vital to understanding the intricacies of voter reform policies, and his analysis shows that there is great variability in policy effects depending on the state. This thesis will seek to control for these differences by using state factor variables as controls.

Another important piece of research around specific registration reforms was Jinahi Yu's study of Online Voter Registration (OVR) in his *Social Science Quarterly* article, "Does State Online Voter Registration Increase Voter Turnout," which analyzed general elections from 2000-2014. Yu found that OVR increased young voter turnout in presidential election years by three percentage points, and that, for a given voter, using OVR increases their turnout by 18 to 20 percentage points.

The Tufts University Center for Information and research on Civic Learning and Engagement published "Voter Registration among Young People in Midterm Elections." This was a contribution to the literature on voting practices, as the report analyzed data on the methods used by voters to register. The report found that, in states where OVR was available in 2010, just one percent of adults over the age of 30 and four of adults 18-29 registered to vote online.

There is also a body of research that discusses the implications of voter registration reform. Christopher B. Mann, et al. (2020) have investigated attitudes about automatic voter registration through the DMV. Mann, et al. were interested in whether AVR was a partisan issue and attempted to highlight public attitudes about AVR, voter fraud, fairness, and election problems. While this analysis is insightful for policymakers to consider, it does not answer the same questions this thesis hopes to address, which is to explain the difference in registration rates by state based on the registration policies therein.

Thessalia Merivaki's *The Administration of Voter Registration: Expanding the Electorate Across and Within States* is a recent (2020) review of voter registration efforts on the national and state level. In addition to analyzing data, Merivaki discusses questions that cannot be captured by aggregate-level data. In "2.2, Why Don't Americans Register to Vote?" Merivaki looks at the Current Population Survey to examine the reasons Americans have for not registering to vote. According to the 2012 CPS, 4.57 percent of those who did not register to vote in 2012 said that they failed to register because they "did not know where or how." For the purposes of this thesis, this subset of the population is the target of many voter registration reforms such as ERIC and AVR. Merivaki also looks at nationwide "voter registration gaps" and shows a stready increase in the number of eligible but unregistered voters from 2008 to 2016, stressing the urgency for increased ballot access through voter registration reforms.

It is important to note that while voter registration rates are important explainers of voter turnout, which is the ultimate variable of interest for policymakers, other scholars have demonstrated that other factors unrelated to registration rates can influence voter turnout. For example, Stewart and Ansolabehere (2015) demonstrate that long lines to vote discourage voting; Ansolabehere et al. (2000) showed that redistricting policies can give incumbents an advantage in elections; Walker, Herron, and Smith (2019) demonstrated that local rules regarding early voting induces a range of reactions from voters depending on local conditions; and Merivaki et al. (2020) found that rejected provisional ballots in North Carolina prevented people from voting.

Finally, there is a healthy body of research around who in particular bears the brunt of disenfranchisement in the United States. Notably, Dayna Cunningham's article "Who Are to be the Electors" in the *Yale Policy Review* enumerates how minority and poor citizens have been most affected by legislation around voter reform. While this thesis does not discuss demographics, it is important to note the importance of this work. Any findings in this thesis will have social implications, and researchers like Cunningham explain exactly how voter reform can correct racial and economic inequalities in the United States.

4 Hypotheses

Though this analysis will include several control variables in its regression approach, coefficients on the following variables, which are the policies that vary across states and time, will be estimated. The expected sign of the coefficient and rationale for each variable is outline below:

ERIC Membership

Expected Impact on Registration Rate:

(/) If a state is a part of the Electronic Registration Information Center, registration rates should not change significantly.

Rationale:

ERIC makes sure that voters are not registered in multiple states, but also mails registration instructions to citizens who are eligible to vote but not registered. These mailings should cancel out purges.

Same Day Registration (SDR)

Expected Impact on Registration Rate:

(+) If a state offers citizens same day registration, registration rates will increase.

Rationale:

Same day registration means that registering to vote is less restricted, and that people who may have forgotten to register before election day are able to. Since more voters are being included, an increase in rates is expected.

Online Voter Registration (OVR)

Expected Impact on Registration Rate:

(+) If online registration is allowed, registration rate will increase.

Rationale:

Allowing online registration increases convenience, making it easier to register.

Automatic Voter Registration (AVR) through a State's DMV

Expected Impact on Registration Rate:

(+) Automatic Voter Registration will increase the proportion of registered voters.

Rationale:

Defaulting nearly all citizens (those who drive) to become registered will significantly increase registration rates; less meaningful in urban areas.

AVR through other State Bodies

Expected Impact on Registration Rate:

(+) Automatic Voter Registration will increase the proportion of registered voters.

Rationale:

An increase is expected, but probably not on the same magnitude as DMV AVR, since driving is common.

Proof Documents (ID Required)

Expected Impact on Registration Rate:

(-) An increase in the burden of proof for a citizen to register will decrease registration.

Rationale:

Needing to provide proof documents is more burdensome than no proof documents.

Voter Roll Purges (Removals)

Expected Impact on Registration Rate:

(-) Purging voter rolls will decrease the number of registered voters by an amount proportional to the purge.

Rationale:

Unless a significant number of people whose rolls were purged re-register, people who were once registered will now not be.

State Legislature Partisanship

Expected Impact on Registration Rate:

The proportion of registered voters will increase after an election cycle when control of a state legislature flips blue.

Rationale:

Policies that are associated with enfranchising more voters are associated with the Democratic party, while policies such as gerrymandering and voter suppression are associated with the Republican party.

Governor Partisanship

Expected Impact on Registration Rate:

The proportion of registered voters will increase after an election cycle when the governorship flips blue.

Rationale:

Policies that are associated with enfranchising more voters are associated with the Democratic party, while policies such as gerrymandering and voter suppression are associated with the Republican party.

Secretary of State Partisanship

Expected Impact on Registration Rate:

The proportion of registered voters will increase after an election cycle when the Secretary of State becomes a Democrat.

Rationale:

Policies that are associated with enfranchising more voters are associated with the Democratic party, while policies such as gerrymandering and voter suppression are associated with the Republican party.

5 Research Design

This study will be observational in nature, looking at past general and midterm elections (every two years) in all fifty states, and track voter reform policies in each state during each election year, using the parameters outline in the equation below. For ERIC, SDR, OVR, AVR, Legislature Partisanship, Governor Partisanship, and Secretary of State Partisanship, binary trackers [0,1] will be used to determine whether or not a state had implemented such a policy prior to that election. For the variable that tracks proof documents, or whether or not an ID is required in order to register, [0, 1, 2] will be used to determine whether a state does not require an ID to register (0), requires an ID, but is not strict about what type of ID (1), or has a strict governmental ID requirement (2). Removal data will be normalized by dividing the total amount of voters purged in a state before a given election year by the total number of eligible voters in that state. This will create a variable Removal/VEP that will track the severity of purges in a state for any given election year.

In the regression equation below, i is the state/election combination (observation), n are the coefficients on the reform variables, and is the average registration rate across all states. The terms control for state- and year-effects. These controls are especially salient when thinking about how high turnout was in 2020, either because of the 45th president, the COVID-19 pandemic, or both. Controlling for year effects will make sure that exogenous events do not influence the coefficients of interest.

While a multivariable regression analysis might be helpful, the covariance of these policies is likely high, as certain states might tend to implement many of them and other states might implement none. To account for this covariability, other models should be considered. One useful model for teasing out the effects of covaried variables is the Least Absolute Shrinkage and Selection Operator model, or the LASSO model. This linear regression model, which is a type of machine-learning model, uses shrinkage. Fewer parameters and a relatively smaller dataset mean that our predictions from OLS might contain more error. The shrinkage from LASSO accounts for this by eliminating variables that are weakest in association with our target variable. LASSO employs L1 linearization, which will shrink the effects of unimportant variables, often to zero. The point of shrinkage is to reduce variance and dramatically increase readability, hopefully helpful in informing policymakers as to which reforms are most impactful when it comes to increasing registration rates.

A simple multivariable regression controlling for state and year fixed-effects will initially estimate the impact each voting reform policy has on registration rates:

$$RegistrationRate_i = \alpha + \beta_1 ERIC_{i,t} + \beta_2 SDR_{i,t} + \beta_3 OVR_{i,t} + \beta_4 AVR_{DMV\ i,t} + \beta_4 AVR_{Other\ i,t}$$
$$+\ \beta_5 IDRequired_{i,t} + \beta_6 Removals_{i,t} + \beta_7 RepLeg_{i,t} + \beta_8 RepGov_{i,t} + \beta_9 RepSoS_{i,t} + \gamma Year_t + \gamma State_i + \varepsilon_{i,t}$$

Data

Vote.org, the National Conference of State Legislatures, and the Brennan Center for Justice had nearly all of the information regarding state-by-state registration policies. However, these data were not in an easily readable form. Fortunately, Wikipedia had readable data tables with the same information that was easily ingestible into a .csv file, though there were some inconsistencies. Inconcistencies were corrected with numbers from the Brennan Center, and after scanning through to make sure that the correct numbers were in, the data were compiled into a master set. Each cell represents a policy status for a policy type in a given state in a given year. The data from the Brennan Center accounted for many of the columns of the explanatory variables examined.

Political Science Professor Michael P. McDonald of the University of Florida has compiled an incredible dataset of state populations and voting-eligible populations for all 50 states in every election year since 2000. This is the response variable, and is imperative to the analysis.

Competitiveness data was extracted from Harvard's opensource database of US election results. The absolute margin of victory of one party over another was used as a proxy for competitiveness.

5.1 Limitation: Removal Data

Unfortunately, state-level data on removals (or purges) is not easily accessible. While the federal government published these data from 2010-2018, 2020 is not yet available, and the federal government did not publicly report this data prior to 2010. While some states published their removal data prior to 2010, others did not, which renders the removal variable incomplete in years prior to 2010.

However, a subset analysis will allow us to determine whether the removal data's being missing poses a non-trivial threat to this analysis. We will take the subset of the data for which we have removals, and run an OLS regression estimating the coefficients with an without the removal data, and then compare the difference between the estimates for the coefficients.

Shown in the table below, our subset analysis from 2010-2018 reveals that including removal data does not have a meaningful impact on the point-estimates of other variables, allowing us to proceed with the analysis without removal data. The negligible impact of removal data is likely attributed to the covariance of removals with state fixed effects.

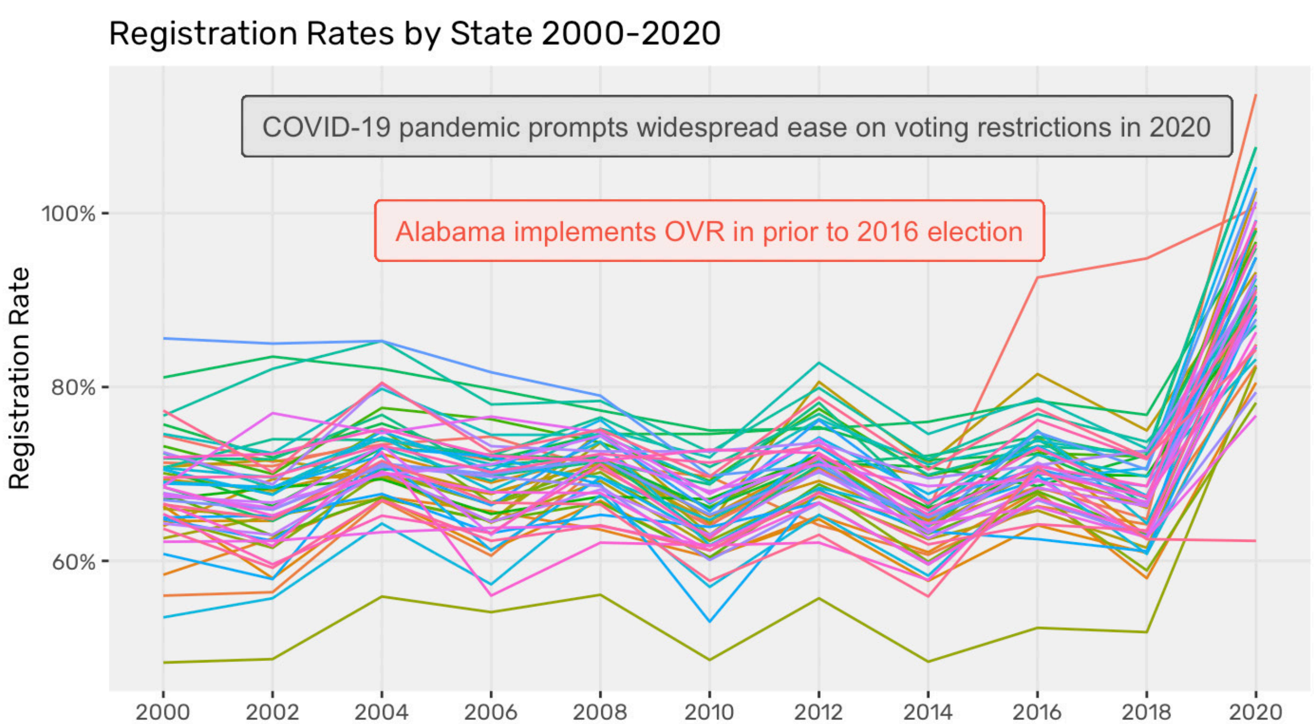
Subset Analysis for Removal Data **TABLE 5.1**

Variable	With Removal Estimate	No Removal Estimate	Difference
ERIC	0.0145	0.0165	-0.00202
SDR	0.0158	0.0147	0.00113
OVR	0.0027	0.0022	0.00051
AVR, DMV	-0.0105	-0.0078	-0.00273
AVR, Other	-0.0119	-0.0145	0.00258
ID Required	0.0065	0.0057	0.00087
Republican Leg.	0.0094	0.0076	0.00176
Republican Gov.	0.0126	0.0127	-0.00008
Republican SoS	-0.0054	-0.0060	0.00057

6 Results

In this section, I discuss the results of my analyses, starting with preliminary single-variable regressions, and then moving on to the OLS multivariable regression described in *5: Research Design*, and then the LASSO analysis. Finally, this section will also include a deeper investigation of ERIC in *Section 6.4*. As will be discussed in this section, ERIC was identified by multiple models to be a significant driver of registration rates. *Section 6.4* outlines the results of differences-in-differences analyses for ERIC, seeking to clarify its effect.

FIGURE 6



Fluctuations in registration rates every two years are driven by the fact that General Elections, which are perceived as more consequential, usher increased registration rates as opposed to Midterm Elections

6.1 Preliminary Analysis

The results from the preliminary single-variable regressions are shown below. The purpose of the preliminary analysis is to identify correlations between the analyzed variables and registration rates. The preliminary analysis will serve as context for the multivariable regression to follow, and then the LASSO analysis.

Single-variable OLS Regressions **TABLE 6.1**

Single-variable regressions include factors for State and Year as controls

Variable	Coefficient Estimate	*t*-score	Pr(>|*t*|)	
Removals/VEP	-0.1190352	-1.401	0.162725	
AVR (DMV)	0.0151857	1.627	0.104379	
AVR (Other)	0.0336807	2.851	0.004535	**
OVR	0.0281618	4.680	0.000004	***
SDR	0.0132027	1.710	0.087969	
ERIC	0.0276369	4.362	0.000016	***
ID Requirements	0.0021062	0.527	0.598373	
Republican Governor	0.0026059	0.641	0.521963	
Republican-controlled Legislature	0.0075220	1.340	0.180720	
Republican Secretary of State	0.0001757	0.035	0.971799	

Significance codes: 0 = ‘ *** ’ | 0.001 = ‘ ** ’ | 0.01 = ‘ * ’ | 0.1 = ‘ . ’

The preliminary single-variable regressions above indicate that AVR (non-DMV agency), OVR, and ERIC drive registration rates, all in the positive direction. For each of these three variables identified, the single-variable regressions estimate roughly a three percentage point increase in a state's registration rate when the given variable is offered. Specifically, the model estimates that a state newly offering automatic voter registration through a non-DMV agency is associated with a 3.37 percentage point increase in registration rates; a state newly offering online voter registration is associated with a 2.82 percentage point increase in registration rates; and that a state joining ERIC is associated with a 2.76 percentage point increase in registration rates. These preliminary results are further investigated with multivariable OLS regressions and LASSO analyses.

6.2 Multivariable OLS Regression

Our preliminary analysis identified three significant variables that we might expect to be significant for the OLS multivariable regression. In terms of other preliminary considerations, see Table A1 in *Appendix A* for the covariance between selected variables and Table A2 in *Appendix B* for a check on whether lagged variables are more significant than non-lagged variables.

Multivariable OLS Regression **TABLE 6.2**

Includes factors for State and Year as controls

Variable	Coefficient Estimate	t-score	Pr(>|t|)	
AVR (DMV)	-0.0056404	-0.494	0.621873	
AVR (Other)	0.0310616	2.202	0.028134	*
OVR	0.0201087	3.022	0.002641	***
SDR	0.0085407	1.062	0.288563	
ERIC	0.0181778	2.609	0.009365	**
ID Requirements	0.0009180	0.230	0.817887	
Republican Governor	0.0037183	0.849	0.396364	
Republican-controlled Legislature	0.0095256	1.684	0.092856	.
Republican Secretary of State	0.0010157	0.193	0.847305	
factor(YEAR), 2020	0.1992733	20.102	0.000000	***

Significance codes: 0 = ‘ *** ’ | 0.001 = ‘ ** ’ | 0.01 = ‘ * ’ | 0.1 = ‘ . ’

Our multivariable regression produced some interesting findings. First, we find that ERIC, OVR, and AVR (Other), are all statistically significantly associated with an increase in voter registration rates for a given state in any given election year. To be more precise, we can interpret the coefficient estimates from our OLS multivariable regression:

If a state that is not a part of ERIC decides to join ERIC, we can expect the registration rate in that state to increase by 1.82 percentage points. This prediction is statistically significant at the p=0.01 level. This result was surprising to say the least. At the onset of this project, I thought that including ERIC was an important control. It turns out that ERIC was one of the most significant explainers of registration rates out of the selected policies, and that it is associated with a nontrivial increase in registration rates. In a state like California, for example, which is not a part of ERIC, a two-percentage-point increase in registrations can mean as many as 500,000 newly registered voters. Notwithstanding the caveats of the OLS regression, which is compromised by high variance, this result should not be taken lightly, and we will return to ERIC when discussing the LASSO model in *6.3: LASSO Regression*.

Next, if a state that did not have online voter registration decides to implement OVR before an election cycle, we can expect voter registration rates for that next election cycle to increase by 2.01 percentage points. This prediction is statistically significant at the p=0.01 level. This matches our hypothesis about online registration: making registration more convenient will increase registration rates. Looking at the top figure on page 9, we can see, even with the naked eye, the power of OVR in Alabama's registration rate changes from 2014 to 2016.

Finally, our analysis found that if a state implements automatic voter registration through a department other than the state's DMV, we can expect registration rates in the following election to increase by 3.11 percentage points. This result was statistically significant at the p=0.05 level. This result was likely driven, at least in part, by Alaska's implementing AVR through their permanenet dividend fund, which is incredibly popular.

Finally, I wanted to take a brief moment to highlight the power of the COVID-19 pandemic, and more broadly the 2020 general election. At the bottom of the table on page 10, I've isolated the fixed effect for the year 2020 to show the impact of the 45th president, the COVID-19 pandemic, social unrest for racial justice, explicit changes in voting policies because of the pandemic, and other events in 2020 on registration rates. This OLS model found that the events of 2020 were associated with nearly a *20 percentage point increase* in registration rates, with a *t*-score of over 20! This nugget of information is tangential to the goals of this thesis, but is interesting nonetheless.

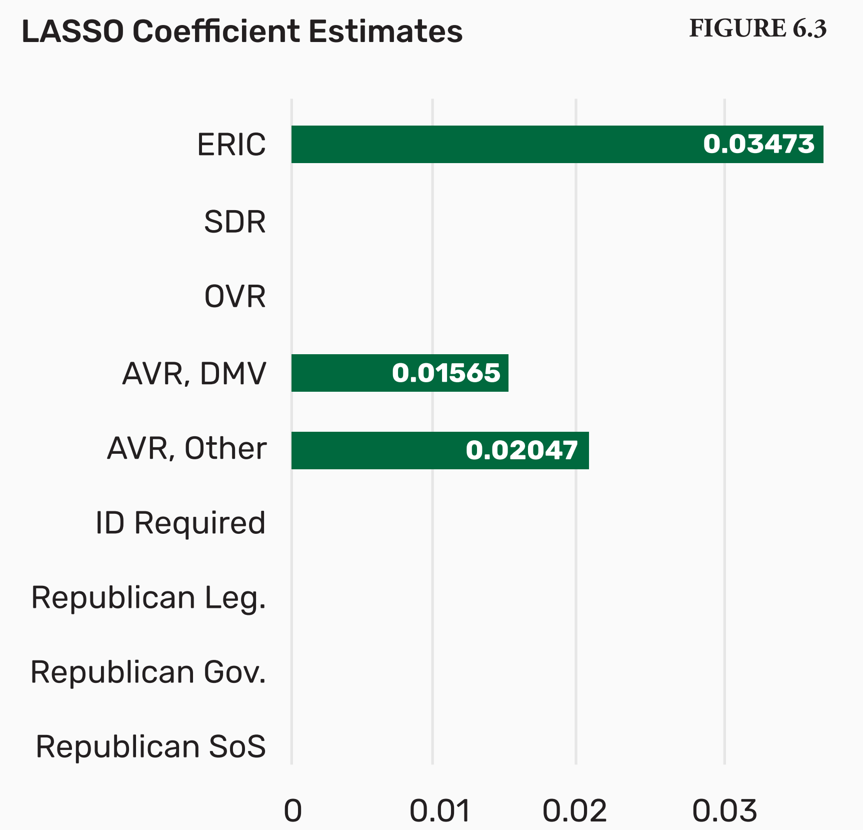
More so than what the model found to be statistically significant, it is worth noting what was *not* found to be significant by this model. Namely, we hypothesized that SDR, AVR (DMV), less stringent ID requirements, and partisan government control would have a statistically significant impact on registration rates, but the model found those variables to be insignificant.

However, before concluding that Automatic Voter Registration through the DMV and other policies not worth fighting for (they are worth fighting for!), we should account for the shortcomings of the simple OLS regression we ran, which had few parameters and relatively few observations. The LASSO analysis will help address these concerns.

6.3 LASSO Regression

Aforementioned, a LASSO model is useful for a dataset like ours because it has relatively few parameters and relatively few observations. The aim of a LASSO model is to *shrink* the coefficients of unimportant variables by running simulations and a basic machine learning model. The expected result is fewer significant variables and more easily interpretable data. For our purposes, then, we will expect only a few policies to stand out from our dataset and 'survive' the LASSO's shrinkage. The results of that analysis are shown to the right, and more details on the code used to run this analysis are found in the appendix.

FIGURE 6.3



6.3.1 Key Findings from LASSO

The first key takeaway from the LASSO regression was that the coefficients on SDR, OVR, ID Requirements, and partisan control were all shrunk to zero. This is exactly what we were hoping the LASSO would do. **Please note the following important caveat:** This result does *not* mean that SDR, OVR, ID Requirements and partisan control are unimportant in the fight for increased ballot access. I'd like to explicitly state that this study should *not* lead to any conclusions that these variables are unimportant. As discussed in *3: Literature*, many studies have been done that show the link between, say, OVR, and an increase in voter turnout. The purpose of this study is to investigate the factors that impact registration rates, and contribute a new wrinkle to the discourse on voting policies.

Two out of the three significant variables from this model are unsurprising: both types of automatic voter registration studied are determined to have a significant impact on registration rates. This supports our initial hypothesis inspired by Sunstein, Thaler, and Kahneman: default bias is incredibly powerful. When voters don't have to exert effort in order to gain access to the ballot (a guaranteed right), more voters will have access to the ballot.

Finally, perhaps the most important finding of this thesis is the degree to which ERIC membership was a significant explainer of registration rates in states. ERIC gets very little media attention, but this analysis predicts that becoming a part of ERIC can increase a state's registration rate by nearly 3.5 percentage points.

Before proceeding with the assumption that ERIC and AVR are important explainers of registration rates generally, it is important to consider whether we can generalize these results across all elections, midterm and general.

LASSO REGRESSION COEFFICIENTS **TABLE 6.3.2**

LASSO Regressions Subsetted by Year
Factor dummies for State and Year included in all regressions

			Coefficient Estimate
Variable	**All Years**	**Midterm Elections**	**General Elections**
AVR (DMV)	0.01565354	0	0.01866320
AVR (Other)	0.02047096	0	0.09106864
OVR	0	0	0.02295378
SDR	0	0	0
ERIC	0.03472969	0	0.03715017
ID Requirements	0	0	0
Republican Governor	0	0	0
Republican-controlled Legislature	0	0	0
Republican Secretary of State	0	0	0

6.3.2 Additional LASSO Models, Refined Findings

Subsetting our analysis by type of election might tell a story about the types of citizens being reached by these policies. Unregistered voters, might be unengaged, by virtue of not taking the initiative to register, and thus might be less likely to vote in (and register for) midterm elections than they would be in general elections, which include voting for president of the United States. Therefore, two more LASSO analysis subsetted by type of election were run to investigate whether policy measures like ERIC and AVR were associated with increases in registration rates for midterm election years.

The LASSO model for just midterm election years predicted no significant coefficients, bolstering the hypothesis that **measures that target unengaged voters might not be as effective for midterm elections**. To further investigate this result, I ran another simple OLS multivariable regression with just midterm election years. This regression found no significant variables at the p=0.05 level. While the OLS regression is a quick glace at the data, this brief finding further supports the hypothesis that the effects of voter registration policies vary based on whether the next election is a midterm or a general election.

However, when looking at just general elections compared with all elections, the coefficient estimate of ERIC increased from 3.47 percentage points to 3.72 percentage points; the coefficient of AVR (DMV) increased from 1.57 percentage points; the coefficient of AVR (Other) increased from 2.05 percentage points to 9.12 percentage points; and OVR went from being an insignificant variable to having a coefficient of 2.30 percentage points.

Although it did not survive the LASSO's shrinkage for all years, OVR emerges as an explainer of registration rates for general election years, verifying previous findings touting its ability to increase rates.

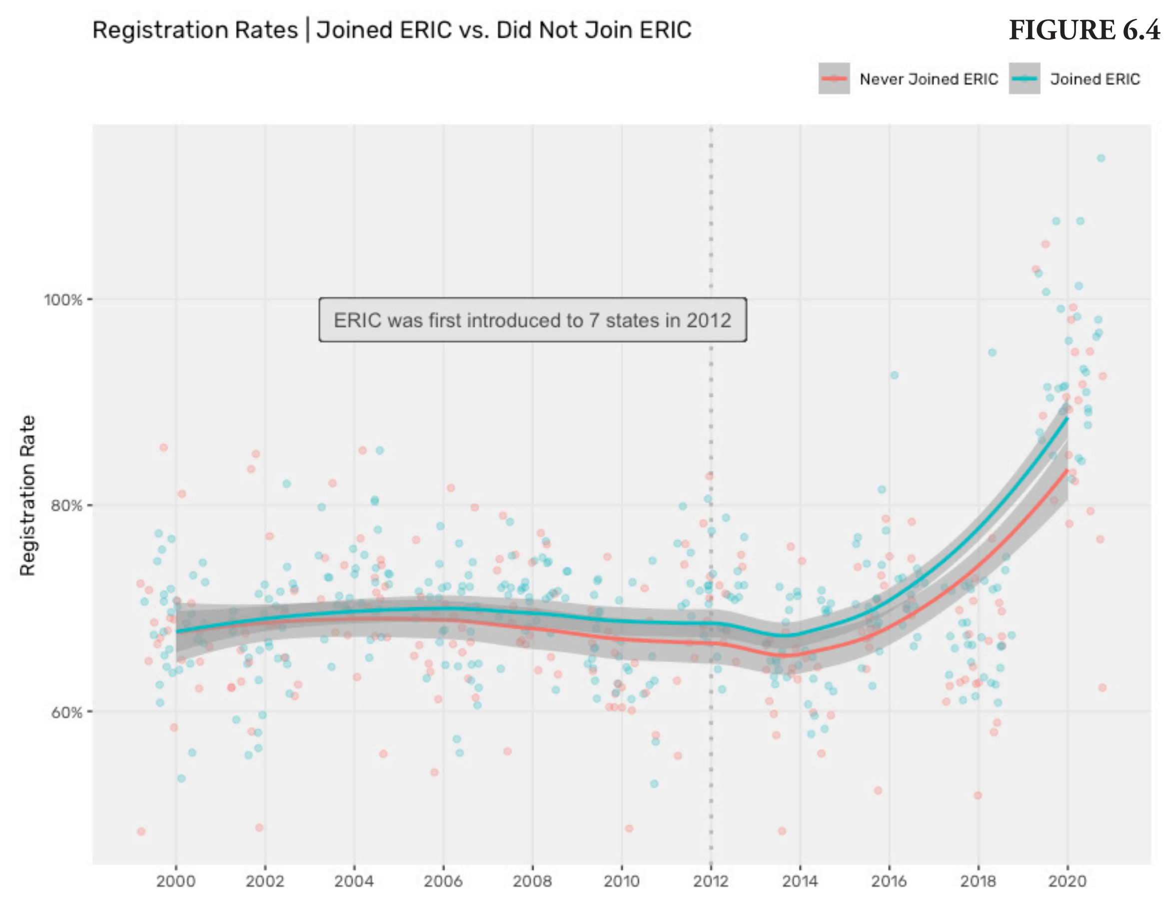
6.4 A Further Investigation of ERIC using Differences-in-Differences

The current literature already does an excellent job of linking AVR and OVR to increased voter turnout, which is the ultimate variable of interest (though outside the scope of this thesis), and many of these studies have made causal claims. ERIC, on the other hand, has been minimally studied and should be investigated further before making any sort of causal claim or policy recommendation. While the LASSO analysis is extremely effective at feature selection, or telling us which variables matter, neither the OLS regression (*Section 6.2*) nor the LASSO models (*Section 6.3*) are sufficient for making a causal claim for ERIC increasing registration rates.

To investigate the strength of a causality argument, a difference-in-differences approach will be used. First, we can examine whether parallel trends persist in years prior to ERIC's introduction in 2012. As we can see in Figure 6.4, the group of states that joined ERIC and the group that did not maintain the same trajectory before 2012, with their margins of error overlapping. While the lines are not perfectly parallel, they are pretty close, allowing us to proceed with the analysis.

After the cutoff, both groups of states experience increases in registration rates, but states that joined ERIC pull ahead of those that did not. The question is whether the difference is significant. If so, then there is a strong argument that joining ERIC increases registration rates.

FIGURE 6.4



To test for significance in differences-in-differences, we will create a dummy cohort variable for whether a state entered ERIC, and interact it with ERIC in a regression. In addition to looking at the cohort of all states that entered ERIC, as shown in Figure 6.4, we will examine each cohort of states that entered ERIC before a given election year. For example, in Table 6.4, *'2014'* means that a difference-in-differences test was conducted with the 'treatment group' being only states that joined ERIC after the 2012 election and before the 2014 election. Results are shown below. For difference-in-differences graphs for all cohorts as show in Figure 6.4, see *Appendix H*.

Difference-in-Differences Tests for ERIC by Entering Cohort **TABLE 6.4**

All tests include controls for Year and State; 'Controlled' tests include OVR, and both types of AVR

		Pr(>|*t*|)			
Analyzed Cohort	**Election Type**			*With Controls*	
All Years		0.000016	***	0.009524	**
2012	General	0.000154	***	0.066116	.
2014	Midterm	0.912081		0.199538	
2016	General	0.001723	**	0.242926	
2018	Midterm	0.707678		0.609227	
2020	General	0.000006	***	0.000029	***
2014 and 2016	General	0.238471		0.930004	
2018 and 2020	General	0.000325	***	0.011493	*

Significance codes: 0.0001 = ' *** ' | 0.001 = ' ** ' | 0.01 = ' * ' | 0.1 = ' . '
Cohorts with two year groups are analyzed using general election years only

The results of the difference-in-differences tests for ERIC continue to tell the story that was being told by the LASSO models: ERIC explains changes in registration rates, but ERIC's impact is significant in general election years, and negligible in midterm election years. As will be expounded upon in the *Discussion*, this is likely due to the types of citizens targetted by ERIC. Since ERIC mails letters to folks who are unregistered, ERIC is targetting a sample of citizens who did not take the initiative to register. For these "unengaged" citizens, a letter in the mail with information on how to register to vote in a midterm election would likely not be interesting. However, a letter clearly explaining the most efficient way to register prior to a presidential election might seem important enough to merit the cost of filling out the necessary documents to register, depending the state.

Is ERIC useless to states that join prior to a midterm? Joined cohorts at the bottom of Table 6.4 groups states that joined ERIC before a general election cycle. While the model does not predict that ERIC was helpful to states that entered prior to the 2014 midterm election (five states), the 2018/2020 cohort showed significance for the difference in differences due to ERIC, suggesting that states who entered ERIC prior to the 2018 midterm elections reaped the benefits in the 2020 election.

7 Discussion

Throughout the United States, state lawmakers are proposing legislation that would make it far more difficult to register to vote. These efforts are rationalized, generally, as a response to baseless claims of voter fraud in the 2020 presidential election. Georgia State Senator Jeff Mullis, who echoed many baseless claims of voter fraud, says “we believe just to assume that people need to vote might not be the right way.” Mullis was defending legislation that would end automatic voter registration and no-excuse absentee voting in Georgia. The bill had passed the Ethics Committee in February 2021.

The list of proposed legislation goes on. GOP lawmakers in Georgia, Ohio, Minnesota, and Montana have put forth legislation threatening to end SDR, OVR, AVR, and other progressive measures, which have proven incredibly successful. For example, in Georgia, AVR added more than 300,000 voters to the rolls! Still, lawmakers are also making strides in passing legislation allowing AVR and OVR; such is the case in Delaware, Hawaii, Maine, and elsewhere.

For activists hoping to expand ballot access by protecting AVR provisions and fighting for legislation that implements AVR, this study should be a vote of confidence that their time and energy is being well spent–the models in this study predicted increases in registration rates ranging from 1.57 to 9.11 percentage points associated with AVR. On the other hand, targetted efforts to attack AVR are well thought-out, and are certainly a threat to widespread ballot access. Thus, this study bolsters strategies by voter advocacy groups and grassroots organizers to fervently defend AVR and call attention to efforts to dismantle it, and other progressive voter registration policies across the nation.

Despite all of the national fireworks around voter registration, there are many reasons why Americans do not register to vote that have nothing to do with policy. Some Americans do not know where or how to register. ERIC reaches out to voters by mail to give them a gentle nudge to register to vote. ERIC would hopefully capture those citizens who would like to register but don’t know how (if AVR hadn’t already). However, this group of Americans makes up only about four percent of the American population that did not register (Merivaki, 35). The most common reason given for not registering to vote in the Current Population Survey was “not interested in the election or not involved in politics.” In 2012, over 43 percent of non-registered Americans reported not registering due to this ‘lack of engagement’ (Merivaki, 35).

The results of this thesis suggest that this group of Americans might not be stuck outside the fold of American democracy. It appears as though reforms that make voter registration easier can capture unengaged citizens when the stakes seem higher–in general election years. In 2004, Highton discussed the preventative costs of registering to vote, and how the NVRA of 1993 helped lower those costs by making registration more convenient. For some contemporary unengaged citizens, it appears that ERIC helps lower the costs of registration enough to make registering worth it–but only when they can vote for president.

Selected ERIC Member State Bylaws

Data Sharing:

States must be able to procure a full voter roll as well as records from state licensing/identification agencies (like the DMV) for all residents with active records–not just registered voters.

Reaching Out to Eligible Voters:

Members commit to contacting eligible but unregistered residents identified by ERIC, educating them on the most efficient means to register to vote.

Improving Roll Accuracy:

ERIC sends states reports that show voters who have moved out of state, voters who have died, and those who have duplicate registrations. This helps states maintain voter roll accuracy.

8 Conclusion

ERIC publishes statistics on its website every year, transparently reporting "List Accuracy" actions (removals), as well as the number of records it handles. Since its inception in 2012, ERIC boasts that it has identified over 55 million potential voters who were unregistered. On the flip-side, ERIC has identified nearly 6 million cross-state movers, in-state duplicates, and deceased records, along with over 10 million in-state updates.

However, despite the incredible success ERIC has deomnstrated in increasing registration rates and maintaining roll accuracy, twenty states have yet to join. Several hypotheses might explain why this is the case. It is important to point out the variety of states that are not a part of ERIC (see *Appendix F*).

Many smaller states in the Northeast and the middle of the country are not a part of ERIC, but the big blue giants, New York and California, also have yet to join. Why might this be? The bigger states have professionalized legislatures that probably tackle their own needs for roll maintenence because they have the resources. On the ERIC website, the "challenges in maintaining the accuracy of voter rolls" is explicitly mentioned as the impetus for such an organization, and the website even says that states, before joining ERIC, kept voter records in handwritten paper form. New York and California likely did not face many of these challenges, and would thus likely not benefit from joining ERIC.

This suggests that our predictive model might not be generalizable to all states. For example, California has implemented AVR through its DMV, likely capturing many unregistered voters, since most people drive in California.

On the other hand, a state like New Hampshire, which offers neither OVR nor AVR, might benefit from ERIC, which would capture unregistered voters that would otherwise have been captured through AVR and OVR. Thus, legislators looking to expand ballot access where AVR and OVR might be politically unfeasible can look to ERIC as a pragmatic way to advance an agenda that promotes expanded access to the ballot and secured voter rolls. Additionally, smaller-scale state governments, like those of New Hampshire, Montana, Maine, and other small states, would benefit immensely from the data management services offered by ERIC, which states currently handle on their own to maintain rolls.

In 2019, New Hampshire lawmakers passed legislation to enroll the Granite State in ERIC (House Bill 315). However, the bill was opposed by the Secretary of State, and ultimately vetoed by the governor. This study hopes to be part of a non-partisan conversation about the benefits of joining ERIC. While ERIC might seem to some like an unecessary technology that is expensive and cumbersome, ERIC is anything but. Besides a one-time fee of $25,000, small states can pay as little as $15,000 yearly to be a part of ERIC, and the only requirement is to send ERIC updated registration data and moter vehicle licensing data every two months. Essentially, that is a few data files sent over securely each time–not a logistical nightmare.

Besides the governments that might benefit from joining ERIC, the people who will benefit most from this legislation are Americans who are disillusioned with the political system and see no point in making the effort to register. A number worth repeating is 43 percent: the proportion of people who cite "uninterested in the election" or "not involved in politics" as the reason for not registering to vote. Gentle nudges, like ERIC mailings, that offer these folks a chance to easily register can bring people back into American politics, increasing engagement in our democracy.

This study predicts that, in general election years, ERIC can play a part in increasing voter registration rates, nudging civically unengaged citizens back into the political arena. States that are not a part of the Electronic Registration Information Center should consider joining.

Sources

A. Cassidy, Christina, and Hannah Fingerhut. "AP-NORC Poll: Majority in US Back Easier Voter Registration." AP NEWS, Associated Press, 2 Apr. 2021, apnews.com/article/ap-norc-poll-us-majority-back-easier-voter-registration-d4c6c40628aa4ddc56fbbd372d30dd04.

"Alabama Secretary of State Initiates ERIC Mailing Assistance Program." Alabama Secretary of State, www.sos.alabama.gov/node/1369.

Ansolabehere, Stephen, and James M. Snyder Jr. "Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency.." American Journal of Political Science (Wiley-Blackwell), vol. 44, no. 1, Jan. 2000, p. 17. EBSCOhost, doi:10.2307/2669290.

"Automatic Voter Registration Bills, 2015-Present." Brennan Center for Justice, 10 July 2019, www.brennancenter.org/our-work/research-reports/automatic-voter-registration-bills-2015-present.

"Automatic Voter Registration, a Summary." Brennan Center for Justice, 10 July 2019, www.brennancenter.org/our-work/research-reports/automatic-voter-registration-summary.

"Automatic Voter Registration." Alaska Department of Revenue | Permanent Fund Dividend Division, pfd.alaska.gov/Automatic-Voter-Registration.

"Automatic Voter Registration." IVote, www.ivoteforamerica.org/automatic_voter_registration.

"Automatic Voter Registration." National Conference of State Legislatures, www.ncsl.org/research/elections-and-campaigns/automatic-voter-registration.aspx.

Battaglia, Roman, and Sophia Schmidt. "Automatic Voter Registration Bill Passes State Senate." Delaware First Media, www.delawarepublic.org/post/automatic-voter-registration-bill-passes-state-senate.

Brater, Jonathan. "Automatic Voter Registration in Oregon a Huge Success." Brennan Center for Justice, 8 Apr. 2016, www.brennancenter.org/our-work/analysis-opinion/automatic-voter-registration-oregon-huge-success.

Brooks, James. "Alaska Has More Registered Voters than Eligible Residents." Juneau Empire, Juneau Empire, 1 Nov. 2018, www.juneauempire.com/news/alaska-has-more-registered-voters-than-eligible-residents/.

Brown, Robert D, and Justin Wedeking. "People Who Have Their Tickets But Do Not Use Them: 'Motor Voter,' Registration, and Turnout Revisited." American politics research 34.4 (2006): 479–504. Web.

Buchanan, Tyer. "Ohio House GOP Proposes Voting Law Changes." Review Times Here Are Some Key Components to the Bill and How the Changes Would Differ from Current Ohio Election Practices, reviewtimes.com/news/317601/ohio-house-gop-proposes-voting-law-changes/.

Contreras, Gonzalo, et al. "The Return of Censitary Suffrage? The Effects of Automatic Voter Registration and Voluntary Voting in Chile." Democratization, vol. 23, no. 3, 2015, pp. 520–544., doi:10.1080/13510347.2014.986720.

Cunningham, Dayna L. "Who Are to Be the Electors - A Reflection on the History of Voter Registration in the United States." Yale Law and Policy Review, 1991.

Sources

"Current Secretaries of State (Table)." Wikipedia, Wikimedia Foundation, 27 Apr. 2021, en.wikipedia.org/wiki/Secretary_of_state_(U.S._state_government).

"Election Administration and Voting Survey (EAVS) Comprehensive Report." Studies and Reports | U.S. Election Assistance Commission, www.eac.gov/research-and-data/studies-and-reports.

"Electronic Registration Information Center." Wikipedia, Wikimedia Foundation, 3 May 2021, en.wikipedia.org/wiki/Electronic_Registration_Information_Center.

"Former Governors." National Governors Association, www.nga.org/former-governors/.

Graves, Leslie. "Automatic Voter Registration." Ballotpedia, ballotpedia.org/Automatic_voter_registration.

Graves, Leslie. "Secretary of State (State Executive Office)." Ballotpedia, ballotpedia.org/Secretary_of_State_(state_executive_office).

Hanmer, Michael J. *Discount Voting: Voter Registration Reforms and Their Effects.* Cambridge University Press, 2012.

Highton, Benjamin. "Voter Registration and Turnout in the United States." Perspectives on Politics, vol. 2, no. 03, 2004, doi:10.1017/s1537592704040307.

"The Impact of the National Voter Registration Act of 1993 on the Administration of Elections for Federal Office 2009-2010 | A Report to the 112th Congress." United States Election Assistance Commission, 30 June 2011, www.eac.gov/sites/default/files/eac_assets/1/28/2010%20NVRA%20FINAL%20REPORT.pdf.

Jensen West Hawaii, Chelsea. "Lawmakers Pass Automatic Voter Registration Bill." West Hawaii Today, 28 Apr. 2021, www.westhawaiitoday.com/2021/04/28/hawaii-news/lawmakers-pass-automatic-voter-registration-bill/.

Kawashima-Ginsberg, Kei. Center for Information & Research on Civic Learning and Engagement, 2014, Voter Registration among Young People in Midterm Elections. https://circle.tufts.edu/sites/default/files/2019-12/FS_RegistrationInMidterms_2014.pdf

Keith, Theo. "Minnesota Senate GOP Looks to Limit Same-Day Voter Registration." FOX 9 Minneapolis-St. Paul, FOX 9 Minneapolis-St. Paul, 7 Apr. 2021, www.fox9.com/news/minnesota-senate-gop-looks-to-limit-same-day-voter-registration.

"Left Turn; Automatic Voter Registration." The Economist, vol. 417, no. 8965, Nov 21, 2015, pp. 26. ProQuest, https://search-proquest-com.dartmouth.idm.oclc.org/magazines/left-turn-automatic-voter-registration/docview/1735348931/se-2?accountid=10422.

Mann, Christopher B., et al. "Framing Automatic Voter Registration: Partisanship and Public Understanding of Automatic Voter Registration." American Politics Research, vol. 48, no. 6, 2020, pp. 693–699., doi:10.1177/1532673x20922525.

McDonald, Michael P. "Voter Turnout Data." United States Elections Project, University of Florida Department of Political Science, www.electproject.org/home/voter-turnout/voter-turnout-data.

Sources

Merivaki, Thessalia. *The Administration of Voter Registration* : Expanding the Electorate Across and Within the States . 1st ed. 2021. Cham, Switzerland: Palgrave Macmillan, 2021. Web.

Merivaki, Thessalia, and Daniel A Smith. “A Failsafe for Voters? Cast and Rejected Provisional Ballots in North Carolina.” Political research quarterly 73.1 (2020): 65–78. Web.

“Montana Governor Approves Ending Same-Day Voter Registration.” Flathead Beacon, Associated Press, 20 Apr. 2021, flatheadbeacon.com/2021/04/19/montana-governor-approves-ending-same-day-voter-registration/.

Nichanian, Daniel, et al. “The Voting Rights Manifesto: a State-by-State Plan to Defend Democracy.” Vox, Vox, 15 Dec. 2016, www.vox.com/the-big-idea/2016/12/15/13957714/voting-rights-states-ballot-franchise-civil-rights.

Niesse, Mark. “Georgia Senate Panel Backs End to Automatic Voter Registration, No-Excuse Absentee Voting.” Ajc, The Atlanta Journal-Constitution, 26 Feb. 2021, www.ajc.com/politics/senate-panel-backs-bill-to-end-automatic-voter-registration-in-georgia/HLQU3OUNKNDYPJHGXXWQRPMC7U/.

“Online Voter Registration.” National Conference of State Legislatures, www.ncsl.org/research/elections-and-campaigns/electronic-or-online-voter-registration.aspx.

“Oregon Motor Voter Registration Manual.” Oregon Secretary of State, Oregon Elections Division, Sept. 2017, sos.oregon.gov/voting/Documents/oregon-motor-voter-registration-manual-final.pdf.

“Partisan Composition of State Legislatures 1990-2000, 2002-2014.” National Conference of State Legislatures, www.ncsl.org/documents/statevote/legiscontrol_1990_2000.pdf, www.ncsl.org/documents/statevote/legiscontrol_2002_2014.pdf.

“Party Control of Nebraska State Government.” Ballotpedia, ballotpedia.org/Party_control_of_Nebraska_state_government.

Perez, Myrna, and Stephen Eide. “Should Voter Registration Be Automatic?” Junior Scholastic/Current Events, vol. 122, no. 11, 20 Apr. 2020.

Pérez, Myrna. “Voter Purges: Executive Summary.” Brennan Center for Justice at New York University School of Law, www.brennancenter.org/sites/default/files/2020-01/Purge.Exec_.pdf.

Rakich, Nathaniel. “What Happened When 2.2 Million People Were Automatically Registered To Vote.” FiveThirtyEight, FiveThirtyEight, 10 Oct. 2019, fivethirtyeight.com/features/what-happened-when-2-2-million-people-were-automatically-registered-to-vote/.

“Same Day Voter Registration.” National Conference of State Legislatures, www.ncsl.org/research/elections-and-campaigns/same-day-registration.aspx.

“Secretary of State’s Office Breaks 3.7 Million Registered Voters Ahead of November 3 General Election.” Alabama Secretary of State, 2 Nov. 2020, www.sos.alabama.gov/newsroom/secretary-states-office-breaks-37-million-registered-voters-ahead-november-3-general.

Sources

St. Louis Public Radio | By Jason Rosenbaum. “Election Officials Say Online Registration Expanded Voter Rolls For Presidential Primary.” St. Louis Public Radio, 4 Aug. 2020, news.stlpublicradio.org/politics-issues/2020-02-16/election-officials-say-online-registration-expanded-voter-rolls-for-presidential-primary.

Staff. “Automatic Voter Registration Grows Rolls a Lot, Study Finds.” The Fulcrum, The Fulcrum, 26 Dec. 2019, thefulcrum.us/automatic-voter-registration-2.

“State Partisan Composition.” National Conference of State Legislatures, www.ncsl.org/research/about-state-legislatures/partisan-composition.aspx.

“States That Have Enacted Voter ID Laws, Strict vs. Non-Strict Classification.” National Conference of State Legislatures, www.ncsl.org/documents/LSSS/VoterIDdoc.pdf.
Thistle, Scott. “Maine Bill Would Allow Online Voter Registration by 2023.” Governing, Governing, 21 Apr. 2021, www.governing.com/next/maine-bill-would-allow-online-voter-registration-by-2023.html.

Stewart, Charles, and Stephen Ansolabehere. “Waiting to Vote.” Election law journal 14.1 (2015): 47–53. Web.

“Sununu Vetoes Democratic Voter Verification Overhaul .” Concord Monitor, Concord Monitor, 27 Sept. 2019, www.concordmonitor.com/New-Hampshire-Gov-Chris-Sununu-vetoes-CrossCheck-voting-replacement-ERIC-28871212.

“United States : Connecticut Department of Motor Vehicles Launches Automatic Voter Registration Initiative.” MENA Report, 2016. ProQuest, https://search-proquest-com.dartmouth.idm.oclc.org/trade-journals/united-states-connecticut-department-motor/docview/1817710192/se-2?accountid=10422.

“United States : Group of Former Aides to Obama, Clinton Pledges to Invest $10M to make Voter Registration Automatic.” MENA Report, 2015. ProQuest, https://search-proquest-com.dartmouth.idm.oclc.org/trade-journals/united-states-group-former-aides-obama-clinton/docview/1732709024/se-2?accountid=10422.

“United States : Illinois Senate Democrats Propose Plan for Automatic Voter Registration.” MENA Report, 2015. ProQuest, https://search-proquest-com.dartmouth.idm.oclc.org/trade-journals/united-states-illinois-senate-democrats-propose/docview/1724084854/se-2?accountid=10422.

“U.S. President 1976–2020.” MIT Election Data Science Lab, Harvard Dataverse, 13 Jan. 2021, dataverse.harvard.edu/dataset.xhtml?persistentId=doi%3A10.7910%2FDVN%2F42MVDX.

“Voter Identification Laws in the United States | Laws by State (Table).” Wikipedia, Wikimedia Foundation, 24 Apr. 2021, en.wikipedia.org/wiki/Voter_identification_laws_in_the_United_States#Laws_by_state.

“Voter Identification Requirements: Voter ID Laws.” National Conference of State Legislatures, www.ncsl.org/research/elections-and-campaigns/voter-id.aspx#Two.

“Voter List Accuracy.” National Conference of State Legislatures, www.ncsl.org/research/elections-and-campaigns/voter-list-accuracy.aspx.

“Voter Registration & Voter Statistics.” Alaska Division of Elections, www.elections.alaska.gov/doc/info/statsAGE.php.

Sources

"Voter Registration Deadlines." Vote.org, www.vote.org/voter-registration-deadlines/.

"Voter Registration Deadlines." Vote.org, www.vote.org/voter-registration-deadlines/.

"Voter Registration." Ballotpedia, ballotpedia.org/Voter_registration.

Walker, Hannah L., et al. "Early Voting Changes and Voter Turnout: North Carolina in the 2016 General Election." Political Behavior, vol. 41, no. 4, Dec. 2019, pp. 841–869. EBSCOhost, doi:10.1007/s11109-018-9473-5.

"Who We Are." Electronic Registration Information Center, ericstates.org/who-we-are/.

Yu, Jinhai. "Does State Online Voter Registration Increase Voter Turnout?*." Social Science Quarterly, vol. 100, no. 3, 2019, pp. 620–634., doi:10.1111/ssqu.12598.

Appendix A: Covariance between variables

Before proceeding with an OLS multivariable regression, strong covariance between multiple variables in the model should be dismissed. Covariance between selected varables is shown in the table below:

Correlation between selected variables **TABLE A**

Variables are selected based on qualitative logical links; Acknowledgement: Professor J. Ferwerda

Variable 1	Variable 2	Covariance
AVR (DMV)	*AVR (Other)*	0.6216727
AVR (DMV)	*Republican Governor*	-0.09094705
Republican-controlled Legislature	*Republican Governor*	0.3667954
SDR	*ERIC*	0.1652187
ERIC	*Lagged ERIC*	0.06225435
Competitiveness (Absolute Margin)	*ERIC*	-0.04616273

Appendix B: Bivariate Regressions with Lagged Variables

Before proceeding with an OLS multivariable regression, strong covariance between multiple variables in the model should be dismissed. Covariance between selected varables is shown in the table below:

Preliminary Examination of Lagged Variable Significance **TABLE B**

Bivariate regressions include VAR, Lagged VAR, and factors for State and Year as controls

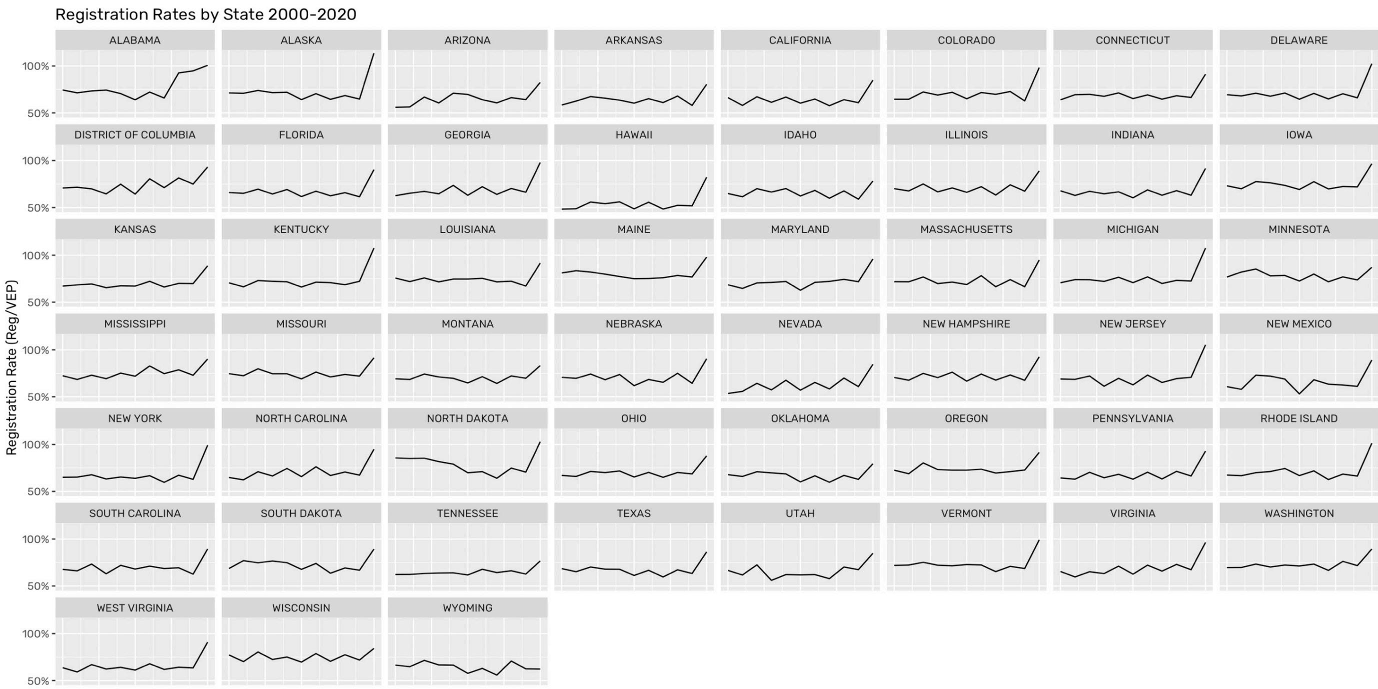
Lagged Variable	**Pr(>|*t*|)**	
Removals/VEP	0.867203	
AVR (DMV)	0.982710	
AVR (Other)	0.011675	*
OVR	0.673916	
SDR	0.383947	
ERIC	0.930197	
ID Requirements	0.732980	
Republican Governor	0.941572	
Republican-controlled Legislature	0.348880	
Republican Secretary of State	0.508944	

Significance codes: 0 = ‘ *** ’ | 0.001 = ‘ ** ’ | 0.01 = ‘ * ’ | 0.1 = ‘ . ’

Although lagged AVR (Other) is significant according to this analysis, non-lagged AVR (Other) is far more significant, with a *p*-value of 0.003668. This analysis shows that adding lags to the main analysis is unnecessary.

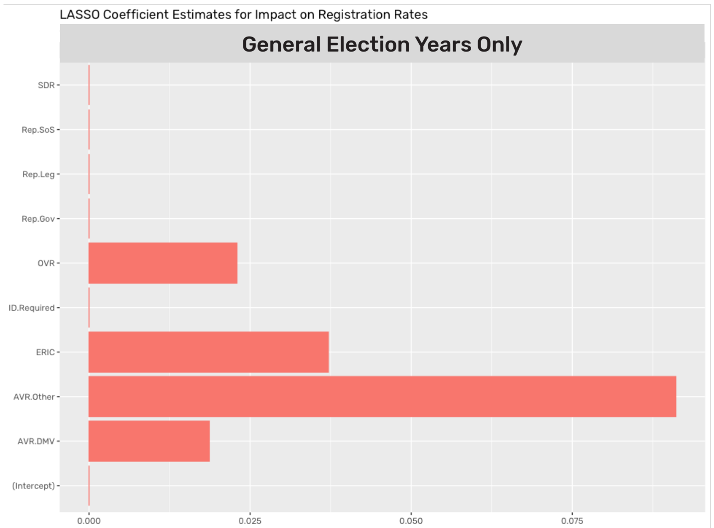
Appendix C: Registration Rates 2000-2020 by State

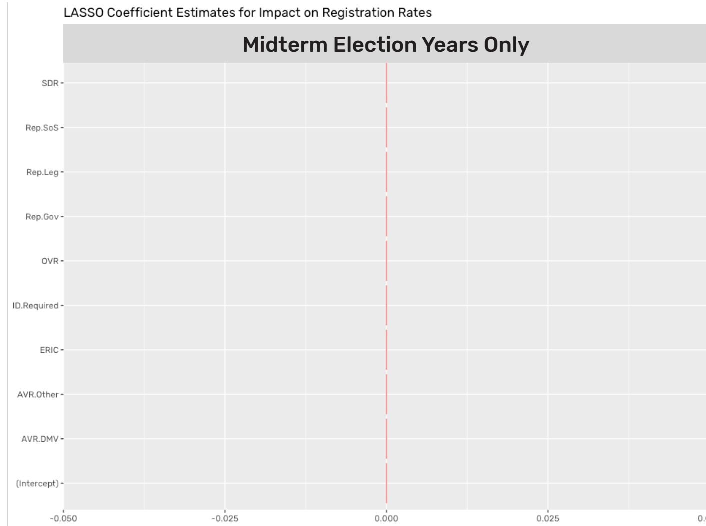
FIGURE C

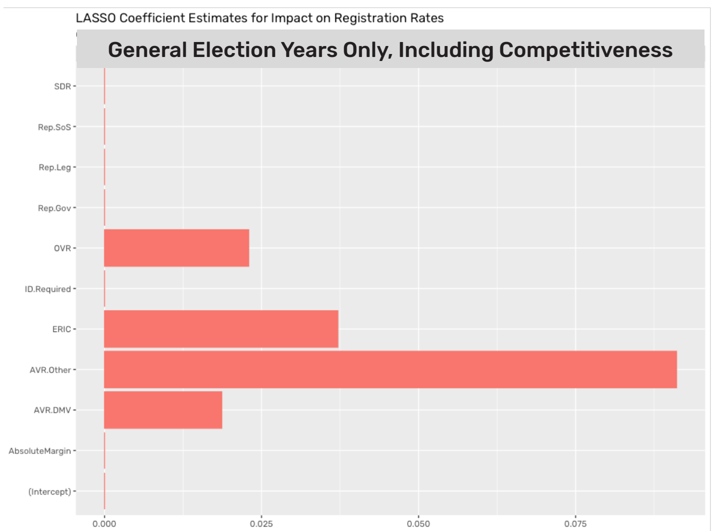


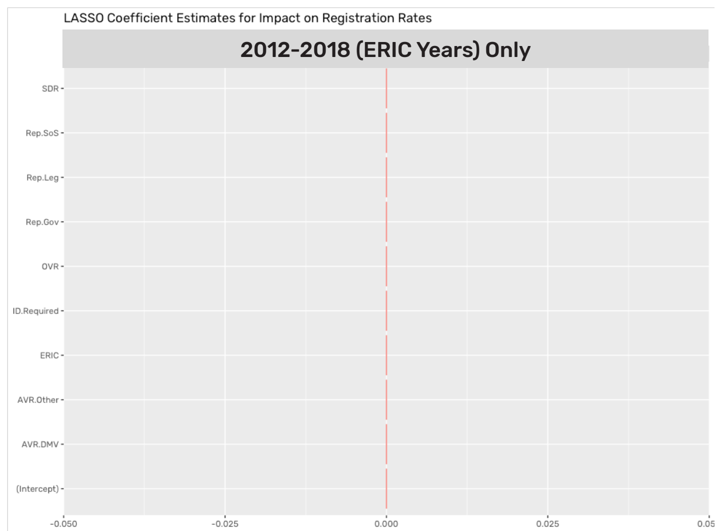
Appendix D: LASSO results with additional controls

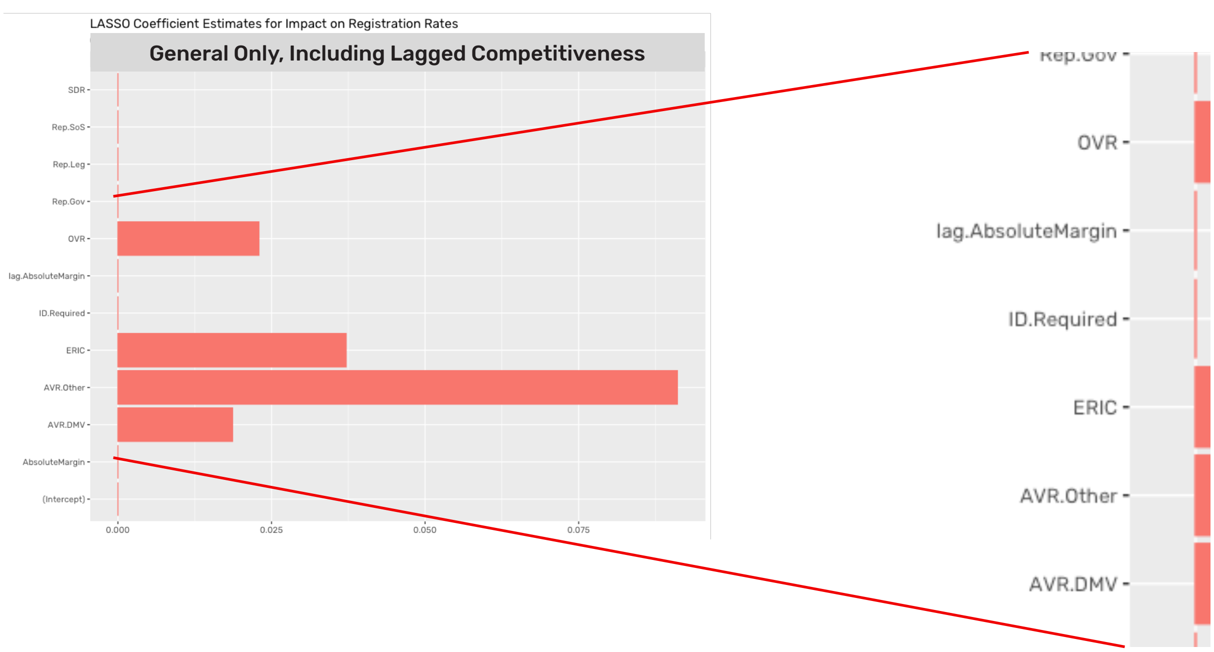
FIGURE SET D











Appendix E: Data Deep Dive

Washington, D.C. in the Data:

For the master dataset, Washington, D.C. is included as a state. We treat D.C. as a state because it apportions electoral votes in national elections. For the column "Governor," we will use the Mayor of D.C. For the column "Legislature Control," we will use the partisanship of City Council members.

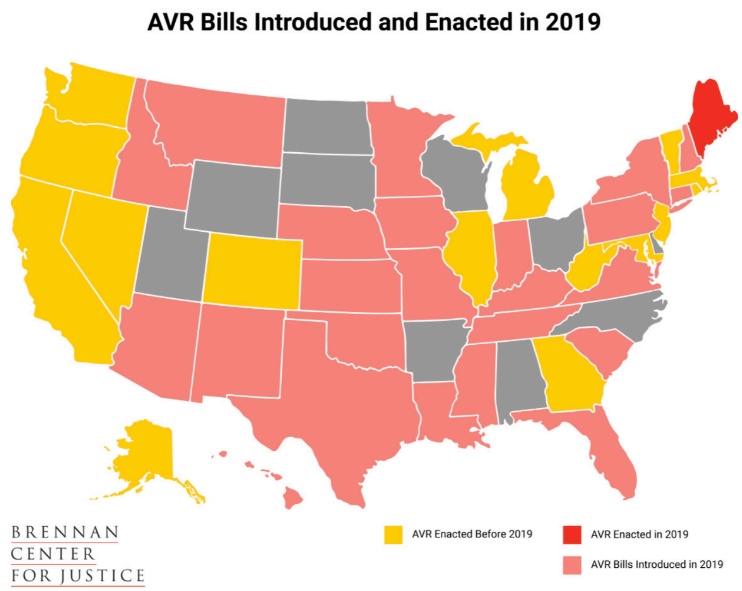
Nebraska in the Data:

Nebraska's state legislature is technically nonpartisan, but we determined legislature control by looking at the party affiliation of each member of the legislature.

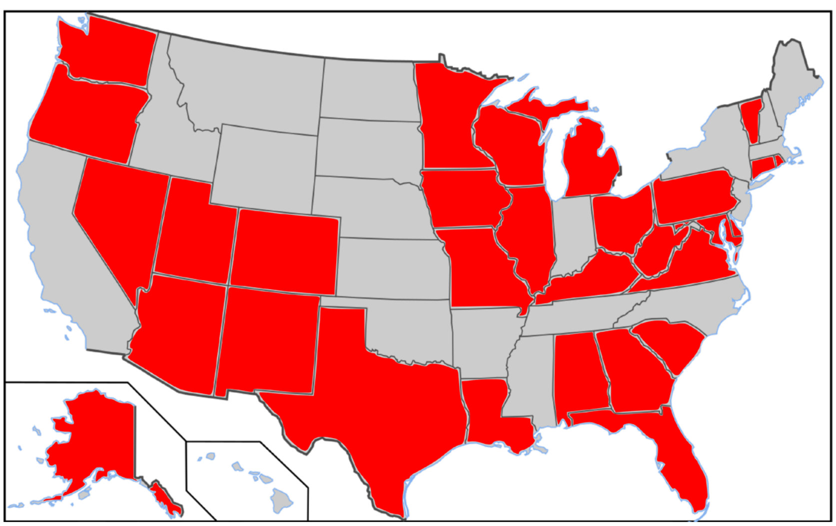
When there is no Secretary of State:

Alaska, Hawaii, and Utah don't have a Secretary of State. In those state, the Lieutenant Governor runs elections, and so the Lieutenant Governor's partisanship will be used for the "Secretary of State" column in states that don't have a State Secretary.

Appendix F: Recent AVR Efforts



Appendix G: ERIC Member States

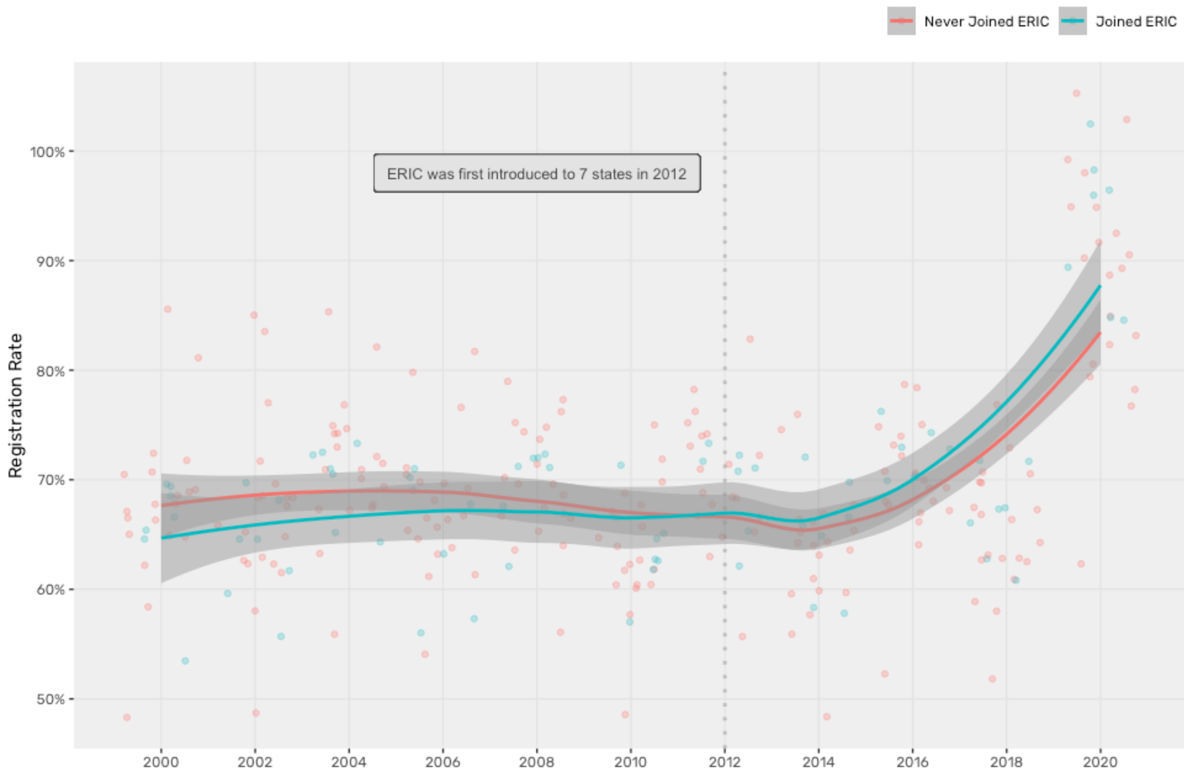


Appendix H: Differences in Differences Visualizations

FIGURE SET H

Registration Rates | Joined ERIC in 2012 vs. Did Not Join ERIC

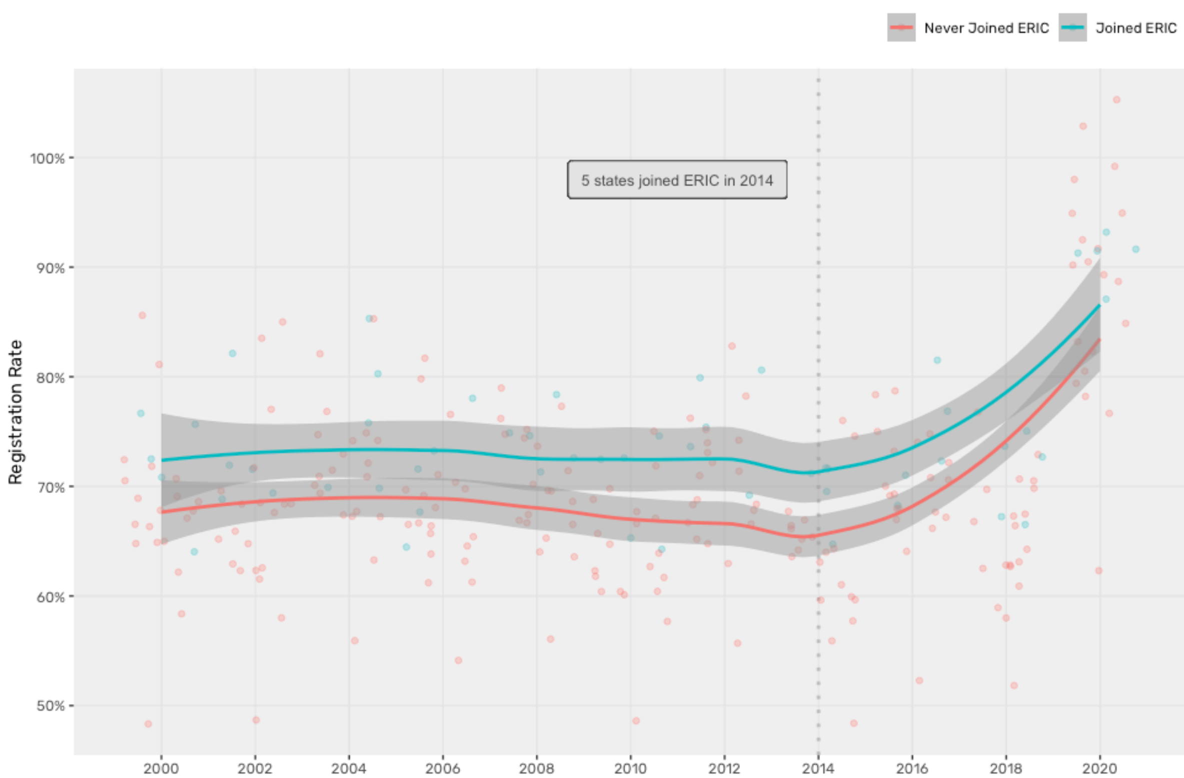
States that joined ERIC, but Not in 2012, are Excluded



Joined "in" YEAR means that the state joined immediately prior to that election year.

Registration Rates | Joined ERIC in 2014 vs. Did Not Join ERIC

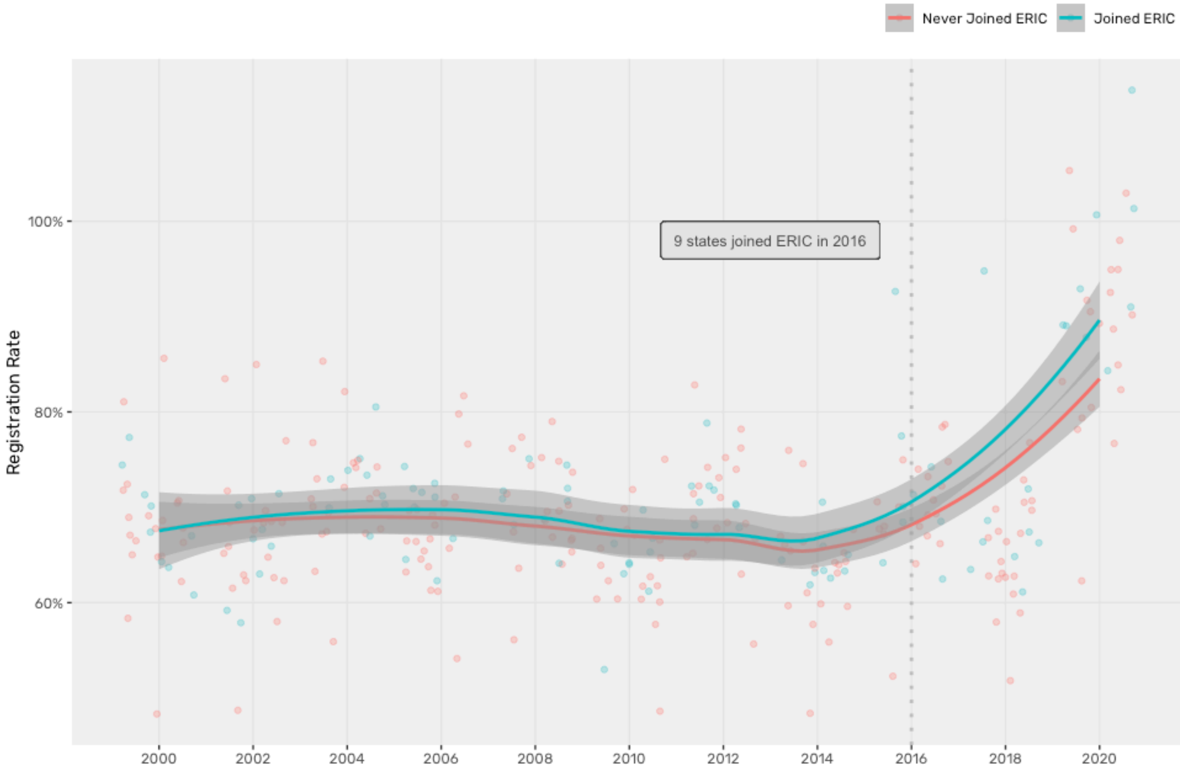
States that joined ERIC, but Not in 2014, are Excluded



Joined "in" YEAR means that the state joined immediately prior to that election year.

Registration Rates | Joined ERIC in 2016 vs. Did Not Join ERIC

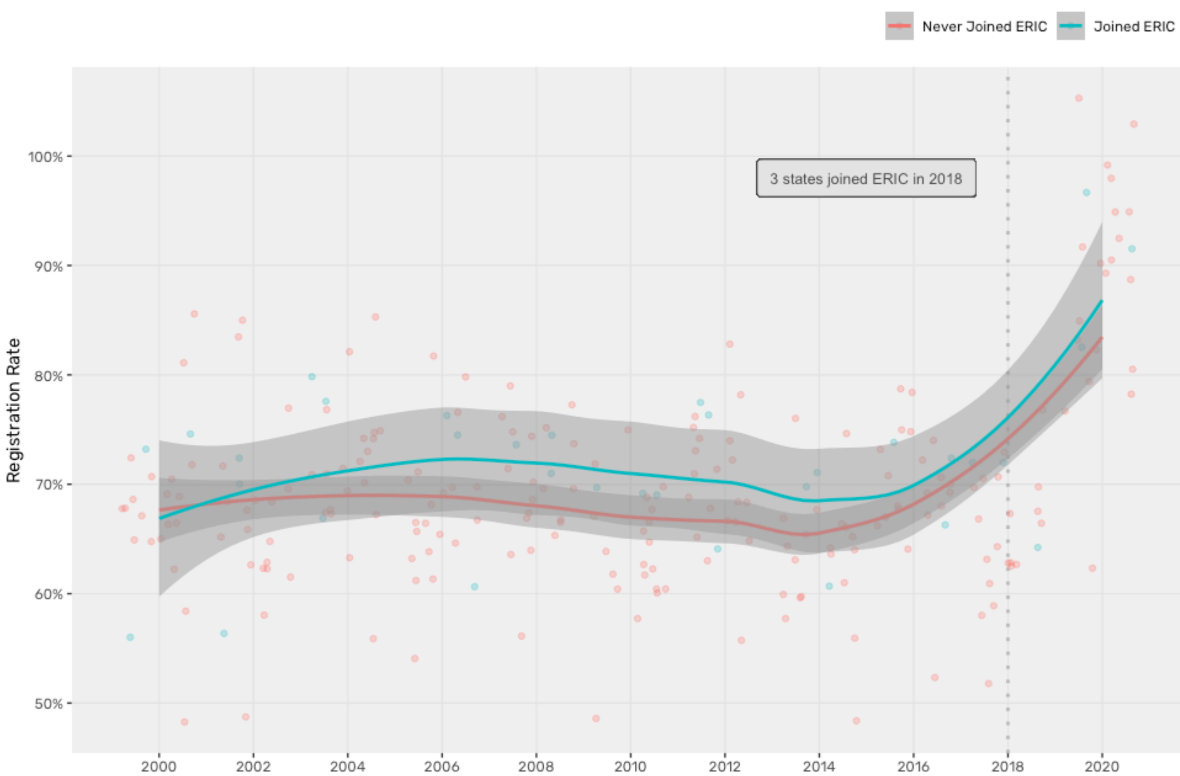
States that joined ERIC, but Not in 2016, are Excluded



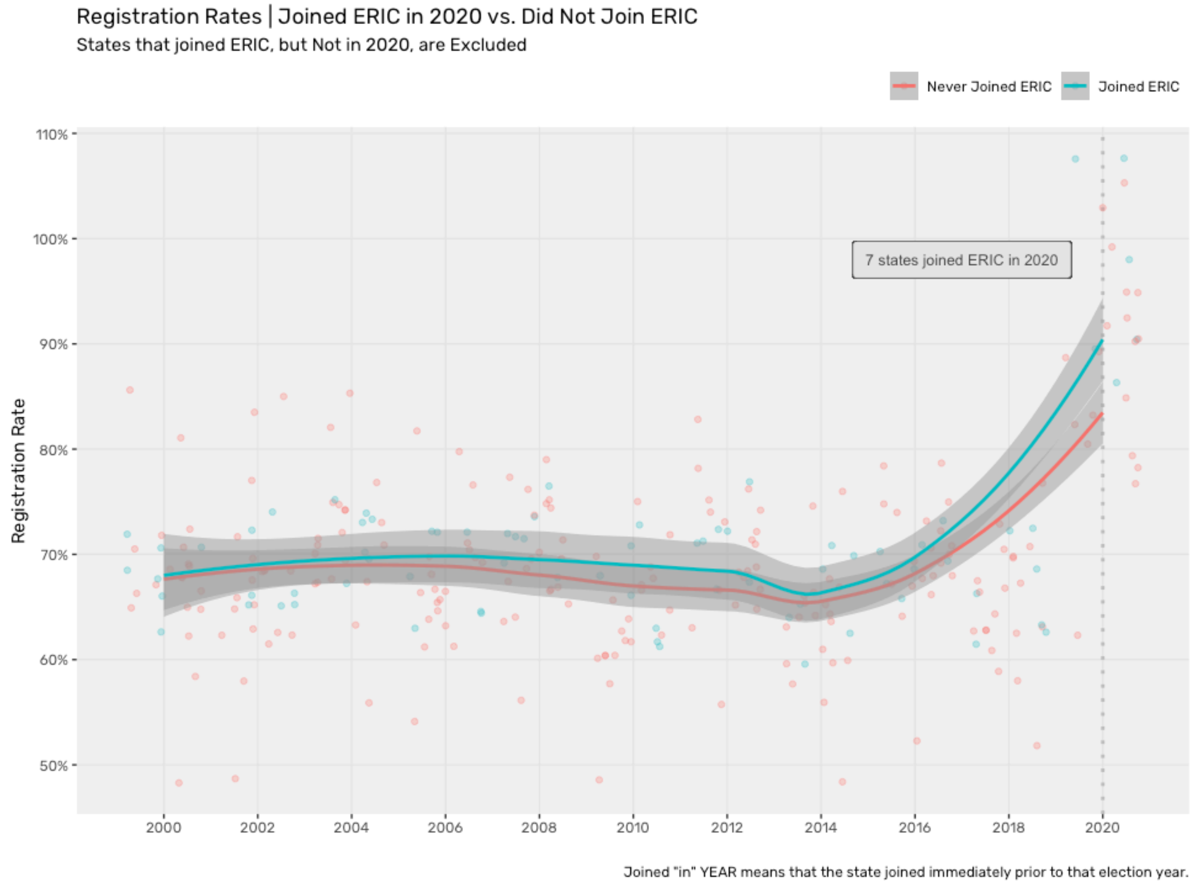
Joined "in" YEAR means that the state joined immediately prior to that election year.

Registration Rates | Joined ERIC in 2018 vs. Did Not Join ERIC

States that joined ERIC, but Not in 2018, are Excluded

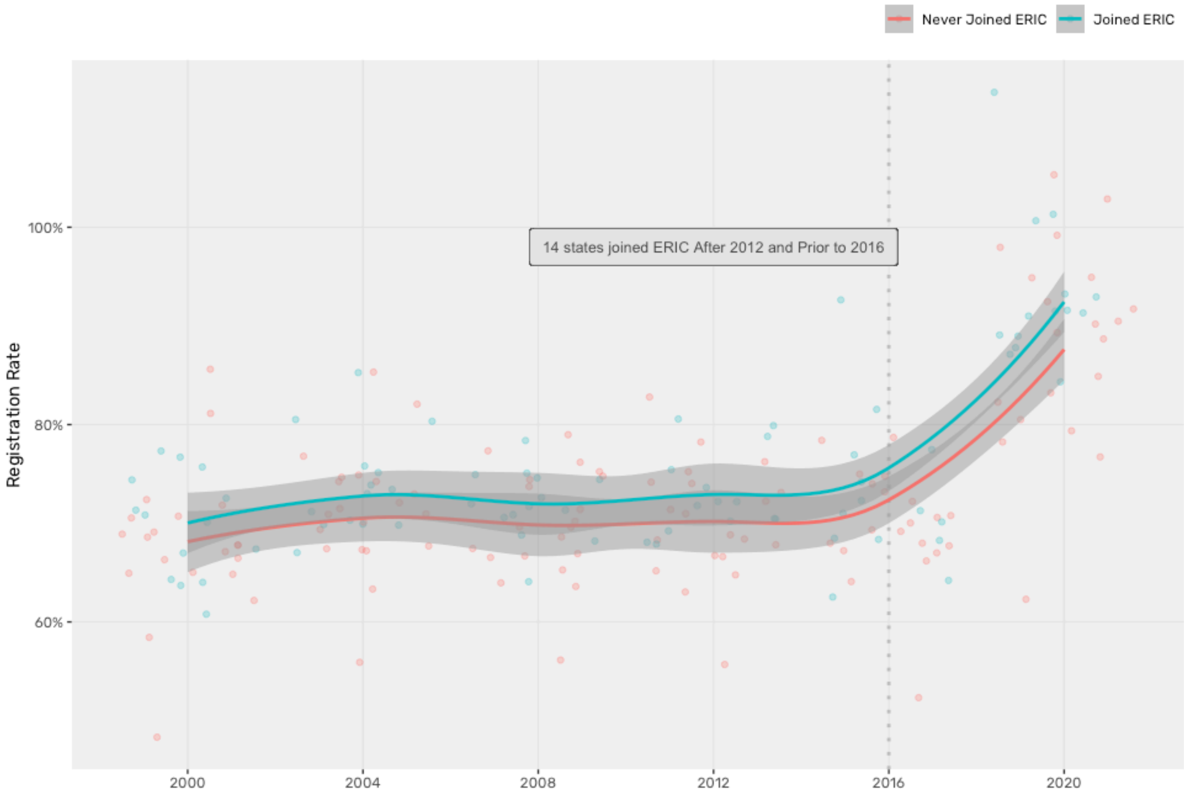


Joined "in" YEAR means that the state joined immediately prior to that election year.

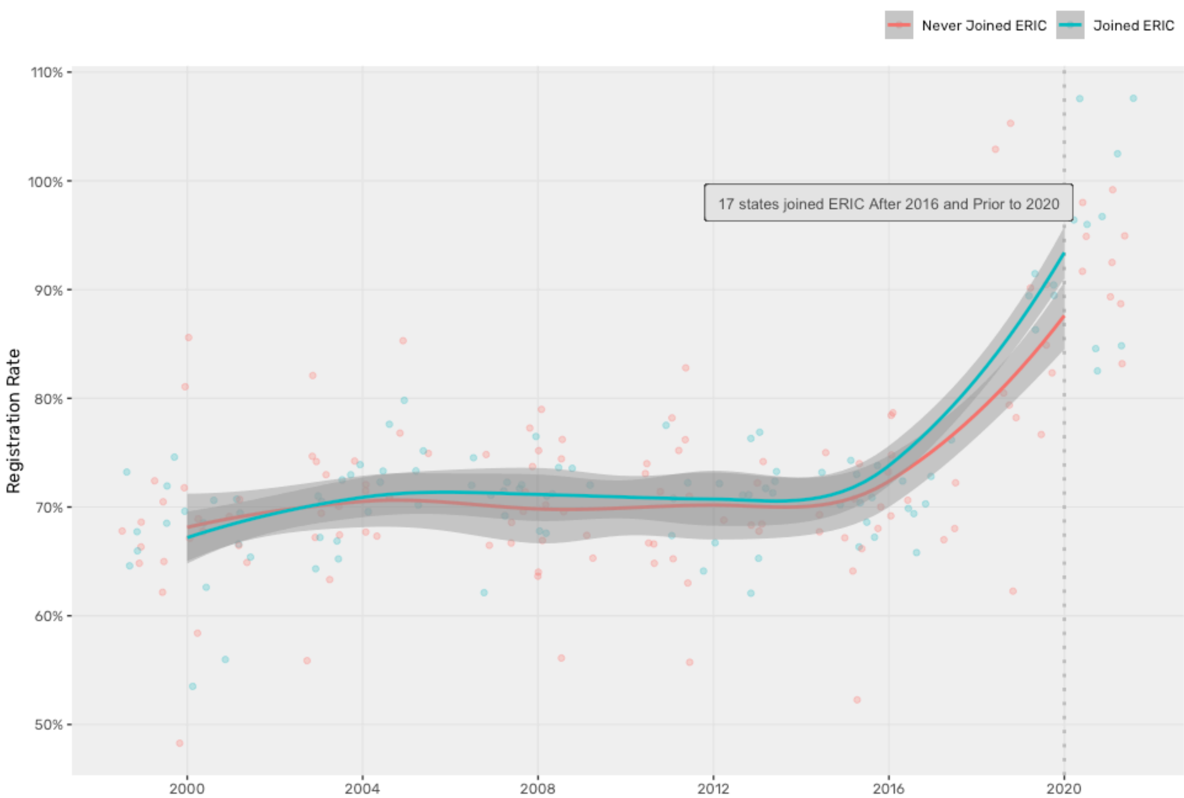


Registration Rates | Joined ERIC in 2020 vs. Did Not Join ERIC
States that joined ERIC, but Not in 2020, are Excluded
Never Joined ERIC
Joined ERIC
7 states joined ERIC in 2020
Registration Rate
110%
100%
90%
80%
70%
60%
50%
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
Joined "in" YEAR means that the state joined immediately prior to that election year.

Registration Rates | Joined ERIC 2012-2016 vs. Did Not Join ERIC | General Elections Only
States that joined ERIC, but Not in 2016, are Excluded



Registration Rates | Joined ERIC After 2016, Before 2020 vs. Did Not Join ERIC
General Election Years Only



Appendix X: R Code

R code for my analysis:

Below is the R code I used to conduct my analysis. To access my master dataset, please use this link: https://docs.google.com/spreadsheets/d/1qkRuo1PMIkZc3bX1An4zSZzOFRI8Q6YVFtBoj9iqDrU/edit?usp=sharing

```
###############################################################################
########################### QSS HONORS THESIS #################################
###############################################################################
#########################      EITAN DARWISH       ############################
###############################################################################
# Data Analysis File ##########################################################
# March 4th 2020 ##############################################################
###############################################################################

# LOAD LIBRARIES -------------------------------------------------------------
library(data.table)
library(dplyr)
library(ggplot2)
library(tidyverse)
library(plm)
library(glmnet)
library(broom)

# LOAD DATA ------------------------------------------------------------------
df <- read.csv(file = "MASTER_DATA_3_15.csv", header = TRUE, sep = ",")

# Competitiveness Dataset
comp <- read.csv(file="1976-2020-president.csv", header = TRUE, sep = ",")
comp <- comp %>% filter(year >= 1996) %>%
  mutate(YEAR = year) %>% mutate(STATE = state) %>%
  select(YEAR, STATE, party_simplified, candidatevotes, totalvotes) %>%
  filter(party_simplified == "DEMOCRAT") %>%
  mutate(DemVotesPercent = as.numeric(candidatevotes/totalvotes)) %>%
  mutate(AbsoluteMargin = abs(DemVotesPercent-0.5)) # uses democrat votes as anchor
comp <- comp %>% select(YEAR, STATE, AbsoluteMargin)
comp <- comp %>% distinct(YEAR, STATE, .keep_all = TRUE)
## Add lagged competitiveness
comp <- comp %>% mutate(lag.AbsoluteMargin = dplyr::lag(AbsoluteMargin, n=51,
                                                         default=NA)) %>%
  filter(YEAR >= 2000)

## SUBSET MASTER DATASET
#### 2012 - 2018
df_ERIC_Years <- df %>% filter(YEAR < 2020) %>% filter(YEAR >= 2012)
#### Midterm Election Years
df_midterm <- df %>% filter(YEAR == 2002 | YEAR == 2006 | YEAR == 2010 |
                              YEAR == 2014 | YEAR == 2018)
#### General Election Years
# merge with competitiveness (absolute margin of victory/defeat)
df_general <- merge(df, comp, by=c("YEAR","STATE"))
```

```r
# PLOT DATA ---------------------------------------------------------------
## Facet wrapped by state
ggplot(data = df, mapping = aes(x = YEAR, y = Registered.VEP)) +
  geom_line() + facet_wrap(~STATE) + ggtitle(df$STATE) +
  ggtitle("Registration Rates by State 2000-2020") +
  theme(text=element_text(family="Rubik")) +
  theme(axis.title.x = element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank()) +
  scale_y_continuous(name="Registration Rate (Reg/VEP)", breaks=c(0.5, 1),
                     labels=c("50%", "100%"))

## All in one graph
ggplot(data = df, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates by State 2000-2020") +
  theme(legend.position = "none") +
  theme(text=element_text(family="Rubik")) +
  geom_line(aes(color=STATE)) +
  scale_x_continuous(breaks=c(2000,2002,2004,2006,2008,2010,2012,
                              2014,2016,2018,2020)) +
  scale_y_continuous(labels = scales::percent) +
  xlab("") +
  ylab("Registration Rate") +
  scale_color_hue() +
  # callouts
  geom_label(
    label = "Alabama implements OVR in prior to 2016 election",
    x = 2010,
    y = 0.98,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    label.size = 0.35,
    color = "#f55442",
    fill = "#fcedeb"
  ) +
  geom_label(
    label = "COVID-19 pandemic prompts widespread ease on voting restrictions in 202
    x = 2010.5,
    y = 1.1,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    label.size = 0.35,
    color = "#4a4a4a",
    fill = "#e6e6e6"
  ) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0)
  ) +
  labs(caption = "Fluctuations in registration rates every two years are driven by t
```

```
# PRELIMINARY ANALYSIS ---------------------------------------------------------
## Significance Checks with Single-Variable Regression ---------------------
### Removals ################ p=0.162
result <- lm(Registered.VEP ~ Removal.VEP + factor(STATE) + factor(YEAR),data=df)
summary(result)
### AVR | DMV ############### p=0.104
result <- lm(Registered.VEP ~ AVR.DMV + factor(STATE) + factor(YEAR),data=df)
summary(result)
### AVR | Other ############ p=0.0045 **
result <- lm(Registered.VEP ~ AVR.Other + factor(STATE) + factor(YEAR),data=df)
summary(result)
### AVR | Other | midterm elections
result <- lm(Registered.VEP ~ AVR.Other + factor(STATE) + factor(YEAR),data=df_midte
summary(result) # Not significant
### OVR #################### p=0.0000 ***
result <- lm(Registered.VEP ~ OVR + factor(STATE) + factor(YEAR),data=df)
summary(result)
### OVR | midterm elections. still significant, but estimate decreases
result <- lm(Registered.VEP ~ OVR + factor(STATE) + factor(YEAR),data=df_midterm)
summary(result)
### SDR #################### p=0.0880 .
result <- lm(Registered.VEP ~ SDR + factor(STATE) + factor(YEAR),data=df)
summary(result)
### ID.Required ############ p=0.5983
result <- lm(Registered.VEP ~ ID.Required + factor(STATE) + factor(YEAR),data=df)
summary(result)
### ERIC ################### p=0.0000 ***
result <- lm(Registered.VEP ~ ERIC + factor(STATE) + factor(YEAR),data=df)
summary(result)
### ERIC | midterm elections. still significant, but estimate slightly decreases
result <- lm(Registered.VEP ~ ERIC + factor(STATE) + factor(YEAR),data=df_midterm)
summary(result)
### Republican Governor ##### p=0.522
result <- lm(Registered.VEP ~ Rep.Gov + factor(STATE) + factor(YEAR),data=df)
summary(result)
### Republican Legislature ## p=0.181
result <- lm(Registered.VEP ~ Rep.Leg + factor(STATE) + factor(YEAR),data=df)
summary(result)
### Rep SoS ################# p=0.972
result <- lm(Registered.VEP ~ Rep.SoS + factor(STATE) + factor(YEAR),data=df)
summary(result)
```

```
### Lagged Variables, Bivariate Regression --------------------------------
df_lag <- plm::pdata.frame(df,index=c('STATE','YEAR'))
#### Lagged AVR | DMV #######
df_lag$lag_Removal.VEP <- plm::lag(df$Removal.VEP,2)
result <- lm(Registered.VEP ~ Removal.VEP + lag_Removal.VEP + factor(STATE) + factor(YEAR),data=df_lag)
summary(result)
#### Lagged AVR | DMV #######
df_lag$lag_AVR.DMV <- plm::lag(df$AVR.DMV,2)
result <- lm(Registered.VEP ~ AVR.DMV + lag_AVR.DMV + factor(STATE) + factor(YEAR),data=df_lag)
summary(result)
#### Lagged AVR | Other #####
df_lag$lag_AVR.Other <- plm::lag(df$AVR.Other,2)
result <- lm(Registered.VEP ~ AVR.Other + lag_AVR.Other + factor(STATE) + factor(YEAR),data=df_lag)
summary(result)
#### Lagged OVR #############
df_lag$lag_OVR <- plm::lag(df$OVR,2)
result <- lm(Registered.VEP ~ OVR + lag_OVR + factor(STATE) + factor(YEAR),data=df_lag)
summary(result)
#### Lagged SDR #############
df_lag$lag_SDR <- plm::lag(df$SDR,2)
result <- lm(Registered.VEP ~ SDR + lag_SDR + factor(STATE) + factor(YEAR),data=df_lag)
summary(result)
#### Lagged ERIC ############
df_lag$lag_ERIC <- plm::lag(df$ERIC,2)
result <- lm(Registered.VEP ~ ERIC + lag_ERIC + factor(STATE) + factor(YEAR),data=df_lag)
summary(result)
#### Lagged ID.Required #####
df_lag$lag_ID.Required <- plm::lag(df$ID.Required,2)
result <- lm(Registered.VEP ~ ID.Required + lag_ID.Required + factor(STATE) + factor(YEAR),data=df_lag)
summary(result)
#### Lagged Rep Gov #########
df_lag$lag_Rep.Gov <- plm::lag(df$Rep.Gov,2)
result <- lm(Registered.VEP ~ Rep.Gov + lag_Rep.Gov + factor(STATE) + factor(YEAR),data=df_lag)
summary(result)
#### Lagged Rep Leg #########
df_lag$lag_Rep.Leg <- plm::lag(df$Rep.Leg,2)
result <- lm(Registered.VEP ~ Rep.Leg + lag_Rep.Leg + factor(STATE) + factor(YEAR),data=df_lag)
summary(result)
#### Lagged Rep SoS #########
df_lag$lag_Rep.SoS <- plm::lag(df$Rep.SoS,2)
result <- lm(Registered.VEP ~ Rep.SoS + lag_Rep.SoS + factor(STATE) + factor(YEAR),data=df_lag)
summary(result)
```

```
## Correlation between variables ----------------------------------------

cor(df$AVR.DMV,df$AVR.Other,use='complete.obs')                # 0.6217
cor(df$AVR.DMV,df$Rep.Gov,use='complete.obs')                  # -.091
cor(df$AVR.DMV,df$ERIC,use='complete.obs')                     # 0.038
cor(df$Rep.Leg,df$Rep.Gov,use='complete.obs')                  # 0.367
cor(df$ERIC,df$SDR,use='complete.obs')                         # 0.165
cor(df$ERIC,df_lag$ERIC,use='complete.obs')                    # 0.062
cor(df_general$AbsoluteMargin,df_general$ERIC,use='complete.obs') # -0.046
```

```r
## Subset Analysis ------------------------------------------------------------
## How much of a difference does the removal data make?

df_sub <- df %>% filter(YEAR >= 2010) %>% filter(YEAR <= 2018)

result <- lm(Registered.VEP
              ~ Removal.VEP + AVR.DMV + AVR.Other + OVR + SDR + ERIC + Rep.Gov + Rep.Leg + Rep.SoS
              + factor(STATE) + factor(YEAR),data=df_sub)
with_removal <- write.csv(tidy(result), "w_removal_sub.csv")
with_removal <- read.csv("w_removal_sub.csv")
names(with_removal)[3] <- "WithRemovalEstimate"

result2 <- lm(Registered.VEP
              ~ AVR.DMV + AVR.Other + OVR + SDR + ERIC + Rep.Gov + Rep.Leg + Rep.SoS
              + factor(STATE) + factor(YEAR),data=df_sub)
without_removal <- write.csv(tidy(result2), "no_removal_sub.csv")
without_removal <- read.csv("no_removal_sub.csv")
names(without_removal)[3] <- "NoRemovalEstimate"

comp <- merge(with_removal, without_removal, by=c("term"))
comp <- comp %>% mutate(diff = WithRemovalEstimate - NoRemovalEstimate) %>%
  select(term, WithRemovalEstimate, NoRemovalEstimate, diff)
print(head(comp, 4))
# All other variables are state and year factors
print(tail(comp, 5))

################################################################################
## OLS Multivariable Regressions -----------------------------------------------
################################################################################
## All Years
resultLM <- lm(Registered.VEP
               ~ AVR.DMV + AVR.Other + OVR + SDR + ERIC + Rep.Gov + Rep.Leg + Rep.SoS
               + ID.Required + factor(STATE) + factor(YEAR),data=df)
summary(resultLM) ## ERIC, OVR, AVR Other are significant (RepLeg less so)

# Midterm Years Only
resultLM <- lm(Registered.VEP
               ~ AVR.DMV + AVR.Other + OVR + SDR + ERIC + Rep.Gov + Rep.Leg + Rep.SoS
               + ID.Required + factor(STATE) + factor(YEAR),data=df_midterm)
summary(resultLM) ## None are significant at the 0.05 level

## OLS From 2012-2018 ----------------------------------------------------------
resultLM_ERIC_Years <- lm(Registered.VEP
               ~ AVR.DMV + AVR.Other + OVR + SDR + ERIC + Rep.Gov + Rep.Leg + Rep.SoS
               + ID.Required + factor(STATE) + factor(YEAR),data=df_ERIC_Years)
summary(resultLM_ERIC_Years) ## ERIC and SDR are significant

## OLS General Elections Only --------------------------------------------------
resultLM_General <- lm(Registered.VEP
                          ~ AVR.DMV + AVR.Other + OVR + SDR + ERIC + Rep.Gov + Rep.Leg + Rep.SoS
                          + ID.Required + factor(STATE) + factor(YEAR),data=df_general)
summary(resultLM_General) ## AVR (Other), OVR, and ERIC (less so) are significant

## OLS General Elections Only, with Competitiveness ----------------------------
resultLM_General_comp <- lm(Registered.VEP
                       ~ AVR.DMV + AVR.Other + OVR + SDR + ERIC + Rep.Gov + Rep.Leg + Rep.SoS
                       + ID.Required + AbsoluteMargin + factor(STATE) + factor(YEAR),data=df_general)
summary(resultLM_General_comp) ## AVR (Other), OVR, and ERIC (less so) are significant
```

```
################################################################################
################################################################################
# LASSO Models -------------------------------------------------------------
################################################################################
# Creating a function ------------------------------------------------------
run_lasso <- function(lasso_dataframe, subtitle_lasso) {
  x = model.matrix(Registered.VEP~.,lasso_dataframe) # matrix of predictors
  y = lasso_dataframe$Registered.VEP                  # vector y values

  set.seed(123)                                     # replicable results

  lasso_model <- cv.glmnet(x, y, alpha=1)           # alpha=1 is lasso

  best_lambda_lasso <- lasso_model$lambda.1se       # largest lambda in 1 SE

  lasso_coef <- lasso_model$glmnet.fit$beta[,       # retrieve coefficients
                                            lasso_model$glmnet.fit$lambda    # at lambda.1se
                                            == best_lambda_lasso]
  coef_l = data.table(lasso = lasso_coef)           # build table
  coef_l[, feature := names(lasso_coef)]            # add feature names
  to_plot_l = melt(coef_l                           # label table
                   , id.vars='feature'
                   , variable.name = 'model'
                   , value.name = 'coefficient')
  print(to_plot_l)
  ggplot(data=to_plot_l,                            # plot coefficients
         aes(x=feature, y=coefficient,
             fill=model, color=model)) +
    theme(text=element_text(family="Rubik"),
          legend.position = "none",
          axis.title.x = element_blank(),
          axis.title.y = element_blank()) +
    ggtitle("LASSO Coefficient Estimates for Impact on Registration Rates") +
    labs(subtitle=subtitle_lasso) +
    scale_color_hue() +
    coord_flip() +
    geom_bar(stat='identity') +
    facet_wrap(~ model) + guides(fill=FALSE)
}
```

```
# LASSO | Preliminary -----------------------------------------------------
df_lasso <- df %>% select(Registered.VEP, ERIC, SDR, OVR,
                          AVR.DMV, AVR.Other,
                          ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)

run_lasso(df_lasso, "2000-2020") ## ERIC, AVR.Other, and AVR.DMV explain variation in registratic
# -------------------------------------------------------------------------
# LASSO | 2012-2018 (ERIC Years) Only -------------------------------------
df_lasso2 <- df_ERIC_Years %>% select(Registered.VEP, ERIC, SDR, OVR,
                          AVR.DMV, AVR.Other,
                          ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso2, "2012-2018 (ERIC Years) Only") ## No variables are significant
# -------------------------------------------------------------------------
# LASSO | 2012 - 2018 (ERIC Years) Only | Including Removals --------------
df_lasso3 <- df_ERIC_Years %>% select(Registered.VEP, ERIC, SDR, OVR,
                          AVR.DMV, AVR.Other,
                          ID.Required, Rep.Gov, Rep.Leg, Rep.SoS, Removal.VEP)
run_lasso(df_lasso3, "2012-2018 (ERIC Years) Only | Including Removals") ## No variables are sigr
# -------------------------------------------------------------------------
# LASSO | General Election Years Only -------------------------------------
df_lasso4 <- df_general %>% select(Registered.VEP, ERIC, SDR, OVR,
                                    AVR.DMV, AVR.Other,
                                    ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso4, "General Election Years Only")
## AVR.DMV, OVR, ERIC, and AVR.Other are all significant
# -------------------------------------------------------------------------
# LASSO | General Election Years Only | With Competitiveness -------------
df_lasso5 <- df_general %>% select(Registered.VEP, ERIC, SDR, OVR,
                                 AVR.DMV, AVR.Other, AbsoluteMargin,
                                 ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso5, "General Election Years Only, Including Competitiveness")
## Same results as without competitiveness
# -------------------------------------------------------------------------
# LASSO | General Election Years Only | With Lagged Competitiveness -------
df_lasso6 <- df_general %>% select(Registered.VEP, ERIC, SDR, OVR,
                                 AVR.DMV, AVR.Other, AbsoluteMargin,
                                 lag.AbsoluteMargin,
                                 ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso6, "General Election Years Only, Including Lagged Competitiveness") ## Same res
# -------------------------------------------------------------------------
# LASSO | Midterm Election Years Only -------------------------------------
df_lasso_midterm <- df_midterm %>% select(Registered.VEP, ERIC, SDR, OVR,
                                          AVR.DMV, AVR.Other,
                                          ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso_midterm, "Midterm Election Years Only")
# -------------------------------------------------------------------------
# LASSO | General Election Years Only | With Lagged Competitiveness -------
df_lasso7 <- df_general %>% select(Registered.VEP, ERIC, SDR, OVR,
                                 AVR.DMV, AVR.Other, lag.AbsoluteMargin,
                                 ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso7, "General Election Years Only with Lagged Competitiveness")
```

```r
# LASSO | Preliminary -----------------------------------------------------------
df_lasso <- df %>% select(Registered.VEP, ERIC, SDR, OVR,
                          AVR.DMV, AVR.Other,
                          ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)

run_lasso(df_lasso, "2000-2020") ## ERIC, AVR.Other, and AVR.DMV explain variation in registratic
# -------------------------------------------------------------------------------
# LASSO | 2012-2018 (ERIC Years) Only ---------------------------------------------
df_lasso2 <- df_ERIC_Years %>% select(Registered.VEP, ERIC, SDR, OVR,
                          AVR.DMV, AVR.Other,
                          ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso2, "2012-2018 (ERIC Years) Only") ## No variables are significant
# -------------------------------------------------------------------------------
# LASSO | 2012 - 2018 (ERIC Years) Only | Including Removals --------------
df_lasso3 <- df_ERIC_Years %>% select(Registered.VEP, ERIC, SDR, OVR,
                          AVR.DMV, AVR.Other,
                          ID.Required, Rep.Gov, Rep.Leg, Rep.SoS, Removal.VEP)
run_lasso(df_lasso3, "2012-2018 (ERIC Years) Only | Including Removals") ## No variables are sigr
# -------------------------------------------------------------------------------
# LASSO | General Election Years Only ----------------------------------------
df_lasso4 <- df_general %>% select(Registered.VEP, ERIC, SDR, OVR,
                                   AVR.DMV, AVR.Other,
                                   ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso4, "General Election Years Only")
## AVR.DMV, OVR, ERIC, and AVR.Other are all significant
# -------------------------------------------------------------------------------
# LASSO | General Election Years Only | With Competitiveness -------------
df_lasso5 <- df_general %>% select(Registered.VEP, ERIC, SDR, OVR,
                                AVR.DMV, AVR.Other, AbsoluteMargin,
                                ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso5, "General Election Years Only, Including Competitiveness")
## Same results as without competitiveness
# -------------------------------------------------------------------------------
# LASSO | General Election Years Only | With Lagged Competitiveness -------
df_lasso6 <- df_general %>% select(Registered.VEP, ERIC, SDR, OVR,
                                AVR.DMV, AVR.Other, AbsoluteMargin,
                                lag.AbsoluteMargin,
                                ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso6, "General Election Years Only, Including Lagged Competitiveness") ## Same res
# -------------------------------------------------------------------------------
# LASSO | Midterm Election Years Only ----------------------------------------
df_lasso_midterm <- df_midterm %>% select(Registered.VEP, ERIC, SDR, OVR,
                                      AVR.DMV, AVR.Other,
                                      ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso_midterm, "Midterm Election Years Only")
# -------------------------------------------------------------------------------
# LASSO | General Election Years Only | With Lagged Competitiveness -------
df_lasso7 <- df_general %>% select(Registered.VEP, ERIC, SDR, OVR,
                                AVR.DMV, AVR.Other, lag.AbsoluteMargin,
                                ID.Required, Rep.Gov, Rep.Leg, Rep.SoS)
run_lasso(df_lasso7, "General Election Years Only with Lagged Competitiveness")
```

```
###########################################################################
# Differences-in-Differences to verify ERIC causality --------------------
###########################################################################
df <- df %>% mutate(joined_ERIC = as.factor(ifelse(
      STATE %in% c("ALABAMA", "ALASKA", "ARIZONA", "COLORADO", "CONNECTICUT",
                   "DELAWARE", "FLORIDA", "GEORGIA", "ILLINOIS", "IOWA",
                   "KENTUCKY", "LOUISIANA", "MARYLAND", "MICHIGAN",
                   "MINNESOTA", "MISSOURI", "NEVADA", "NEW MEXICO", "OHIO",
                   "OREGON", "PENNSYLVANIA", "RHODE ISLAND", "SOUTH CAROLINA",
                   "TEXAS", "UTAH", "VERMONT", "VIRGINIA", "WASHINGTON",
                   "DISTRICT OF COLUMBIA", "WEST VIRGINIA", "WISCONSIN"),
      1, 0))
) ## Variable for states that end up joining ERIC
resultLM <- lm(Registered.VEP  ## All years, as shown in main graph
               ~ joined_ERIC + joined_ERIC*ERIC + factor(STATE) + factor(YEAR),data=df)
summary(resultLM) # ERIC *** (p=0.000)
resultLM <- lm(Registered.VEP  ### add controls
               ~ joined_ERIC + joined_ERIC*ERIC + OVR + AVR.DMV + AVR.Other +
                 factor(STATE) + factor(YEAR),data=df)
summary(resultLM) # ERIC ** p=(0.0095)
## Main Graph: States that Joined ERIC vs. Not
ggplot(data = df, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates | Joined ERIC vs. Did Not Join ERIC") +
  theme(text=element_text(family="Rubik")) +
  geom_vline(xintercept = 2012, linetype="dotted", color="gray", size=1) +
  geom_jitter(aes(color=joined_ERIC), alpha=0.3) +
  geom_smooth(aes(color=joined_ERIC), alpha=0.5) +
  geom_label(
    label = "ERIC was first introduced to 7 states in 2012",
    x = 2008,
    y = 0.98,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    size = 3.25,
    color = "#4a4a4a",
    fill = "#e6e6e6"
  ) +
  scale_x_continuous(breaks=c(2000,2002,2004,2006,2008,2010,2012,
                              2014,2016,2018,2020)) +
  scale_y_continuous(labels = scales::percent) +
  xlab("") +
  ylab("Registration Rate") +
  scale_color_hue(labels = c("Never Joined ERIC", "Joined ERIC")) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0),
    legend.position = "top",
    legend.justification = "right",
    legend.title = element_blank()
  )
```

```r
##############################################################################
# Differences-in-Differences to verify ERIC causality --------------------
##############################################################################
df <- df %>% mutate(joined_ERIC = as.factor(ifelse(
      STATE %in% c("ALABAMA", "ALASKA", "ARIZONA", "COLORADO", "CONNECTICUT",
                   "DELAWARE", "FLORIDA", "GEORGIA", "ILLINOIS", "IOWA",
                   "KENTUCKY", "LOUISIANA", "MARYLAND", "MICHIGAN",
                   "MINNESOTA", "MISSOURI", "NEVADA", "NEW MEXICO", "OHIO",
                   "OREGON", "PENNSYLVANIA", "RHODE ISLAND", "SOUTH CAROLINA",
                   "TEXAS", "UTAH", "VERMONT", "VIRGINIA", "WASHINGTON",
                   "DISTRICT OF COLUMBIA", "WEST VIRGINIA", "WISCONSIN"),
      1, 0))
) ## Variable for states that end up joining ERIC
resultLM <- lm(Registered.VEP  ## All years, as shown in main graph
               ~ joined_ERIC + joined_ERIC*ERIC + factor(STATE) + factor(YEAR),data=df)
summary(resultLM) # ERIC *** (p=0.000)
resultLM <- lm(Registered.VEP  ### add controls
               ~ joined_ERIC + joined_ERIC*ERIC + OVR + AVR.DMV + AVR.Other +
                 factor(STATE) + factor(YEAR),data=df)
summary(resultLM) # ERIC ** p=(0.0095)
## Main Graph: States that Joined ERIC vs. Not
ggplot(data = df, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates | Joined ERIC vs. Did Not Join ERIC") +
  theme(text=element_text(family="Rubik")) +
  geom_vline(xintercept = 2012, linetype="dotted", color="gray", size=1) +
  geom_jitter(aes(color=joined_ERIC), alpha=0.3) +
  geom_smooth(aes(color=joined_ERIC), alpha=0.5) +
  geom_label(
    label = "ERIC was first introduced to 7 states in 2012",
    x = 2008,
    y = 0.98,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    size = 3.25,
    color = "#4a4a4a",
    fill = "#e6e6e6"
  ) +
  scale_x_continuous(breaks=c(2000,2002,2004,2006,2008,2010,2012,
                              2014,2016,2018,2020)) +
  scale_y_continuous(labels = scales::percent) +
  xlab("") +
  ylab("Registration Rate") +
  scale_color_hue(labels = c("Never Joined ERIC", "Joined ERIC")) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0),
    legend.position = "top",
    legend.justification = "right",
    legend.title = element_blank()
  )
```

```
##############################################################################
# COHORT ANALYSIS ----------------------------------------------------------
##############################################################################
###  2012  ##################################################################
##############################################################################
# States that joined ERIC in 2012 (Excluding Other ERIC states from later years)
df_ERIC_12 <- df %>% # filter out states that joined ERIC but not in 2012
  filter(!STATE %in% c("ALABAMA", "ALASKA", "ARIZONA", "CONNECTICUT", "FLORIDA",
                       "GEORGIA", "ILLINOIS", "IOWA", "KENTUCKY", "LOUISIANA",
                       "MICHIGAN", "MINNESOTA", "MISSOURI", "NEW MEXICO", "OHIO",
                       "OREGON", "PENNSYLVANIA", "RHODE ISLAND", "SOUTH CAROLINA",
                       "TEXAS", "VERMONT", "DISTRICT OF COLUMBIA",
                       "WEST VIRGINIA", "WISCONSIN")) %>%
  mutate(joined_ERIC = as.factor(ifelse( # create joined_ERIC dummy
  STATE %in% c("COLORADO", "DELAWARE", "MARYLAND", "NEVADA", "UTAH",
               "VIRGINIA", "WASHINGTON"),
  1, 0))
)
resultLM <- lm(Registered.VEP ## Diff-in-Diff Regression
               ~ joined_ERIC + joined_ERIC*ERIC + factor(STATE) + factor(YEAR),data=df_ERIC_12)
summary(resultLM) # ERIC *** (p=0.000154)
resultLM <- lm(Registered.VEP ### add controls
               ~ joined_ERIC + joined_ERIC*ERIC + OVR + AVR.DMV + AVR.Other +
                 factor(STATE) + factor(YEAR),data=df_ERIC_12)
summary(resultLM) # ERIC . (p=0.066116); OVR **
## Plot
ggplot(data = df_ERIC_12, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates | Joined ERIC in 2012 vs. Did Not Join ERIC") +
  labs(subtitle="States that joined ERIC, but Not in 2012, are Excluded",
       caption='Joined "in" YEAR means that the state joined immediately prior to that election year.') +
  theme(text=element_text(family="Rubik")) +
  geom_vline(xintercept = 2012, linetype="dotted", color="gray", size=1) +
  geom_jitter(aes(color=joined_ERIC), alpha=0.3) +
  geom_smooth(aes(color=joined_ERIC), alpha=0.5) +
  geom_label(
    label = "ERIC was first introduced to 7 states in 2012", x = 2008,y = 0.98,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    size = 3.25,
    color = "#4a4a4a",
    fill = "#e6e6e6"
  ) + scale_x_continuous(breaks=c(2000,2002,2004,2006,2008,2010,2012,2014,2016,2018,2020)) +
  scale_y_continuous(labels = scales::percent) + xlab("") + ylab("Registration Rate") +
  scale_color_hue(labels = c("Never Joined ERIC", "Joined ERIC")) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0),
    legend.position = "top", legend.justification = "right", legend.title = element_blank()
  )
```

```
##########################################################################
###  2014  ######## (midterm) ############################################
##########################################################################
# States that joined ERIC prior to 2014 (Excluding Other ERIC states from later years)
df_ERIC_14 <- df %>% # filter out states that joined ERIC but not in 2014
  filter(!STATE %in% c("COLORADO", "DELAWARE", "MARYLAND", "NEVADA", "UTAH",
                       "VIRGINIA", "WASHINGTON","ALABAMA", "ALASKA", "ARIZONA",
                       "FLORIDA", "GEORGIA", "ILLINOIS", "IOWA", "KENTUCKY",
                       "MICHIGAN", "MISSOURI", "NEW MEXICO", "OHIO",
                       "PENNSYLVANIA", "RHODE ISLAND", "SOUTH CAROLINA",
                       "TEXAS", "VERMONT",  "WEST VIRGINIA", "WISCONSIN")) %>%
  mutate(joined_ERIC = as.factor(ifelse(  # create joined_ERIC (in 2014) dummy
    STATE %in% c("CONNECTICUT","DISTRICT OF COLUMBIA","LOUISIANA",
                 "MINNESOTA", "OREGON"),
    1, 0))
  )
resultLM <- lm(Registered.VEP ## Diff-in-Diff Regression
               ~ joined_ERIC + joined_ERIC*ERIC + factor(STATE) + factor(YEAR),data=df_ERIC_14)
summary(resultLM) # Not Significant
resultLM <- lm(Registered.VEP ### add controls
               ~ joined_ERIC + joined_ERIC*ERIC + OVR + AVR.DMV + AVR.Other +
                 factor(STATE) + factor(YEAR),data=df_ERIC_14)
summary(resultLM) # Not Significant
## Plot
ggplot(data = df_ERIC_14, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates | Joined ERIC in 2014 vs. Did Not Join ERIC") +
  labs(subtitle="States that joined ERIC, but Not in 2014, are Excluded",
       caption='Joined "in" YEAR means that the state joined immediately prior to that election year.') +
  theme(text=element_text(family="Rubik")) +
  geom_vline(xintercept = 2014, linetype="dotted", color="gray", size=1) +
  geom_jitter(aes(color=joined_ERIC), alpha=0.3) +
  geom_smooth(aes(color=joined_ERIC), alpha=0.5) +
  geom_label(
    label = "5 states joined ERIC in 2014",
    x = 2011,y = 0.98,label.padding = unit(0.55, "lines"), # Rectangle size around label
    size = 3.25,color = "#4a4a4a",fill = "#e6e6e6"
  ) +
  scale_x_continuous(breaks=c(2000,2002,2004,2006,2008,2010,2012,2014,2016,2018,2020)) +
  scale_y_continuous(labels = scales::percent) +
  xlab("") +
  ylab("Registration Rate") +
  scale_color_hue(labels = c("Never Joined ERIC", "Joined ERIC")) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0),
    legend.position = "top",legend.justification = "right",legend.title = element_blank()
  )
```

```r
######################################################################
###  2016  ###########################################################
######################################################################
# States that joined ERIC prior to 2016 (Excluding Other ERIC states from later years)
df_ERIC_16 <- df %>% # filter out states that joined ERIC but not in 2016
  filter(!STATE %in% c("ARIZONA", "FLORIDA", "GEORGIA", "IOWA", "KENTUCKY",
                       "MICHIGAN", "MISSOURI", "SOUTH CAROLINA", "TEXAS", "VERMONT",
                       "COLORADO", "DELAWARE", "MARYLAND", "NEVADA", "UTAH",
                       "VIRGINIA", "WASHINGTON",
                       "CONNECTICUT","DISTRICT OF COLUMBIA","LOUISIANA",
                       "MINNESOTA", "OREGON")) %>%
  # create joined_ERIC (in 2016) dummy
  mutate(joined_ERIC = as.factor(ifelse(
    STATE %in% c("ALABAMA", "ALASKA", "ILLINOIS", "NEW MEXICO", "OHIO",
                 "PENNSYLVANIA", "RHODE ISLAND", "WEST VIRGINIA", "WISCONSIN"),
    1, 0))
  )
resultLM <- lm(Registered.VEP ## Diff-in-Diff Regression
               ~ joined_ERIC + joined_ERIC*ERIC + factor(STATE) + factor(YEAR),data=df_ERIC_16)
summary(resultLM) # ERIC ** (p=0.001723)
resultLM <- lm(Registered.VEP ### add controls
               ~ joined_ERIC + joined_ERIC*ERIC + OVR + AVR.DMV + AVR.Other +
                 factor(STATE) + factor(YEAR),data=df_ERIC_16)
summary(resultLM) # Not Significant; OVR, AVR.Other
## Plot
ggplot(data = df_ERIC_16, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates | Joined ERIC in 2016 vs. Did Not Join ERIC") +
  labs(subtitle="States that joined ERIC, but Not in 2016, are Excluded",
       caption='Joined "in" YEAR means that the state joined immediately prior to that election year.') +
  theme(text=element_text(family="Rubik")) +
  geom_vline(xintercept = 2016, linetype="dotted", color="gray", size=1) +
  geom_jitter(aes(color=joined_ERIC), alpha=0.3) +
  geom_smooth(aes(color=joined_ERIC), alpha=0.5) +
  geom_label(
    label = "9 states joined ERIC in 2016", x = 2013,y = 0.98,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    size = 3.25,color = "#4a4a4a",fill = "#e6e6e6"
  ) +
  scale_x_continuous(breaks=c(2000,2002,2004,2006,2008,2010,2012,
                              2014,2016,2018,2020)) +
  scale_y_continuous(labels = scales::percent) +
  xlab("") +ylab("Registration Rate") + scale_color_hue(labels = c("Never Joined ERIC", "Joined ERIC")) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0),
    legend.position = "top",
    legend.justification = "right",
    legend.title = element_blank()
  )
```

```r
###############################################################################
###  2018  ######## (midterm) ##############################################
###############################################################################
# States that joined ERIC prior to 2018 (Excluding Other ERIC states from later years)
df_ERIC_18 <- df %>% # filter out states that joined ERIC but not in 2016
  filter(!STATE %in% c("FLORIDA", "GEORGIA", "KENTUCKY",
                       "MICHIGAN", "SOUTH CAROLINA", "TEXAS", "VERMONT",
                       "COLORADO", "DELAWARE", "MARYLAND", "NEVADA", "UTAH",
                       "VIRGINIA", "WASHINGTON",
                       "CONNECTICUT","DISTRICT OF COLUMBIA","LOUISIANA",
                       "MINNESOTA", "OREGON",
                       "ALABAMA", "ALASKA", "ILLINOIS", "NEW MEXICO", "OHIO",
                       "PENNSYLVANIA", "RHODE ISLAND", "WEST VIRGINIA", "WISCONSIN")
                       ) %>%
  mutate(joined_ERIC = as.factor(ifelse( # create joined_ERIC (in 2018) dummy
    STATE %in% c("ARIZONA", "IOWA", "MISSOURI"),
    1, 0))
  )
resultLM <- lm(Registered.VEP ## Diff-in-Diff Regression
               ~ joined_ERIC + joined_ERIC*ERIC + factor(STATE) + factor(YEAR),data=df_ERIC_18)
summary(resultLM) # Not Significant (midterm)
resultLM <- lm(Registered.VEP ### add controls
               ~ joined_ERIC + joined_ERIC*ERIC + OVR + AVR.DMV + AVR.Other +
                 factor(STATE) + factor(YEAR),data=df_ERIC_18)
summary(resultLM) # Not Significant; OVR, AVR.Other
## Plot
ggplot(data = df_ERIC_18, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates | Joined ERIC in 2018 vs. Did Not Join ERIC") +
  labs(subtitle="States that joined ERIC, but Not in 2018, are Excluded",
       caption='Joined "in" YEAR means that the state joined immediately prior to that election year.') +
  theme(text=element_text(family="Rubik")) +
  geom_vline(xintercept = 2018, linetype="dotted", color="gray", size=1) +
  geom_jitter(aes(color=joined_ERIC), alpha=0.3) +
  geom_smooth(aes(color=joined_ERIC), alpha=0.5) +
  geom_label(
    label = "3 states joined ERIC in 2018", x = 2015,y = 0.98,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    size = 3.25,color = "#4a4a4a",fill = "#e6e6e6"
  ) +
  scale_x_continuous(breaks=c(2000,2002,2004,2006,2008,2010,2012,2014,2016,2018,2020)) +
  scale_y_continuous(labels = scales::percent) +
  xlab("") + ylab("Registration Rate") +
  scale_color_hue(labels = c("Never Joined ERIC", "Joined ERIC")) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0),
    legend.position = "top",legend.justification = "right",legend.title = element_blank()
  )
```

```r
##########################################################################
###  2020  ################################################################
##########################################################################
# States that joined ERIC prior to 2020 (Excluding Other ERIC states from later years)
df_ERIC_20 <- df %>% # filter out states that joined ERIC but not in 2016
  filter(!STATE %in% c("COLORADO", "DELAWARE", "MARYLAND", "NEVADA", "UTAH",
                       "VIRGINIA", "WASHINGTON",
                       "CONNECTICUT","DISTRICT OF COLUMBIA","LOUISIANA",
                       "MINNESOTA", "OREGON",
                       "ALABAMA", "ALASKA", "ILLINOIS", "NEW MEXICO", "OHIO",
                       "PENNSYLVANIA", "RHODE ISLAND", "WEST VIRGINIA", "WISCONSIN",
                       "ARIZONA", "IOWA", "MISSOURI")
  ) %>%
  mutate(joined_ERIC = as.factor(ifelse( # create joined_ERIC (in 2020) dummy
    STATE %in% c("FLORIDA", "GEORGIA", "KENTUCKY",
                 "MICHIGAN", "SOUTH CAROLINA", "TEXAS", "VERMONT"),
    1, 0))
  )
resultLM <- lm(Registered.VEP ## Diff-in-Diff Regression
               ~ joined_ERIC + joined_ERIC*ERIC + factor(STATE) + factor(YEAR),data=df_ERIC_20)
summary(resultLM) # ERIC *** (p=0.00000)
resultLM <- lm(Registered.VEP ### add controls
               ~ joined_ERIC + joined_ERIC*ERIC + OVR + AVR.DMV + AVR.Other +
                 factor(STATE) + factor(YEAR),data=df_ERIC_20)
summary(resultLM) # ERIC *** (p=0.00003); OVR
## Plot
ggplot(data = df_ERIC_20, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates | Joined ERIC in 2020 vs. Did Not Join ERIC") +
  labs(subtitle="States that joined ERIC, but Not in 2020, are Excluded",
       caption='Joined "in" YEAR means that the state joined immediately prior to that election year.') +
  theme(text=element_text(family="Rubik")) +
  geom_vline(xintercept = 2020, linetype="dotted", color="gray", size=1) +
  geom_jitter(aes(color=joined_ERIC), alpha=0.3) +
  geom_smooth(aes(color=joined_ERIC), alpha=0.5) +
  geom_label(
    label = "7 states joined ERIC in 2020", x = 2017,y = 0.98,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    size = 3.25,color = "#4a4a4a",fill = "#e6e6e6"
  ) +
  scale_x_continuous(breaks=c(2000,2002,2004,2006,2008,2010,2012,2014,2016,2018,2020)) +
  scale_y_continuous(labels = scales::percent) + xlab("") + ylab("Registration Rate") +
  scale_color_hue(labels = c("Never Joined ERIC", "Joined ERIC")) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0),
    legend.position = "top",legend.justification = "right",legend.title = element_blank()
  )
```

```
# COHORT ANALYSIS | GENERAL ELECTIONS ONLY -------------------------------
##########################################################################
###  2012  ###############################################################
##########################################################################
df_ERIC_12_general <- df_general %>% # filter out states that joined ERIC but not in 2012
  filter(!STATE %in% c("ALABAMA", "ALASKA", "ARIZONA", "CONNECTICUT",
                       "FLORIDA", "GEORGIA", "ILLINOIS", "IOWA",
                       "KENTUCKY", "LOUISIANA", "MICHIGAN", "MINNESOTA",
                       "MISSOURI", "NEW MEXICO", "OHIO", "OREGON",
                       "PENNSYLVANIA", "RHODE ISLAND", "SOUTH CAROLINA",
                       "TEXAS", "VERMONT", "DISTRICT OF COLUMBIA",
                       "WEST VIRGINIA", "WISCONSIN")) %>%
  mutate(joined_ERIC = as.factor(ifelse(  # create joined_ERIC dummy
    STATE %in% c("COLORADO", "DELAWARE", "MARYLAND", "NEVADA", "UTAH",
                 "VIRGINIA", "WASHINGTON"),
    1, 0))
  )
resultLM <- lm(Registered.VEP ## Diff-in-Diff Regression
               ~ joined_ERIC + joined_ERIC*ERIC + factor(STATE) + factor(YEAR),data=df_ERIC_12_general)
summary(resultLM) # ERIC * (p=0.017883)
resultLM <- lm(Registered.VEP ### add controls
               ~ joined_ERIC + joined_ERIC*ERIC + OVR + AVR.DMV + AVR.Other +
                 factor(STATE) + factor(YEAR),data=df_ERIC_12_general)
summary(resultLM) # Not Significant; OVR *
## Plot
ggplot(data = df_ERIC_12_general, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates | Joined ERIC in 2012 vs. Did Not Join ERIC | General Elections Only") +
  labs(subtitle="States that joined ERIC, but Not in 2012, are Excluded",
       caption='Joined "in" YEAR means that the state joined immediately prior to that election year.') +
  theme(text=element_text(family="Rubik")) +
  geom_vline(xintercept = 2012, linetype="dotted", color="gray", size=1) +
  geom_jitter(aes(color=joined_ERIC), alpha=0.3) +
  geom_smooth(aes(color=joined_ERIC), alpha=0.5) +
  geom_label(
    label = "ERIC was first introduced to 7 states in 2012", x = 2008,y = 0.98,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    size = 3.25,color = "#4a4a4a",fill = "#e6e6e6"
  ) +
  scale_x_continuous(breaks=c(2000,2004,2008,2012,2016,2020)) +
  scale_y_continuous(labels = scales::percent) +xlab("") +ylab("Registration Rate") +
  scale_color_hue(labels = c("Never Joined ERIC", "Joined ERIC")) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0),
    legend.position = "top",
    legend.justification = "right",
    legend.title = element_blank()
  )
```

```r
###########################################################################
###  2016  ################################################################
###########################################################################
# states that joined prior to 2014 and prior to 2016
df_ERIC_16_general <- df_general %>%   # filter out states that joined ERIC but not in 2016
  filter(!STATE %in% c("ARIZONA", "FLORIDA", "GEORGIA", "IOWA", "KENTUCKY",
                       "MICHIGAN", "MISSOURI", "SOUTH CAROLINA", "TEXAS", "VERMONT",
                       "COLORADO", "DELAWARE", "MARYLAND", "NEVADA", "UTAH",
                       "VIRGINIA", "WASHINGTON")) %>%
  mutate(joined_ERIC = as.factor(ifelse( # create joined_ERIC (in 2016) dummy
    STATE %in% c("CONNECTICUT","DISTRICT OF COLUMBIA","LOUISIANA",
                 "MINNESOTA", "OREGON", # 2014
                 "ALABAMA", "ALASKA", "ILLINOIS", "NEW MEXICO", "OHIO",
                 "PENNSYLVANIA", "RHODE ISLAND", "WEST VIRGINIA", "WISCONSIN"), # 2016
    1, 0))
  )
resultLM <- lm(Registered.VEP ## Diff-in-Diff Regression
               ~ joined_ERIC + joined_ERIC*ERIC + factor(STATE) + factor(YEAR),data=df_ERIC_16_general)
summary(resultLM) # Not Significant p = 0.238471
resultLM <- lm(Registered.VEP ### add controls
               ~ joined_ERIC + joined_ERIC*ERIC + OVR + AVR.DMV + AVR.Other +
                 factor(STATE) + factor(YEAR),data=df_ERIC_16_general)
summary(resultLM) # Not Significant; OVR, AVR.Other
## Plot
ggplot(data = df_ERIC_16_general, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates | Joined ERIC 2012-2016 vs. Did Not Join ERIC | General Elections Only") +
  labs(subtitle="States that joined ERIC, but Not in 2016, are Excluded") +
  theme(text=element_text(family="Rubik")) +
  geom_vline(xintercept = 2016, linetype="dotted", color="gray", size=1) +
  geom_jitter(aes(color=joined_ERIC), alpha=0.3) +
  geom_smooth(aes(color=joined_ERIC), alpha=0.5) +
  geom_label(
    label = "14 states joined ERIC After 2012 and Prior to 2016",x = 2012,y = 0.98,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    size = 3.25,color = "#4a4a4a",fill = "#e6e6e6"
  ) +
  scale_x_continuous(breaks=c(2000,2004,2008,2012,2016,2020)) +
  scale_y_continuous(labels = scales::percent) +
  xlab("") +
  ylab("Registration Rate") +
  scale_color_hue(labels = c("Never Joined ERIC", "Joined ERIC")) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0),
    legend.position = "top",
    legend.justification = "right",
    legend.title = element_blank()
  )
```

```r
###############################################################################
###  2020  ###################################################################
###############################################################################
df_ERIC_20_general <- df_general %>% # states that joined prior to 2018 and prior to 2020
  filter(!STATE %in% c("CONNECTICUT","DISTRICT OF COLUMBIA","LOUISIANA",
                       "MINNESOTA", "OREGON", # 2014
                       "ALABAMA", "ALASKA", "ILLINOIS", "NEW MEXICO", "OHIO",
                       "PENNSYLVANIA", "RHODE ISLAND", "WEST VIRGINIA", "WISCONSIN")) %>%
  mutate(joined_ERIC = as.factor(ifelse( # create joined_ERIC (in 2020) dummy
    STATE %in% c("ARIZONA", "FLORIDA", "GEORGIA", "IOWA", "KENTUCKY",
                 "MICHIGAN", "MISSOURI", "SOUTH CAROLINA", "TEXAS", "VERMONT",
                 "COLORADO", "DELAWARE", "MARYLAND", "NEVADA", "UTAH",
                 "VIRGINIA", "WASHINGTON"),
    1, 0))
  )
resultLM <- lm(Registered.VEP ## Diff-in-Diff Regression
               ~ joined_ERIC + joined_ERIC*ERIC + factor(STATE) + factor(YEAR),data=df_ERIC_20_general)
summary(resultLM) # ERIC *** (p=0.000325)
resultLM <- lm(Registered.VEP ### add controls
               ~ joined_ERIC + joined_ERIC*ERIC + OVR + AVR.DMV + AVR.Other +
                 factor(STATE) + factor(YEAR),data=df_ERIC_20_general)
summary(resultLM) # ERIC * (p=0.011493); OVR
## Plot
ggplot(data = df_ERIC_20_general, mapping = aes(x = YEAR, y = Registered.VEP)) +
  ggtitle("Registration Rates | Joined ERIC After 2016, Before 2020 vs. Did Not Join ERIC") +
  labs(subtitle="General Election Years Only") +
  theme(text=element_text(family="Rubik")) +
  geom_vline(xintercept = 2020, linetype="dotted", color="gray", size=1) +
  geom_jitter(aes(color=joined_ERIC), alpha=0.3) +
  geom_smooth(aes(color=joined_ERIC), alpha=0.5) +
  geom_label(
    label = "17 states joined ERIC After 2016 and Prior to 2020", x = 2016,y = 0.98,
    label.padding = unit(0.55, "lines"), # Rectangle size around label
    size = 3.25,color = "#4a4a4a",fill = "#e6e6e6"
  ) +
  scale_x_continuous(breaks=c(2000,2004,2008,2012,2016,2020)) +
  scale_y_continuous(labels = scales::percent) +
  xlab("") +
  ylab("Registration Rate") +
  scale_color_hue(labels = c("Never Joined ERIC", "Joined ERIC")) +
  theme(
    panel.background = element_rect(fill = "#f2f2f2"),
    panel.grid.major.x = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.x = element_line(size = 0),
    panel.grid.major.y = element_line(color = "#e6e6e6", size = 0.5),
    panel.grid.minor.y = element_line(size = 0),
    legend.position = "top",
    legend.justification = "right",
    legend.title = element_blank()
  )
```