The Keys to the White House

Forecasting the U.S presidential elections

What influences the voting share?

Surname	Name	Email address on MS Teams/ARIEL
Gervasoni	Sofia	sofia.gervasoni@studenti.unimi.it
Paoletti	Luca	luca.paoletti@studenti.unimi.it
Saini	Chiara	chiara.saini@studenti.unimi.it
Scaffidi Caruso	Giuseppe	giuseppe.scaffidicaruso@studenti.unimi.it

Abstract: The following paper aims to understand which variables influence the voting share at a State level and predict the outcome of the 2012 and 2020 USA elections. The study of Americans' voting behaviour was carried out by applying a linear model and using panel data. The examined period is between 1984 and 2020, and the analysis is conducted for each of the 51 states of the USA. Firstly, Fair's theoretical model was used as a guideline for constructing the variables of economic condition and incumbency (measures of economic performance such as GDP growth and inflation, the familiarity of the eligible candidate, and preference for the party in power). Subsequently, more variables were included to improve the explanatory potential of the model. For this intent, variables that describe the unemployment rate and the presence of capital punishment in the different states were introduced. Finally, the technique of the Double Selection Lasso (DSL) was used and the model obtained produced better forecasts. We were able to point out that the variables that better explained the democratic voting share are the growth in real GDP pc, the incumbency variables and the Unemployment Rate. They all presented a positive correlation with the dependent variable (democratic share).

<u>Index</u>

Part 1 - Microeconometrics

Data

For this analysis, we ran three different regressions with the same dependent variable, Democratic share (we called it "dem_share"). The Democratic vote share of the two parties is defined as *100*D/(D + R);* where D and R are respectively the number of votes for Democratic and Republican candidates. To obtain the dependent variable, we used the data provided by Harvard Dataverse ("U.S. President 1976– 2020"). The difference between the models (regressions) is given by the independent variables used. We distinct our three models as:

1. *Model 1*: we rearranged Fair's model for panel data, using annual values;
2. *Model 2*: we used the rearranged Fair's model and added further variables, useful to explain the voting behavior at a State Level (e.g., unemployment rate, death penalty);
3. *Model 3*: the final model was obtained after running the LASSO on *Model 2*.

For the ***Rearranged Fair's model*** at point ***1***, we used the following variables:

- ***I***: assumes values 1 (a Democratic presidential incumbent at the time of the election) and -1 (Republican presidential incumbent). This is our control variable.
- ***DPER***: assumes values 1 (Democratic presidential incumbent is running again), -1 (Republican presidential incumbent is running again), and 0 (otherwise). This is the incumbency variable.
- ***DUR***: assumes values 0 when R or D party has been in the White House for one term; 1 if the Democratic party has been in the White House for two consecutive terms; -1 if the Republican party has been in the White House for two consecutive terms, -1.25 if the Republican party has been in the White House for three consecutive terms.
- ***G***: growth rate of real per capita GDP in the first four quarters of the on-term election year (annual rate).
- ***P***: absolute value of the growth rate of the GDP deflator in the first 16 quarters of the administration (annual rate).
- ***Z***: number of years in the four years of administration in which the growth rate of real per capita GDP is greater than the average yearly growth for each State.

To calculate the economic variables, we used annual Real/Chained GDP (GDPR), annual Nominal/Current GDP (GDP), and annual Population by State (POP).

To calculate these variables we used the following formulas:

$$G = [(Y_{16}/Y_{12})^\wedge(4/3) - 1] \cdot 100$$

$$P = [(GDPD_{16}/GDPD_{16}(-1))^\wedge(4/15) - 1] \cdot 100$$

Where (−1) indicates the previous four-year election period.

To construct Z we need to define the growth rate in a given year, which for year k is $gk=[(Yk/Yk-1) -1]\cdot 100$. Therefore, *Z* is the number of years in the four years of administration in which gk is greater than the average growth for each State.

Knowing that:

- *Y = GDPR/POP*, is the real per capita GDP;
- *GDPD = GDP/GDPR*, is the GDP deflator.

For *Model 2*, we added two more variables to *Model 1:* ***Unemployment Rate*** as a proxy of economic issues and ***Death Penalty*** as a proxy of social and cultural representation.

The ratio behind our variable (***Unemployment Rate***) derived from the change in the insured unemployment rate between the election year and two years prior to the election year.

For panel data, we assumed data on a State level. The reasoning has been done on two types of data: the overall US unemployment rate (UR) and the US Insured Unemployment Rate (IUR).

The first one indicates the percentage changes in US unemployment; the second is a nonstandard measure of unemployment (related to the first), which equals the percentage of the labour force collecting State or federal unemployment benefits. This measure does not include the unemployed and those whose benefits have been delayed or have run out. However, data for the US insured unemployment rate are not available for all States and periods under analysis.

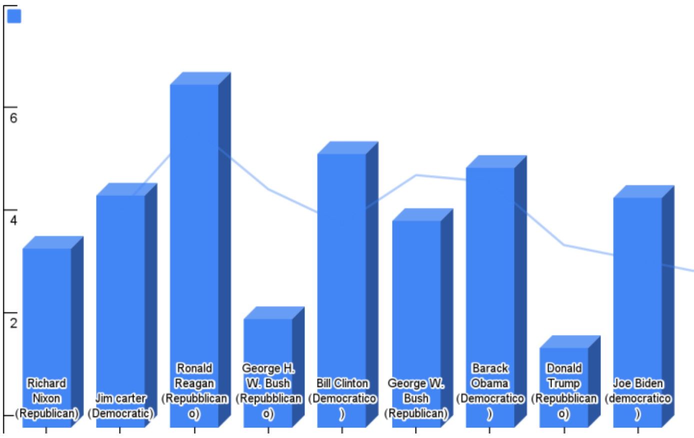
To solve this issue, we split our study into two **Trees:**

1. We modelled the overall data in the first one, producing the assumption to set in our analysis;
2. Starting from the process above, we fitted the assumption on our Panel data set, which was based on the US insured unemployment rate but suited to the US unemployment rate[1].

In the **first Tree,** we computed the correlation between the IUR[2] and the UR[3]; this allowed us to identify the grade of strength between the two variables to understand if the assumption on IUR could be carried out on the UR. The result from the computed correlation was **0.7**, meaning that the measure employed in this computation would provide a fair picture of the actual employment condition in each state. Moreover, we formulated the framework where we applied the logical computation to understand the correlation between these variables and the US presidential election (see *Graph 1)*.

The first passage was:

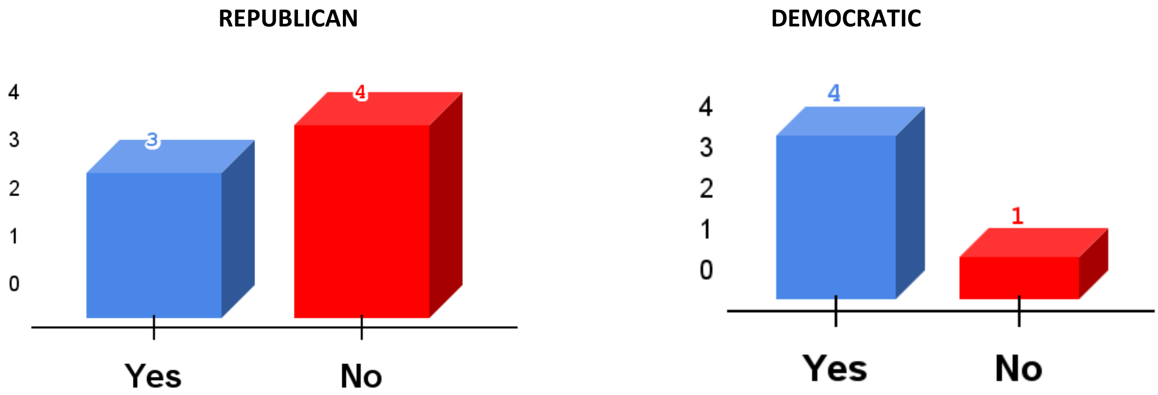
- Computing the average of the **IUR** in order to have a yearly observation (data were provided monthly);
- Computing the absolute (**ABS**)&(%) changing in the two previous years (according to the previous assumption);
- Computing the **cumulate** two years prior to better understand if the insurance decrease was weak or strong.



**Notice the EMA under the data to better understand the trend.*

Graph 1. Variation between different levels of IUR during the mandates.

Graph 2. unemployment benefits standpoint: work of Republicans vs work of Democrats.



---

[1] The main problem is, not all data are available for the timeframe of our interest. For this reason, we face the choice to set the assumption on the main US unemployment rate. URL:[Annual Unemployment Rates by State | Iowa Community Indicators Program]

[2] Insurance Unemployment Rate. URL:[Insured Unemployment Rate (IURNSA) | FRED | St. Louis Fed]

[3] Unemployment Rate.

From the unemployment benefits standpoint, the work of Republicans is worse than Democrats', as shown in *Graph 2*. **Yes** means that the assumption of "decreasing in i*nsurance unemployment*" was valid in the last two years prior to the presidential election for the incumbent party. **No** represents the opposite result.
The **second Tree** represents the adjustment of our assumption on the panel data set.
Knowing that the correlation is 0.7, we can apply the same computation to the UR panel data. We can yield the final table with the decrease or increase in the IUR for each state using the same ratio.
Once again, the conclusion is that the work of the Republicans, referring to unemployment, is worse than the Democrats'. When Republicans were in office, unemployment benefits were significantly reduced during the last two years of their presidency before the new election.
To deepen our analysis, we thought it might be interesting to introduce a variable linked to the **Death Penalty** in the various American states. Indeed, we have noticed that the States where the Death Penalty is currently active are tendentially Republicans. We have introduced a variable, indicating which states abolished capital punishment between 1984 and 2020. Capital Punishment can be active, abolished, or suspended. Due to lack of data, we have categorized our variable as abolished or not abolished. We call this variable "**Death penalty**," it takes value 1 if the Death Penalty has been abolished, otherwise 0. Data from 1984 to 2020 for each State were collected from Death Penalty Information Center.

In Model 3, we ran the Double-Selection LASSO, and we obtained our final model. This model used the variables ***g_i*** (G*I = Real GDP growth), ***dur, dper, p_i, z_i,*** and ***unemployment.*** We also kept the fixed effects for the States. One problem worth highlighting concerning the extrapolation of the data are the missing quarterly data. For G and P variables, the formula would be better with quarterly data. However, this assumption is impossible to achieve due to a lack of data (our dataset was missing quarterly data before 2005). To overcome this issue, we considered annual data adjusting G and P formulas. Instead of computing the **15th** quarter (three *years and nine months*) of the election year, we computed the **16th** quarter. The economic logic behind the best fitting of the quarterly data, relies on the assumption that the incumbent President in the 15th quarter will be committed to carrying out an electoral campaign trying to widen the economy in order to increase his re-election probability. By looking at the Statistical effects of this adjustment, this minor replacement does not significantly change the consistency of our variable because the values are almost identical[4].

Results

Model 1 - Rearranged Fair's Model

We started our analysis by estimating a fixed effects (FE, at a level State) model, containing the same variables of Fair's rearranged model for panel data and annual values (as above mentioned). We obtained the following model:

$$DEM\_SHARE_{it} = \alpha_i+\beta_0+\beta_1(G_{it}\times I_{it})+\beta_2(P_{it}\times I_{it})+\beta_3(Z_{it}\times I_{it})+\beta_4 I_{it}+\beta_5 DPER_{it}+\beta_6 DUR_{it}+u_{it}$$

Where $i = [1,...,51]$ represents the 51 States in the USA and $t = [1984, 1988,..., 2016, 2020]$ represents the election years between 1984 and 2020. Moreover, αi represents the State (entity) fixed effect, and it could be rewritten using a set of dummy variables that select each State (***i***).

$$\alpha_i = \gamma_2 ALASKA_i + . . . + \gamma_n WYOMING_i$$

Where ALASKA is the first dummy variable, ALASKAi=1 when i=2 (and 0 otherwise). Whereas WYOMING is the last dummy variable, WYOMINGi=1 when i=51 (and 0 otherwise). The reference State is Alabama (i=1), so when i=1, αi=0.
We decided to introduce the FE only for the entities because we are interested in capturing the time-invariant characteristics for each State. Furthermore, trying to run the analysis with FE for time, we found out non-significant values and problems of collinearity, so we decided to keep only the FE at a State Level. Running the regression on this model, we found out significant values for the estimated coefficients *β2, β4, β5, and β6. These* values are statistically distinguishable from zero. The estimates of coefficients *β1* and *β3* do not result statistically significant at any confidence value (statistically non-distinguishable from zero). The majority of estimates of the coefficients γi referred to the dummies of the States FE are statistically significant (since their

[4] Look at the appendix for the dataset, to see the slight difference between most of the variables G and P calculated with quarterly (15th quarter) instead of yearly data (16th quarter).

p values are mostly close to 0), even if for some States we obtained non-significant values at a 95% confidence interval.

- Economic Variables

The variable ***G*I*** (*β1* hat) 's estimated coefficient is negative but statistically not significant. We expected a positive value for this coefficient since an increase in real per capita GDP in the first four quarters of on term election year should increase vote share. This result tells us that if GDP pc is growing during the election year, people tend not to appreciate the work of the party in charge, and it does not make sense. According to our estimate, a 1% increase in Real pc GDP causes a 0.0056289% decrease in vote share. The p-value related to this estimate is 0.925, so we do not reject H0:{*β1=0*} (not statistically significant).
We expected the coefficient of ***Z*I*** to be positive since one extra period of growth in the GDP rate should increase the Democratic vote share. Even if we obtained a positive coefficient for this variable (0.2069457%), it is not significant (pval=0.362). Although the estimated coefficient for this variable is not statistically significant, we agree with its positive sign since it makes no sense to have a negative coefficient. One more period of Real pc GDP above the mean (for each State), indeed, causes an increase in vote share. Z and G, for this model, are not statistically significant, so they do not deserve to be further analysed.
The coefficient of ***P*I*** is the only one that results statistically different from zero. From our regression, we get that a 1% increase in real p.c. GDP deflator in absolute value increases vote share by 0.7702654%. Surprisingly, an increase in inflation is linked to an increase in the percentage of voting share. However, this is possible because, in general, we can say that an increase in inflation (within certain limits) is a sign of a healthy and expanding economy. On the other hand, we can say that deflation, is a sign of an unhealthy economy. We ran a test for the joint significance of economic variables in the model, where the H0:{*β1=β2=β3=0*}. We obtained an F-Statistic=20.06 and a pval=0.0, so we reject H0. This result means that jointly the economic variables are significantly different from zero. This is mainly because the coefficient of ***P*I*** is statistically different from zero.

- Incumbency Variables

For the coefficient of the **DPER** variable, we expected a positive value since the incumbent President is a familiar figure, and the electorate tends to re-vote him/her. Indeed, this variable has a positive coefficient (2.149338%) and is statistically significant (pval=0.001). From this result, we can interpret that a change of candidates does not seem preferable. People get familiar with the incumbent President and tend to re-vote him/her when they are satisfied with his/her administration. Between 1980 and 2020, four Presidents have been re-elected (Reagen, Clinton, W.Bush, and Obama). Furthermore, it is common for Presidents to be re-elected for a second term, as American history teaches us.
For the coefficient of **DUR**, we obtained a positive value for the estimate (4.051801%), and it is also significantly different from zero for every confidence interval since the p-value=0.0. As stated above, in the US, it is common that a President (and so his party) stays in charge for more than one term. Usually, a Party stays in charge for a maximum of two consecutive years. However, between 1980 and 1988, Republicans were elected for three consecutive terms (two terms with Reagan and one term with H.W.Bush).

- Joint significance of state-level fixed effects

In this case, we are testing H0: {γ2=*...*=γn=*0*}. From this Poolability test, we obtained an F-statistic=47.25 and a pval=0. The result suggested keeping the FE for the States in our model, since the coefficients are jointly significantly different from zero.

- Forecasts

Using the fitted values from this regression, we can obtain the 2012 and 2020 election results' forecasts and compare them with the actual ones. According to our model, in 2012, Democrats obtained 342 out of 538 Great Electors' votes. Therefore, Democrats gained 63.57% of the Great Electors' votes, winning the election. The actual result of 2012 was 332 Great Electors' votes for the democrats (61.71%), winning the elections with candidate Barack Obama. The model we created worked well for 2012, with a minimal error of 1.86%.
Applying our model on 2020 data, Democrats obtained 342 out of 538 Great Electors' votes, winning the election. The actual result of 2020 was 306 Great Electors' votes (56.88%), winning the elections with President Biden. Even in this situation, our model worked well, predicting the correct outcome. At the same time, the prediction error was higher (6.69%).

We have to bear in mind that the 2020 elections were affected by the Covid-19 pandemic. We did not include the shock of this crisis in our model. However, *"The COVID-19 pandemic and the 2020 US presidential election"*[5] supported this thesis. This analysis found proof that the COVID-19 pandemic negatively affected the voting share (at the county level) for President Donald Trump between 2016 and 2020. From this study, the negative effect of COVID-19 seems to be stronger in urban counties, States without stay-at-home orders, areas considered battlegrounds between parties, and, most interestingly, States where Trump won in 2016. A counterfactual analysis also suggests that Trump would likely have won the re-election if COVID-19 cases had been 5% lower (see the table with the results of the forecasts in the appendix).

Model 2 - Rearranged Fair's Model, Unemployment rate and Death Penalty

We applied a Double-Selection LASSO to *Model 2*, keeping real per capita GDP (variable g*i) in the regression. We obtained the following model:

$$DEM\_SHARE_{it} = \alpha_i+\beta_0+\beta_1(G_{it}\times I_{it})+\beta_2(P_{it}\times I_{it})+\beta_3(Z_{it}\times I_{it}) +\beta_4 I_{it}+\beta_5 DPER_{it}+\beta_6 DUR_{it}+\beta_7 UNEMP_{it} + \beta_8 DP_{it}+u_{it}$$

For the coefficient *β8* (***Death Penalty***) we found a positive value (4.36927%), as expected, since passing from the reference category ("not abolished") to the other ("abolished"), shows a preference in voting Democratics instead of Republicans (the Democratic vote share increased by 4.36927%). The estimate for this coefficient is statistically different from zero (pval= 0.0).
For the coefficient *β7* (***Unemployment***), we found a negative value (-0.3463784%), but it is not statistically significant (since the pval=0.219). Since Democrats' policies are more effective in decreasing the Unemployment Rate[6], when the unemployment is high, the Democratic candidates have more chances to be elected (higher Democratic share). Therefore, we expect a positive value for this coefficient.
The other variables provide more or less the same result[7]. Since many variables are still not significant (both in models 1 and 2), we are not satisfied with the results obtained. We ran the LASSO on this model to keep only the significant variables to explain the Democratic Share.

Model 3 - Final model (Obtained after DS LASSO)

We applied a Double-Selection LASSO to *Model 2*, keeping real per capita GDP (variable g*i) in the regression. We obtained the following model:

$$DEM\_SHARE_{it} = \alpha_i+\beta_0+\beta_1(G_{it}\times I_{it})+\beta_2\, DPER_{it}+\beta_3 DUR_{it}+\beta_4 UNEMP_{it}+\beta_2(P_{it}\times I_{it})+\beta_3(Z_{it}\times I_{it})+u_{it}$$

In this circumstance, the estimate of ***β1*** (***G*I***) does not result statistically significant (pval= 0.642), yet, we obtained the positive coefficient that was expected (as we said before, when we analyzed the results of Model 1). In this model, a 1% increase in Real pc GDP causes a 0.0324557% increase in vote share. We agree with the sign of the coefficient. In this case, the non-significance of this coefficient could be due to its value not being close to zero. We expected a more significant increase in vote share due to a Real pc GDP growth.
Nevertheless, we are more satisfied with this result than with the previous ones.
The estimated coefficients for the variables ***DPERit*** and ***DURit*** remain positive (respectively 2.068809 and 1.781064) and significant (p-value respectively 0.005 and 0.01). We agree with these results. In this model, for the coefficient of ***UNEMPit***, we obtained a positive value (2.017086), and it also resulted statistically significant (p-value = 0.000). This estimation is coherent with our analysis. As aforementioned, we expected an increase in the Unemployment Rate would increase the Democratic share since Democratic candidates can successfully convince voters that their party is better at solving this problem.
The estimate coefficients for ***P*I*** *and* ***Z*I*** take values respectively 0.0649079 and -0.0386569, but they both do not result statistically significant (p-value respectively 0.515 and 0.874). The estimated coefficient of ***P*It*** takes a positive value (lower than in the OLS regression in *Model 1*). Even if it does not result statistically significant, we agree with the sign of this coefficient (for the reasons explained above). Whereas the estimated coefficient of ***Z*I****,* in this case, takes a negative value and we disagree with it (for the logic explained in the analysis of *Model 1*). In this case, this coefficient results are also not statistically significant. Since we are interested in having real GDP pc in our model, we do not want the LASSO to throw it away. Therefore, we make sure to keep it using a ***Double Selection LASSO***. This method is a three steps process:

---

5 The COVID-19 pandemic and the 2020 US presidential election, Leonardo Baccini, Abel Brodeur, Stephen Weymouth, Journal of Population Economics, 8 December 2020

6 "Unemployment and the Democratic Electoral Advantage", JOHN R. WRIGHT, Ohio State University

7 See the Appendix for the results

1. Select the confounders that are relevant to explain the dependent variable (dem_share),
2. Select the confounders that are relevant to explain the variable that we want to keep in our model, which does not enter in the LASSO penalty (G*I);
3. Run a regressing dem_share on G*I and have the confounders selected by the two LASSOs (***X_demshare*** and ***X_gi***).

In our case, the confounders selected to explain dem_share were ***DUR, DPER,*** *and* ***UNEMP*** (*X_demshare*), whereas the confounders selected to explain ***G*I*** were ***DUR, DPER, UNEMP, P*I,*** *and* ***Z*I*** (***X_gi)***. In the ***Naive-Post selection LASSO***, we only omit ***G*I*** from the LASSO penalty, and we run the regression of ***dem_share*** on ***G*I,*** and the confounders selected by the LASSO (***X_demshare***). Here, ***X_gi*** are omitted, so the LASSO is setting to zero variables correlated only with ***G*I*** (that are ***P*I*** *and* ***Z*I***). In this situation, we are omitting relevant variables, leading to an unbiased and inconsistent estimation of ***β1***. That is why a Double Selection LASSO should be preferred.

- Forecasts

In the final model, we obtained the following results for the forecast of the 2012 and the 2020 elections. In 2012 we had the same results as in Model 1. Otherwise, applying the final model on 2020 data, we noticed a positive change in prediction compared to Model 1. Democrats obtained 285 out of 538 Great Electors' votes (therefore 52.97%), winning the election. The actual result of 2020 was 306 Great Electors' votes (56.88%). The prediction of our final model was closer to the actual outcome of the 2020 elections, with an error of 3.91%[8] (see the table with the results of the forecasts in the appendix).

Part 2 - Causal Inference

First, we have to identify the main idea behind each party's political position. The Democratic policies generally rely on American liberalism. The main points touched by these ideas are ***Fiscal policy, Minimum wage, Health Care, Education, Environment, Trade Agreements.*** Let us focus, for example, on ***Minimum Wage***. This is a central topic for the Democratic party in an electoral campaign. For years, the Democrats wanted to increase the minimum wage threshold to $15. This policy raised concerns from the Republican party on other important Macroeconomics components such as GDP and unemployment. Moreover, the Republicans generally oppose such a sharp increase, labelling it as "bad for business". The policy debate over raising the federal minimum wage to $15 an hour also reveals fissures in the Republican Party, which is straining to appeal to its corporate backers. Some of them believe that more than doubling the minimum wage would take a toll on their profits. Actually, this policy may have unforeseeable effects on the American system, from politics to the economy. On the other hand, the goal of this policy would be to improve the living conditions of low-skill workers. This category of workers is the most affected by this policy. One can expect an unemployment increase because if these employees' minimum wage increases, some of them might be fired. In order to disentangle causality from correlation on this problem, the biggest issue is to understand how this effect can be treated. Excluding the positive correlation between unemployment and wage is the most challenging part of the American task list. It seems quite impossible, but many theories can be applied to avoid the problem. A significant increase in wages will give us a considerable increase in unemployment. We can face these issues by finding the right percentage of wage increase, allowing the unemployment to stay indifferent or at least to be less correlated. This could be a winning point for each party. The approach of finding a good balance between them in order to decrease the correlation seems the best one. To solve the problem stated before, we need to measure such a case to establish which percentage increase of wage could provide an empirical approach to choose the best percentage of wage in order to disentangle the strong correlation among the latter and unemployment. One of the models that could well explain the interpretation of our input is the DiD model. For the coefficient *β8* (***Death Penalty***)*,* we found a positive value (4.36927%), as expected, since passing from the reference category ("not abolished") to the other ("abolished"), there is a preference in voting Democratics instead of Republicans (the democratic vote share increase by 4.36927%). The estimate for this coefficient is statistically different from zero (pval= 0.0). For the coefficient *β7* (***Unemployment***), we found a negative value (-0.3463784%), but it is not statistically significant (since the pval=0.219). Since Democrats' policies are more effective in decreasing the unemployment rate, when

[8] See the table with the results of the forecasts in the appendix

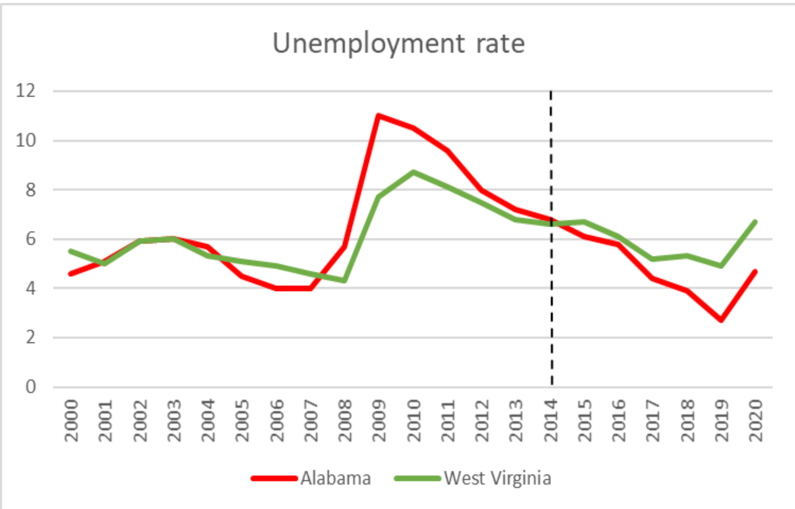
the unemployment is high, the Democratic candidates have more chances to be elected (higher Democratic share). Therefore, we expect a positive value for this coefficient.
The other variables provide more or less the same results. Since many variables are still not significant (both in models 1 and 2), we are not satisfied with the results obtained. We ran the LASSO on this model to keep only the significant variables to explain the Democratic Share.

$$\textbf{DiD} = \overline{Y}_{post}^{treatment} - \overline{Y}_{pre}^{treatment} - (\overline{Y}_{post}^{comparison} - \overline{Y}_{pre}^{comparison})$$

$$\textbf{Yi,t} = \alpha + \beta D_i + \zeta Post_t + \delta (D_i * Post_t) + \varepsilon_{i,t}$$

We applied this method at a State level, comparing two States (one that implemented the increase in the minimum wage in 2014 and one that did not). In the USA, each State can set its minimum wage. That is why this analysis was possible. To understand how a policy that increases the minimum wage can affect the unemployment rate and disentangle causation from correlation, we used the DiD model to compare Alabama and West Virginia (*Graph 3*). Alabama was the control group, whereas West Virginia was the treated group. We considered the minimum-wage law in 2014 in West Virginia, and we looked Before and After this date. In 2014 the minimum wage in West Virginia was increased from $8.00 to $8.75 (hourly)[9]. In the same year, the minimum wage in Alabama did not change and remained constant at $7.25. It is interesting to notice that in 2014 the Governor in Alabama was Republican, while the one in West Virginia was Democratic. As previously said, Democrats tend to increase the minimum wage (or at least they would like to do so), whereas Republicans tend to oppose this policy. Therefore, in this case, being a Democratic state or a Republican State meant implementing the policy or not. As we can see from the graph that compares the Unemployment Rate of the two States between 2000 and 2020, the common trend assumption (on which the DiD was built) appears reliable. As we can see, after 2014, the unemployment in West Virginia presented higher values than the ones in Alabama. The second formula represents the DiD in a regression framework. Both *D* and *Post* are dummy variables. In a panel data analysis related to this case, *D* takes value 0 when the State is Alabama and value 1 when the State is West Virginia. At the same time, the variable *Post* takes value 0 before 2014 and 1 after 2014. The output *Y*, in this case, is the unemployment rate of the two States from 2000 to 2020. In this case, an increase in the minimum wage corresponded to an increase in the unemployment rate. Why do we care about this? Having a Republican or Democratic governor meant implementing or not the minimum wage policy and the subsequent effects on the unemployment rate. In this case, disentangling causation from correlation could be relevant policy-wise to understand the drawbacks of the implemented policy. Finding a balance between the increment in the minimum wage and unemployment is crucial since a high unemployment rate would lead to complications in other economic variables (such as a decrease in GDP). As seen in Part 1 (microeconomics) of this project, there is a positive correlation between GDP growth and voting share. Therefore, the decrease in GDP would be due to a policy implemented by the Democrats (in this case, the increase in the minimum wage, which would lead to a loss in voting share, which is absolutely what a party does not want).

Graph 3. DiD Unemployment rate (2000-2020) Alabama and West Virginia



[9] https://www.epi.org/minimum-wage-tracker/#/min_wage/

Appendix

We thought it would have been interesting to include in our analysis data regarding the demographic composition of the various states. Among the data of our interest were the percentage of minorities, young people, and women. These data would have been significant if the voting percentage of these categories were available.

Unfortunately, even after contacting the United States Census Bureau and the National Education Association's Research Department, we could not obtain these data for the time frame that affects our panel data. Other data of our interest, such as gun ownership per capita and the percentage of the Catholic population, were also not available. Our goal was to carry out an analysis like the one published by Pew Research Center[10], which showed that the young people, female and educated population, tend to vote democratic. In contrast, the Catholic and pro-gun population tends to vote Republican.

[10] In Changing US Electorate, Race and Education Remain Stark Dividing Lines, Pew Research Center, 2 june 2020.

- ***Output regression Model 1 (Fair Model Rearranged)***

Source	SS	df	MS			
				Number of obs	=	510
				F(56, 453)	=	47.23
Model	53887.7899	56	962.281963	Prob > F	=	0.0000
Residual	9229.17228	453	20.3734487	R-squared	=	0.8538
				Adj R-squared	=	0.8357
Total	63116.9622	509	124.00189	Root MSE	=	4.5137

dem_share	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
g_i	-.0056289	.05999	-0.09	0.925	-.1235221	.1122644
p_i	.7702654	.1232831	6.25	0.000	.5279877	1.012543
z_i	.2069457	.2267339	0.91	0.362	-.2386352	.6525265
it	-4.771689	.5033062	-9.48	0.000	-5.760794	-3.782584
dpert	2.149338	.669919	3.21	0.001	.8328036	3.465873
durtfair	4.051801	.6509512	6.22	0.000	2.772543	5.33106
S2	-.8586598	2.028892	-0.42	0.672	-4.845867	3.128547
S3	5.163128	2.019224	2.56	0.011	1.19492	9.131336
S4	4.097141	2.019788	2.03	0.043	.1278242	8.066458
S5	17.15852	2.0216	8.49	0.000	13.18564	21.1314
S6	9.256448	2.026978	4.57	0.000	5.273001	13.2399
S7	15.21258	2.019254	7.53	0.000	11.24431	19.18084
S8	14.51511	2.021214	7.18	0.000	10.54299	18.48723
S9	51.18079	2.020084	25.34	0.000	47.21089	55.15068
S10	7.178657	2.019686	3.55	0.000	3.20954	11.14777
S11	5.40259	2.020005	2.67	0.008	1.432847	9.372333
S12	21.31912	2.027367	10.52	0.000	17.3349	25.30333
S13	-6.31782	2.019201	-3.13	0.002	-10.28598	-2.349656
S14	15.90035	2.019022	7.88	0.000	11.93254	19.86816
S15	2.614989	2.019366	1.29	0.196	-1.353499	6.583476
S16	10.62455	2.020458	5.26	0.000	6.653913	14.59518
S17	-.0959906	2.019975	-0.05	0.962	-4.065676	3.873694
S18	1.724247	2.020842	0.85	0.394	-2.247141	5.695635
S19	4.0673	2.018719	2.01	0.045	.1000833	8.034516
S20	13.06055	2.018723	6.47	0.000	9.093324	17.02777
S21	18.4861	2.019108	9.16	0.000	14.51812	22.45408
S22	21.6134	2.020729	10.70	0.000	17.64223	25.58456
S23	11.4326	2.020404	5.66	0.000	7.462069	15.40312
S24	13.45989	2.018959	6.67	0.000	9.492205	17.42758
S25	1.913605	2.018903	0.95	0.344	-2.053973	5.881183
S26	6.795949	2.01994	3.36	0.001	2.826334	10.76556
S27	3.176753	2.019123	1.57	0.116	-.7912563	7.144762
S28	-3.088461	2.018912	-1.53	0.127	-7.056057	.8791352
S29	8.345137	2.02085	4.13	0.000	4.373733	12.31654
S30	8.447036	2.01912	4.18	0.000	4.479031	12.41504
S31	13.66038	2.019	6.77	0.000	9.692609	17.62814
S32	11.88154	2.020352	5.88	0.000	7.911111	15.85196
S33	20.97511	2.021894	10.37	0.000	17.00166	24.94857
S34	6.125633	2.018972	3.03	0.003	2.157919	10.09335
S35	-1.877053	2.027704	-0.93	0.355	-5.861927	2.107821
S36	7.959567	2.019608	3.94	0.000	3.990604	11.92853
S37	-3.455441	2.019108	-1.71	0.088	-7.423422	.5125399
S38	13.7922	2.02209	6.82	0.000	9.818364	17.76605
S39	11.64972	2.018707	5.77	0.000	7.682522	15.61691
S40	20.77476	2.018719	10.29	0.000	16.80755	24.74198
S41	2.651057	2.019014	1.31	0.190	-1.316738	6.618853
S42	1.135191	2.018975	0.56	0.574	-2.832528	5.102911
S43	3.280311	2.018714	1.62	0.105	-.6868946	7.247516
S44	2.957412	2.01993	1.46	0.144	-1.012183	6.927008
S45	-7.65653	2.019586	-3.79	0.000	-11.62545	-3.687611
S46	19.91434	2.020336	9.86	0.000	15.94395	23.88473
S47	7.857427	2.019909	3.89	0.000	3.887872	11.82698
S48	14.97204	2.019687	7.41	0.000	11.00292	18.94116
S49	3.806187	2.020809	1.88	0.060	-.1651357	7.777509
S50	11.29606	2.021093	5.59	0.000	7.324178	15.26794
S51	-7.100809	2.020035	-3.52	0.000	-11.07061	-3.131006
_cons	40.37977	1.429878	28.24	0.000	37.56975	43.18979

- ***Forecast 2012 with Model 1 (Forecasted vs Real values)***

STATE	***Democratic votes (EST)***	***Republican votes (EST)***	***Democratic votes (Real)***	***Republican votes (Real)***
ALABAMA	0	9	0	9
ALASKA	0	3	0	3
ARIZONA	0	11	0	11
ARKANSAS	0	6	0	6
CALIFORNIA	55	0	55	0
COLORADO	9	0	9	0
CONNECTICUT	7	0	7	0
DELAWARE	3	0	3	0
DISTRICT OF COLUMBIA	3	0	3	0
FLORIDA	29	0	29	0
GEORGIA	0	16	0	16
HAWAII	4	0	4	0
IDAHO	0	4	0	4
ILLINOIS	20	0	20	0
INDIANA	0	11	0	11
IOWA	6	0	6	0
KANSAS	0	6	0	6
KENTUCKY	0	8	0	8
LOUISIANA	0	8	0	8
MAINE	4	0	4	0
MARYLAND	10	0	10	0
MASSACHUSETTS	11	0	11	0
MICHIGAN	16	0	16	0
MINNESOTA	10	0	10	0
MISSISSIPPI	0	6	0	6
MISSOURI	10	0	0	10
MONTANA	0	3	0	3
NEBRASKA	0	5	0	5
NEVADA	6	0	6	0
NEW HAMPSHIRE	4	0	4	0
NEW JERSEY	14	0	14	0
NEW MEXICO	5	0	5	0
NEW YORK	29	0	29	0
NORTH CAROLINA	0	15	0	15
NORTH DAKOTA	0	3	0	3
OHIO	18	0	18	0
OKLAHOMA	0	7	0	7
OREGON	7	0	7	0
PENNSYLVANIA	20	0	20	0
RHODE ISLAND	4	0	4	0
SOUTH CAROLINA	0	9	0	9
SOUTH DAKOTA	0	3	0	3
TENNESSEE	0	11	0	11
TEXAS	0	38	0	38
UTAH	0	6	0	6
VERMONT	3	0	3	0
VIRGINIA	13	0	13	0
WASHINGTON	12	0	12	0
WEST VIRGINIA	0	5	0	5
WISCONSIN	10	0	10	0
WYOMING	0	3	0	3
Total electoral votes	**342**	**196**	**332**	**206**
% of electoral votes	**63.57%**	**36.43%**	**61.71%**	**38.29%**
Winner	*Democratic (Obama)*		*Democratic (Obama)*	

- ***Forecast 2020 with Model 1 (Forecasted vs Real values)***

STATE	Democratic votes (EST)	Republican votes (EST)	Democratic votes (Real)	Republican votes (Real)
ALABAMA	0	9	0	9
ALASKA	0	3	0	3
ARIZONA	0	11	11	0
ARKANSAS	0	6	0	6
CALIFORNIA	55	0	55	0
COLORADO	9	0	9	0
CONNECTICUT	7	0	7	0
DELAWARE	3	0	3	0
DISTRICT OF COLUMBIA	3	0	3	0
FLORIDA	29	0	0	29
GEORGIA	0	16	16	0
HAWAII	4	0	4	0
IDAHO	0	4	0	4
ILLINOIS	20	0	20	0
INDIANA	0	11	0	11
IOWA	6	0	0	6
KANSAS	0	6	0	6
KENTUCKY	0	8	0	8
LOUISIANA	0	8	0	8
MAINE	4	0	4	0
MARYLAND	10	0	10	0
MASSACHUSETTS	11	0	11	0
MICHIGAN	16	0	16	0
MINNESOTA	10	0	10	0
MISSISSIPPI	0	6	0	6
MISSOURI	10	0	0	10
MONTANA	0	3	0	3
NEBRASKA	0	5	0	5
NEVADA	6	0	6	0
NEW HAMPSHIRE	4	0	4	0
NEW JERSEY	14	0	14	0
NEW MEXICO	5	0	5	0
NEW YORK	29	0	29	0
NORTH CAROLINA	0	15	0	15
NORTH DAKOTA	0	3	0	3
OHIO	18	0	0	18
OKLAHOMA	0	7	0	7
OREGON	7	0	7	0
PENNSYLVANIA	20	0	20	0
RHODE ISLAND	4	0	4	0
SOUTH CAROLINA	0	9	0	9
SOUTH DAKOTA	0	3	0	3
TENNESSEE	0	11	0	11
TEXAS	0	38	0	38
UTAH	0	6	0	6
VERMONT	3	0	3	0
VIRGINIA	13	0	13	0
WASHINGTON	12	0	12	0
WEST VIRGINIA	0	5	0	5
WISCONSIN	10	0	10	0
WYOMING	0	3	0	3
Total electoral votes	***342***	***196***	***306***	***232***
% electoral votes	***63.57%***	***36.43%***	***56.88%***	***43.12%***
Winner	***Democratic (Biden)***		***Democratic (Biden)***	

- ***Output regression Model 2 (Fair Model Rearranged + Unemployment + DeathPenalty)***

Source	SS	df	MS
Model	54206.3509	58	934.592258
Residual	8910.61128	451	19.757453
Total	63116.9622	509	124.00189

Number of obs	=	510
F(58, 451)	=	47.30
Prob > F	=	0.0000
R-squared	=	0.8588
Adj R-squared	=	0.8407
Root MSE	=	4.4449

dem_share	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
g_i	-.0235871	.0646257	-0.36	0.715	-.150592	.1034178
p_i	.7829718	.1269492	6.17	0.000	.5334865	1.032457
z_i	.3130693	.2275034	1.38	0.169	-.1340292	.7601677
it	-5.050013	.5955751	-8.48	0.000	-6.22046	-3.879566
dpert	2.235444	.6706075	3.33	0.001	.9175404	3.553347
durtfair	4.019479	.6961896	5.77	0.000	2.6513	5.387657
unempl	-.3463784	.4665658	-0.74	0.458	-1.263291	.5705345
deathpenalty	4.36927	1.088126	4.02	0.000	2.230843	6.507697
S2	-5.144726	2.265968	-2.27	0.024	-9.597892	-.6915602
S3	5.161752	1.98849	2.60	0.010	1.253896	9.069608
S4	-.2784671	2.268077	-0.12	0.902	-4.735777	4.178843
S5	17.25676	1.991461	8.67	0.000	13.34306	21.17045
S6	8.808175	2.000475	4.40	0.000	4.876765	12.73958
S7	13.98248	2.012882	6.95	0.000	10.02668	17.93827
S8	13.62343	2.003543	6.80	0.000	9.685993	17.56087
S9	46.85971	2.262347	20.71	0.000	42.41366	51.30576
S10	7.21095	1.98897	3.63	0.000	3.302152	11.11975
S11	5.410279	1.98927	2.72	0.007	1.50089	9.319668
S12	17.29418	2.25251	7.68	0.000	12.86746	21.7209
S13	-6.325827	1.988516	-3.18	0.002	-10.23373	-2.41792
S14	14.62915	2.014448	7.26	0.000	10.67028	18.58802
S15	2.574741	1.988636	1.29	0.196	-1.333402	6.482885
S16	6.274488	2.265798	2.77	0.006	1.821656	10.72732
S17	-.0875771	1.989672	-0.04	0.965	-3.997755	3.822601
S18	1.660616	1.990128	0.83	0.404	-2.250459	5.571691
S19	4.074776	1.988103	2.05	0.041	.1676803	7.981871
S20	9.10089	2.219289	4.10	0.000	4.739459	13.46232
S21	17.64922	1.999849	8.83	0.000	13.71904	21.5794
S22	17.26628	2.26627	7.62	0.000	12.81252	21.72004
S23	7.062241	2.267921	3.11	0.002	2.605238	11.51924
S24	9.100326	2.265462	4.02	0.000	4.648155	13.5525
S25	1.889539	1.988434	0.95	0.342	-2.018207	5.797284
S26	6.741885	1.98923	3.39	0.001	2.832575	10.65119
S27	3.1955	1.988479	1.61	0.109	-.712335	7.103335
S28	-3.102966	1.988193	-1.56	0.119	-7.010237	.8043059
S29	8.440053	1.99058	4.24	0.000	4.52809	12.35202
S30	7.998918	1.991811	4.02	0.000	4.084535	11.9133
S31	11.95921	2.033126	5.88	0.000	7.963632	15.95479
S32	10.12286	2.037821	4.97	0.000	6.118054	14.12766
S33	19.27725	2.03651	9.47	0.000	15.27502	23.27948
S34	6.156442	1.988587	3.10	0.002	2.248396	10.06449
S35	-6.171598	2.265716	-2.72	0.007	-10.62427	-1.718928
S36	7.910973	1.988945	3.98	0.000	4.002224	11.81972
S37	-3.500556	1.988988	-1.76	0.079	-7.409389	.4082782
S38	13.82258	1.991302	6.94	0.000	9.909196	17.73596
S39	11.66704	1.987985	5.87	0.000	7.760179	15.57391
S40	16.46044	2.260512	7.28	0.000	12.018	20.90289
S41	2.693448	1.988344	1.35	0.176	-1.214121	6.601017
S42	1.08706	1.988655	0.55	0.585	-2.82112	4.995241
S43	3.27449	1.987972	1.65	0.100	-.6323476	7.181328
S44	2.952763	1.989175	1.48	0.138	-.9564391	6.861964
S45	-7.668292	1.990209	-3.85	0.000	-11.57953	-3.757057
S46	15.55137	2.267401	6.86	0.000	11.09539	20.00736
S47	7.883559	1.989643	3.96	0.000	3.973437	11.79368
S48	14.55748	1.991695	7.31	0.000	10.64332	18.47163
S49	-.6358964	2.276853	-0.28	0.780	-5.110453	3.838661
S50	6.820384	2.282426	2.99	0.003	2.334874	11.30589
S51	-7.117766	1.989373	-3.58	0.000	-11.02736	-3.208176
_cons	40.35196	1.40832	28.65	0.000	37.58428	43.11965

- ***Output regression Model 3 (Final model obtained after DS LASSO)***

Source	SS	df	MS			
				Number of obs	=	510
				F(56, 453)	=	40.46
Model	52600.8273	56	939.300488	Prob > F	=	0.0000
Residual	10516.1349	453	23.2144259	R-squared	=	0.8334
				Adj R-squared	=	0.8128
Total	63116.9622	509	124.00189	Root MSE	=	4.8181

dem_share	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
g_i	.0324557	.069694	0.47	0.642	-.104508	.1694195
dpert	2.068809	.7264521	2.85	0.005	.6411746	3.496443
durtfair	1.781064	.6900494	2.58	0.010	.4249685	3.137159
unempl	2.017086	.4165766	4.84	0.000	1.198423	2.835748
p_i	.0649079	.0997051	0.65	0.515	-.131034	.2608498
z_i	-.0386569	.2426991	-0.16	0.874	-.5156128	.4382989
S2	-1.356488	2.165691	-0.63	0.531	-5.612536	2.899559
S3	5.071975	2.15541	2.35	0.019	.836132	9.307818
S4	4.23189	2.15612	1.96	0.050	-.0053489	8.469129
S5	16.63104	2.157339	7.71	0.000	12.3914	20.87067
S6	8.947671	2.164648	4.13	0.000	4.693673	13.20167
S7	14.80361	2.155742	6.87	0.000	10.56712	19.04011
S8	14.33267	2.158225	6.64	0.000	10.09129	18.57404
S9	50.69188	2.156026	23.51	0.000	46.45483	54.92894
S10	6.934167	2.155697	3.22	0.001	2.697761	11.17057
S11	5.292264	2.156238	2.45	0.014	1.054793	9.529735
S12	19.50314	2.172487	8.98	0.000	15.23374	23.77255
S13	-6.255396	2.155456	-2.90	0.004	-10.49133	-2.019463
S14	15.63279	2.15612	7.25	0.000	11.39555	19.87002
S15	2.858046	2.155329	1.33	0.185	-1.377637	7.09373
S16	10.80829	2.157028	5.01	0.000	6.569268	15.04731
S17	-.0223077	2.156701	-0.01	0.992	-4.260688	4.216073
S18	1.989955	2.156857	0.92	0.357	-2.248732	6.228641
S19	4.077822	2.155024	1.89	0.059	-.1572628	8.312906
S20	13.18522	2.155072	6.12	0.000	8.950041	17.4204
S21	18.13397	2.155419	8.41	0.000	13.8981	22.36983
S22	21.31972	2.157676	9.88	0.000	17.07942	25.56002
S23	11.41108	2.156722	5.29	0.000	7.172658	15.6495
S24	13.43454	2.155263	6.23	0.000	9.198985	17.67009
S25	2.144584	2.155151	1.00	0.320	-2.09075	6.379917
S26	6.989342	2.156041	3.24	0.001	2.752259	11.22643
S27	3.205317	2.155424	1.49	0.138	-1.030553	7.441187
S28	-2.87173	2.154922	-1.33	0.183	-7.106614	1.363155
S29	7.849148	2.156529	3.64	0.000	3.611105	12.08719
S30	8.507001	2.155643	3.95	0.000	4.270699	12.7433
S31	13.39971	2.155148	6.22	0.000	9.164376	17.63503
S32	11.66853	2.157076	5.41	0.000	7.42941	15.90765
S33	20.52592	2.15851	9.51	0.000	16.28399	24.76786
S34	5.841951	2.155196	2.71	0.007	1.606528	10.07737
S35	-1.925363	2.165164	-0.89	0.374	-6.180375	2.32965
S36	8.189047	2.155677	3.80	0.000	3.952679	12.42542
S37	-3.250254	2.155774	-1.51	0.132	-7.486812	.986303
S38	13.56218	2.158256	6.28	0.000	9.320747	17.80362
S39	11.51902	2.154821	5.35	0.000	7.284338	15.75371
S40	20.42537	2.155233	9.48	0.000	16.18987	24.66086
S41	2.510444	2.155173	1.16	0.245	-1.724935	6.745822
S42	1.401925	2.155287	0.65	0.516	-2.833675	5.637526
S43	3.338471	2.154869	1.55	0.122	-.8963083	7.57325
S44	2.870418	2.156155	1.33	0.184	-1.366889	7.107725
S45	-7.519873	2.157228	-3.49	0.001	-11.75929	-3.280458
S46	19.9707	2.157249	9.26	0.000	15.73124	24.21016
S47	7.48404	2.156099	3.47	0.001	3.246843	11.72124
S48	14.84766	2.155954	6.89	0.000	10.61074	19.08457
S49	4.176067	2.156564	1.94	0.053	-.0620442	8.414179
S50	11.85299	2.157798	5.49	0.000	7.612458	16.09353
S51	-7.149947	2.156386	-3.32	0.001	-11.38771	-2.912185
_cons	40.40472	1.526541	26.47	0.000	37.40474	43.4047

- ***Forecast 2012 with Model 3 (Forecasted vs Real values)***

STATE	Democratic votes (EST)	Republican votes (EST)	Democratic votes (Real)	Republican votes (Real)
ALABAMA	0	9	0	9
ALASKA	0	3	0	3
ARIZONA	0	11	0	11
ARKANSAS	0	6	0	6
CALIFORNIA	55	0	55	0
COLORADO	9	0	9	0
CONNECTICUT	7	0	7	0
DELAWARE	3	0	3	0
DISTRICT OF COLUMBIA	3	0	3	0
FLORIDA	29	0	29	0
GEORGIA	0	16	0	16
HAWAII	4	0	4	0
IDAHO	0	4	0	4
ILLINOIS	20	0	20	0
INDIANA	0	11	0	11
IOWA	6	0	6	0
KANSAS	0	6	0	6
KENTUCKY	0	8	0	8
LOUISIANA	0	8	0	8
MAINE	4	0	4	0
MARYLAND	10	0	10	0
MASSACHUSETTS	11	0	11	0
MICHIGAN	16	0	16	0
MINNESOTA	10	0	10	0
MISSISSIPPI	0	6	0	6
MISSOURI	10	0	0	10
MONTANA	0	3	0	3
NEBRASKA	0	5	0	5
NEVADA	6	0	6	0
NEW HAMPSHIRE	4	0	4	0
NEW JERSEY	14	0	14	0
NEW MEXICO	5	0	5	0
NEW YORK	29	0	29	0
NORTH CAROLINA	0	15	0	15
NORTH DAKOTA	0	3	0	3
OHIO	18	0	18	0
OKLAHOMA	0	7	0	7
OREGON	7	0	7	0
PENNSYLVANIA	20	0	20	0
RHODE ISLAND	4	0	4	0
SOUTH CAROLINA	0	9	0	9
SOUTH DAKOTA	0	3	0	3
TENNESSEE	0	11	0	11
TEXAS	0	38	0	38
UTAH	0	6	0	6
VERMONT	3	0	3	0
VIRGINIA	13	0	13	0
WASHINGTON	12	0	12	0
WEST VIRGINIA	0	5	0	5
WISCONSIN	10	0	10	0
WYOMING	0	3	0	3
Total Electoral votes	***342***	***196***	***332***	***206***
% Electoral votes	***63.57%***	***36.43%***	***61.71%***	***38.29%***
Winner	***Democratic (Obama)***		***Democratic (Obama)***	

- ***Forecast 2020with Model 3 (Forecasted vs Real values)***

STATE	Democratic votes (EST)	Republican votes (EST)	Democratic votes (Real)	Republican votes (Real)
ALABAMA	0	9	0	9
ALASKA	0	3	0	3
ARIZONA	0	11	11	0
ARKANSAS	0	6	0	6
CALIFORNIA	55	0	55	0
COLORADO	9	0	9	0
CONNECTICUT	7	0	7	0
DELAWARE	3	0	3	0
DISTRICT OF COLUMBIA	3	0	3	0
FLORIDA	0	29	0	29
GEORGIA	0	16	16	0
HAWAII	4	0	4	0
IDAHO	0	4	0	4
ILLINOIS	20	0	20	0
INDIANA	0	11	0	11
IOWA	6	0	0	6
KANSAS	0	6	0	6
KENTUCKY	0	8	0	8
LOUISIANA	0	8	0	8
MAINE	4	0	4	0
MARYLAND	10	0	10	0
MASSACHUSETTS	11	0	11	0
MICHIGAN	16	0	16	0
MINNESOTA	10	0	10	0
MISSISSIPPI	0	6	0	6
MISSOURI	0	10	0	10
MONTANA	0	3	0	3
NEBRASKA	0	5	0	5
NEVADA	6	0	6	0
NEW HAMPSHIRE	4	0	4	0
NEW JERSEY	14	0	14	0
NEW MEXICO	5	0	5	0
NEW YORK	29	0	29	0
NORTH CAROLINA	0	15	0	15
NORTH DAKOTA	0	3	0	3
OHIO	0	18	0	18
OKLAHOMA	0	7	0	7
OREGON	7	0	7	0
PENNSYLVANIA	20	0	20	0
RHODE ISLAND	4	0	4	0
SOUTH CAROLINA	0	9	0	9
SOUTH DAKOTA	0	3	0	3
TENNESSEE	0	11	0	11
TEXAS	0	38	0	38
UTAH	0	6	0	6
VERMONT	3	0	3	0
VIRGINIA	13	0	13	0
WASHINGTON	12	0	12	0
WEST VIRGINIA	0	5	0	5
WISCONSIN	10	0	10	0
WYOMING	0	3	0	3
Total electoral votes	***285***	***253***	***306***	***232***
% electoral votes	***52.97%***	***47.03%***	***56.88%***	***43.12%***
Winner	***Democratic (Biden)***		***Democratic (Biden)***	

- ***Model obtained with Naive Post-Selection LASSO***

Source	SS	df	MS			
				Number of obs	=	510
				F(54, 455)	=	42.09
Model	52589.1392	54	973.872948	Prob > F	=	0.0000
Residual	10527.823	455	23.1380726	R-squared	=	0.8332
				Adj R-squared	=	0.8134
Total	63116.9622	509	124.00189	Root MSE	=	4.8102

dem_share	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
g_i	.0473737	.0577252	0.82	0.412	-.0660674	.1608149
dpert	2.14385	.6703874	3.20	0.001	.8264108	3.46129
durtfair	1.805397	.6136809	2.94	0.003	.5993964	3.011397
unempl	1.927155	.3683926	5.23	0.000	1.203194	2.651117
S2	-1.413777	2.156307	-0.66	0.512	-5.651333	2.82378
S3	5.053341	2.151357	2.35	0.019	.8255135	9.281169
S4	4.251783	2.151576	1.98	0.049	.0235237	8.480042
S5	16.61418	2.153505	7.71	0.000	12.38213	20.84623
S6	8.895401	2.152317	4.13	0.000	4.665687	13.12512
S7	14.81865	2.151872	6.89	0.000	10.58981	19.04749
S8	14.31111	2.151363	6.65	0.000	10.08327	18.53895
S9	50.6958	2.152141	23.56	0.000	46.46643	54.92516
S10	6.911621	2.15191	3.21	0.001	2.682707	11.14054
S11	5.268151	2.15191	2.45	0.015	1.039236	9.497066
S12	19.5645	2.165476	9.03	0.000	15.30893	23.82008
S13	-6.271756	2.15171	-2.91	0.004	-10.50028	-2.043235
S14	15.63397	2.151846	7.27	0.000	11.40518	19.86276
S15	2.86305	2.151265	1.33	0.184	-1.364599	7.090698
S16	10.83252	2.151261	5.04	0.000	6.604881	15.06016
S17	-.041536	2.152793	-0.02	0.985	-4.272186	4.189114
S18	1.998292	2.152051	0.93	0.354	-2.230901	6.227484
S19	4.087782	2.151283	1.90	0.058	-.139901	8.315465
S20	13.17063	2.151415	6.12	0.000	8.942684	17.39857
S21	18.11988	2.151529	8.42	0.000	13.89171	22.34805
S22	21.29743	2.15191	9.90	0.000	17.06851	25.52634
S23	11.3904	2.151978	5.29	0.000	7.161346	15.61944
S24	13.43355	2.151395	6.24	0.000	9.20565	17.66145
S25	2.155326	2.151377	1.00	0.317	-2.072542	6.383194
S26	6.977042	2.151272	3.24	0.001	2.749382	11.2047
S27	3.208064	2.151404	1.49	0.137	-1.019856	7.435984
S28	-2.85439	2.151233	-1.33	0.185	-7.081975	1.373196
S29	7.844445	2.152245	3.64	0.000	3.614872	12.07402
S30	8.485215	2.151864	3.94	0.000	4.25639	12.71404
S31	13.39563	2.151589	6.23	0.000	9.167345	17.62391
S32	11.64756	2.151325	5.41	0.000	7.419791	15.87532
S33	20.49209	2.153042	9.52	0.000	16.26095	24.72323
S34	5.847617	2.151593	2.72	0.007	1.619325	10.07591
S35	-1.954339	2.16012	-0.90	0.366	-6.199388	2.290709
S36	8.177945	2.151273	3.80	0.000	3.95028	12.40561
S37	-3.275816	2.151851	-1.52	0.129	-7.504615	.9529829
S38	13.5207	2.153799	6.28	0.000	9.288069	17.75332
S39	11.51248	2.151254	5.35	0.000	7.284851	15.7401
S40	20.433	2.151658	9.50	0.000	16.20458	24.66142
S41	2.518122	2.151386	1.17	0.242	-1.709764	6.746008
S42	1.388572	2.151507	0.65	0.519	-2.83955	5.616695
S43	3.347938	2.151267	1.56	0.120	-.879714	7.575589
S44	2.843837	2.151639	1.32	0.187	-1.384546	7.072219
S45	-7.541152	2.152465	-3.50	0.001	-11.77116	-3.311146
S46	19.97679	2.151435	9.29	0.000	15.74881	24.20478
S47	7.455844	2.151557	3.47	0.001	3.227622	11.68407
S48	14.82628	2.152001	6.89	0.000	10.59718	19.05537
S49	4.186328	2.152004	1.95	0.052	-.0427723	8.415427
S50	11.82106	2.152562	5.49	0.000	7.590867	16.05126
S51	-7.136203	2.152748	-3.31	0.001	-11.36676	-2.905642
_cons	40.38407	1.523343	26.51	0.000	37.39041	43.37773

- ***Comparison between variables (G and P) calculated with quarterly and yearly variables***

Year	State	G (15th quarter)	P (15th quarter)	G (16th quarter)	P (16th quarter)
2008	ALABAMA	-2.952265582	2.493218635	-2.709069768	2.976632203
2012	ALABAMA	-1.757163691	1.581079237	0.691521939	1.850190536
2016	ALABAMA	1.845018995	1.539000923	1.724035375	1.753243873
2020	ALABAMA	-2.098291004	1.929302445	-4.575844669	2.063401238
2008	ALASKA	-2.112341967	7.83285853	-2.5755573	8.163863338
2012	ALASKA	0.817039137	0.729628403	5.65533459	1.922647123
2016	ALASKA	-1.555105091	1.652905603	-1.936545351	1.772586991
2020	ALASKA	-6.835892609	0.917086823	-7.447380867	1.619558482
2008	ARIZONA	-4.189231067	2.296997893	-3.504896103	2.761778487
2012	ARIZONA	-1.519207616	1.424111582	0.955754365	1.770314655
2016	ARIZONA	1.103078519	1.776478489	1.893031921	1.943843208
2020	ARIZONA	-1.865603608	2.062968313	-3.663041011	2.190836364
2008	ARKANSAS	-6.67579391	2.678465991	-2.506976341	3.007185926
2012	ARKANSAS	-1.730318153	1.58255985	0.154024779	1.782466859
2016	ARKANSAS	-0.23403775	1.193606287	-0.010035158	1.472167218
2020	ARKANSAS	0.142204224	1.892394956	-2.514044434	1.994235121
2008	CALIFORNIA	-0.63233687	2.089143338	0.033149027	2.557261588
2012	CALIFORNIA	1.747787372	1.423430103	2.079329979	1.569419502
2016	CALIFORNIA	2.746600195	1.310063235	3.137256789	1.5245352
2020	CALIFORNIA	-0.146585111	1.62205901	-3.439475636	1.736590177
2008	COLORADO	-2.77783356	2.849032425	-0.331808024	3.226758485
2012	COLORADO	-0.616047633	1.36299598	0.637323959	1.22659807
2016	COLORADO	-0.382741481	0.787325311	0.595210002	0.904870282
2020	COLORADO	-2.636558044	1.560289809	-5.085246971	1.794541166
2008	CONNECTICUT	-1.733759804	2.28128537	-0.78904013	2.763770139
2012	CONNECTICUT	0.511414521	1.589887228	1.079467794	2.042290262
2016	CONNECTICUT	0.885065505	1.933682871	0.645227629	2.168752878
2020	CONNECTICUT	-6.036504904	2.000565791	-7.876776268	2.104870477
2008	DELAWARE	-2.956131485	2.217073509	-5.74895851	2.722961677
2012	DELAWARE	-3.804616097	1.484218636	-2.334036234	1.50525178
2016	DELAWARE	-6.794925828	2.381811047	-8.524778541	2.595316918
2020	DELAWARE	-2.759204802	2.481236867	-5.856691127	2.806579217
2008	DISTRICT OF COLUMBIA	2.308229241	2.687153813	3.855934141	3.393850437
2012	DISTRICT OF COLUMBIA	-2.140590515	1.662682872	-2.636072628	1.898984797
2016	DISTRICT OF COLUMBIA	0.818244391	2.017670504	1.140827961	2.164548069
2020	DISTRICT OF COLUMBIA	-1.455796208	2.162239136	-3.643546916	2.334374209
2008	FLORIDA	-6.660434207	2.316462947	-6.379089106	2.886298244
2012	FLORIDA	-0.471060564	1.412413239	-0.589390199	1.797154558
2016	FLORIDA	0.212884096	1.880047452	1.858979341	2.110366464
2020	FLORIDA	-2.937100213	2.061702153	-5.193576411	2.158840007
2008	GEORGIA	-5.953599783	2.078368757	-5.05668903	2.508852829
2012	GEORGIA	-0.387279225	1.457408831	0.489103351	1.835263549
2016	GEORGIA	2.243482371	1.815198345	3.142272077	2.082707623
2020	GEORGIA	-3.769082489	1.798532559	-6.088764898	1.917440033
2008	HAWAII	-3.085779151	2.566658	-0.939327812	3.247977583
2012	HAWAII	0.077610814	1.520487599	0.855504319	1.892062686
2016	HAWAII	0.930486769	2.001778115	1.965988541	2.212189245
2020	HAWAII	-12.19523404	1.972372951	-13.43184242	2.1020396

2008	IDAHO	-0.097601459	2.075549127	1.115664449	2.306427794
2012	IDAHO	-1.133420143	1.710061323	-0.749520713	1.899376832
2016	IDAHO	3.107427358	1.078770636	2.483629937	1.342740774
2020	IDAHO	-1.200902333	2.044246359	-3.202924293	2.13651824
2008	ILLINOIS	-1.905923701	2.332044375	-1.650553881	2.911472193
2012	ILLINOIS	1.162877733	1.731157036	2.68125253	2.057572406
2016	ILLINOIS	0.174362136	1.725269201	-0.036859534	1.998183092
2020	ILLINOIS	-4.044775393	1.955844263	-6.01494066	2.084843325
2008	INDIANA	-4.654473803	2.098809776	-2.291309249	2.536626213
2012	INDIANA	-0.667176671	1.90934869	0.200825574	2.38145879
2016	INDIANA	2.280818301	1.42336701	1.949564486	1.703473728
2020	INDIANA	-0.685121518	1.653525312	-3.786918177	1.770229794
2008	IOWA	-5.105613269	2.171258293	-3.344799707	2.479152325
2012	IOWA	1.87257732	1.963869019	4.914043105	2.316223512
2016	IOWA	-1.16818992	1.281634085	-1.025287532	1.625750278
2020	IOWA	0.977696298	1.953208973	-2.84146304	2.057646567
2008	KANSAS	0.090939286	2.414799409	1.204047344	2.835807229
2012	KANSAS	-2.254590492	1.967898688	1.373621229	2.199523703
2016	KANSAS	3.712419128	0.864479279	4.292541408	1.153711946
2020	KANSAS	0.374591639	1.733620556	-2.882151396	1.824416159
2008	KENTUCKY	-3.514234255	2.617730648	-1.482330946	3.007521732
2012	KENTUCKY	-1.018781271	1.90973131	1.2965887	2.252078805
2016	KENTUCKY	0.517879742	1.500964234	0.536435251	1.749121628
2020	KENTUCKY	-1.03209668	1.781752066	-4.175038781	1.907544734
2008	LOUISIANA	0.598594925	4.752827368	-0.784959652	6.018268026
2012	LOUISIANA	-2.307200932	3.083749461	-0.517043192	2.400389059
2016	LOUISIANA	-1.230932859	0.368040878	-3.014109373	0.156510832
2020	LOUISIANA	-4.875247719	1.385752707	-7.294748272	1.702473367
2008	MAINE	-2.096911273	2.36136476	-2.038353205	2.909823367
2012	MAINE	-1.477805955	1.596473816	-0.232895133	2.015627358
2016	MAINE	1.518385064	1.90754549	2.467295038	2.183902702
2020	MAINE	0.993907065	2.168024373	-1.939878953	2.25457598
2008	MARYLAND	1.967880952	2.292620939	2.727795446	2.856677013
2012	MARYLAND	-1.368253775	1.332313312	-1.054196114	1.706985881
2016	MARYLAND	3.340959796	1.799262745	4.1659095	1.992557111
2020	MARYLAND	-4.167432461	1.969398616	-5.950194271	2.079960884
2008	MASSACHUSETTS	-0.323660417	1.875527679	2.005982696	2.392509157
2012	MASSACHUSETTS	0.484574756	1.318136513	1.566839615	1.611373952
2016	MASSACHUSETTS	-0.054442305	1.933228782	1.487900373	2.140280829
2020	MASSACHUSETTS	-1.290455569	1.910068627	-4.877103162	2.050543878
2008	MICHIGAN	-6.098597602	1.667528214	-6.9973684	2.06294014
2012	MICHIGAN	-0.300802938	1.525190003	2.322839638	2.029538954
2016	MICHIGAN	2.13549216	1.908788866	2.290188801	2.171034825
2020	MICHIGAN	-3.528863896	1.667510501	-5.890876307	1.72809557
2008	MINNESOTA	0.263292232	2.183870691	-0.031280737	2.692552778
2012	MINNESOTA	0.887874627	1.735638302	1.287346671	2.036380481
2016	MINNESOTA	0.970985775	1.360344507	1.126940971	1.61236335
2020	MINNESOTA	-3.709819945	1.902724642	-5.640617295	2.01284222
2008	MISSISSIPPI	0.065090686	2.668286602	1.191314163	3.188324613
2012	MISSISSIPPI	0.215565416	1.788217138	2.374259181	1.993932578
2016	MISSISSIPPI	1.127796768	1.285617516	1.010503476	1.556074615
2020	MISSISSIPPI	1.560086255	1.900861695	-1.930125554	2.022879875
2008	MISSOURI	-1.005491923	2.384639953	2.031970481	2.862811918
2012	MISSOURI	-0.298234293	1.635198157	1.159753911	1.980796001
2016	MISSOURI	0.834295203	1.746150698	-0.307379999	2.026202834
2020	MISSOURI	-1.566030704	1.984761136	-3.966852884	2.094320821
2008	MONTANA	-2.463187647	3.52614792	-1.489269144	4.047984127
2012	MONTANA	-1.480254531	2.307147851	0.5888459	2.254648075
2016	MONTANA	-2.651327895	0.452983202	-4.127549424	0.627141039
2020	MONTANA	-1.535447336	2.119316682	-3.055785672	2.326267999

2008	NEBRASKA	-1.806757917	2.563230679	-1.057989491	2.906431775
2012	NEBRASKA	-3.471680323	2.146070497	-0.739577084	2.472312419
2016	NEBRASKA	0.605891198	0.915535587	0.428719075	1.287591232
2020	NEBRASKA	1.464823713	2.012596984	-1.040281075	2.097078871
2008	NEVADA	-7.888814457	3.140846171	-6.355253657	3.822587287
2012	NEVADA	-2.808786246	1.561679793	-2.918860533	1.859837373
2016	NEVADA	-0.823497377	2.04374187	1.259203176	2.172808077
2020	NEVADA	-10.31414474	2.038582898	-11.56287352	2.223928561
2008	NEW HAMPSHIRE	-2.410037997	2.005512751	-2.322560285	2.545124837
2012	NEW HAMPSHIRE	-0.325223816	1.335318761	1.423723069	1.7526353
2016	NEW HAMPSHIRE	-0.801673545	1.738574114	2.207630179	1.94726249
2020	NEW HAMPSHIRE	-1.009044061	1.982216325	-3.248900314	2.047257797
2008	NEW JERSEY	1.118134564	2.383834731	2.971347426	2.945979524
2012	NEW JERSEY	1.580430707	1.572058562	2.26294905	2.001104625
2016	NEW JERSEY	1.387072759	1.707080339	1.2357948	1.962226397
2020	NEW JERSEY	-3.598496001	1.798256899	-6.008031409	1.906019623
2008	NEW MEXICO	-1.156902334	3.70084474	-1.019761123	3.939687505
2012	NEW MEXICO	-2.233905565	1.346446201	0.056560249	0.709959019
2016	NEW MEXICO	-0.550736893	0.149058123	-0.973495137	0.184386101
2020	NEW MEXICO	-0.155798689	1.110323613	-3.456180125	1.437823502
2008	NEW YORK	0.083055145	2.258739145	-2.57821814	2.820340241
2012	NEW YORK	4.729279422	1.550710582	4.57482975	1.714280033
2016	NEW YORK	4.460734633	2.468677752	3.028474352	2.71215665
2020	NEW YORK	-3.047442965	2.285500104	-5.783388424	2.537814473
2008	NORTH CAROLINA	0.705630464	2.059230742	0.70360206	2.49266622
2012	NORTH CAROLINA	-0.437378466	1.62441476	-0.731263434	1.967497162
2016	NORTH CAROLINA	0.471623085	1.980535137	0.547936534	2.303615391
2020	NORTH CAROLINA	-3.17845571	2.030886957	-5.094985933	2.177622964
2008	NORTH DAKOTA	6.396277144	3.1162241	8.357424737	3.349195281
2012	NORTH DAKOTA	17.21476719	2.13286212	26.81981717	1.940459623
2016	NORTH DAKOTA	-6.591018797	0.670359163	-8.416787376	0.499456512
2020	NORTH DAKOTA	-1.408737909	0.158846568	-4.194586263	0.635816411
2008	OHIO	-2.126502501	2.248085785	-2.504294387	2.754356547
2012	OHIO	-1.048476364	1.755673621	1.0639059	2.096373599
2016	OHIO	0.838636764	1.522724977	0.969473928	1.752867472
2020	OHIO	-1.819583598	1.823303278	-4.572825969	1.976279553
2008	OKLAHOMA	2.015106907	4.753613627	3.975343769	4.995916699
2012	OKLAHOMA	0.691401446	1.431981263	5.773074312	0.268273619
2016	OKLAHOMA	-4.290505592	1.696636047	-3.310026103	1.665280599
2020	OKLAHOMA	-2.585827663	0.726768371	-7.097178584	1.2882153
2008	OREGON	0.286883677	1.302331944	1.969018744	1.68284015
2012	OREGON	-1.390323928	1.095408519	-0.344066165	1.376099911
2016	OREGON	2.125243182	1.507927564	3.182904074	1.738658825
2020	OREGON	-2.673279209	1.843299358	-4.760276958	1.912114599

2008	PENNSYLVANIA	-0.12934807	2.506106092	2.183893861	3.057117322
2012	PENNSYLVANIA	0.616101422	1.501953522	2.033434567	1.847419315
2016	PENNSYLVANIA	1.08281092	1.254391859	1.167475991	1.450758743
2020	PENNSYLVANIA	-3.147829118	1.654563161	-5.83273036	1.808488841
2008	RHODE ISLAND	-0.460709619	2.34184755	-2.023411929	2.936232014
2012	RHODE ISLAND	0.085167386	1.473057176	0.654131013	1.814311354
2016	RHODE ISLAND	-0.337415166	1.962412221	0.093385265	2.17753998
2020	RHODE ISLAND	-1.967735942	2.115893461	-4.383629476	2.234198557
2008	SOUTH CAROLINA	-5.364040908	2.343665584	-3.877893589	2.866825457
2012	SOUTH CAROLINA	-1.057065171	1.500734763	-0.159349236	2.007389664
2016	SOUTH CAROLINA	1.001484307	1.949814829	2.306181607	2.235313592
2020	SOUTH CAROLINA	-1.263099728	2.014680708	-4.003808863	2.102561508
2008	SOUTH DAKOTA	3.343011947	2.236254253	3.397601358	2.487776149
2012	SOUTH DAKOTA	-8.828609601	1.978307653	0.990001002	1.762039766
2016	SOUTH DAKOTA	-1.99841807	1.401113773	-0.205006389	1.737946177
2020	SOUTH DAKOTA	0.013290693	2.343881679	-0.715800035	2.579644116
2008	TENNESSEE	-2.283546986	2.097633994	-0.879080931	2.574356469
2012	TENNESSEE	0.931101804	1.6346335	3.376815777	2.072481702
2016	TENNESSEE	0.369419035	1.917402208	1.33841421	2.19189318
2020	TENNESSEE	-3.479070706	1.905751874	-6.182615032	2.000345851
2008	TEXAS	-4.819968209	4.012704945	-3.022147388	4.484296303
2012	TEXAS	1.703105662	1.62855732	4.32579924	0.801107165
2016	TEXAS	-1.579307198	0.758496566	-0.97883979	0.680253133
2020	TEXAS	-2.604532063	0.935581604	-5.420770131	1.318268617
2008	UTAH	-3.671936333	2.718903707	-2.848515233	3.23918833
2012	UTAH	-1.871885528	1.56090377	-0.276523046	1.56943184
2016	UTAH	1.718701894	1.592434706	3.112583105	1.736067768
2020	UTAH	0.114864487	1.876419924	-1.723081719	2.093139115
2008	VERMONT	3.698241513	2.022464705	4.196702371	2.488784324
2012	VERMONT	-1.517377941	1.395179118	1.228101088	1.761165674
2016	VERMONT	1.739476554	1.769459827	1.638524326	1.98808594
2020	VERMONT	-2.521431532	2.078733609	-5.412078482	2.174787392
2008	VIRGINIA	0.262091429	2.33329902	0.545068281	2.82731268
2012	VIRGINIA	-1.551903807	1.376638372	-0.88295745	1.711157236
2016	VIRGINIA	0.517830635	1.810090882	0.601623889	2.032821396
2020	VIRGINIA	-2.286494341	1.850016333	-4.166045224	1.959901913
2008	WASHINGTON	-1.218134704	2.247266494	0.808828159	2.796488716
2012	WASHINGTON	1.854335582	1.585383131	2.987612535	1.939158278
2016	WASHINGTON	1.742405512	1.422945627	2.415606388	1.650373314
2020	WASHINGTON	0.30374174	1.667700564	-2.141282306	1.730758144
2008	WEST VIRGINIA	1.997309874	4.172920592	2.254875705	4.532212425
2012	WEST VIRGINIA	-2.742102695	2.345988343	-0.134335533	2.383158544
2016	WEST VIRGINIA	0.14078773	0.364312017	-0.443665136	0.346102492
2020	WEST VIRGINIA	-1.083751877	1.538006362	-4.426879725	1.924972044
2008	WISCONSIN	-2.3756279	2.121639695	-2.332042851	2.578956126
2012	WISCONSIN	-0.206962307	1.624918363	1.747621709	2.080364345
2016	WISCONSIN	0.707471066	1.673499141	1.027313448	1.969696249
2020	WISCONSIN	-2.771189067	1.886215736	-5.480741786	1.973181935
2008	WYOMING	0.933803787	7.947696296	4.485230269	8.35898596
2012	WYOMING	-8.394270304	2.367084412	-5.096728042	0.030440814
2016	WYOMING	-6.879760796	1.452539384	-7.044164801	1.649904007
2020	WYOMING	-6.690981865	1.00470427	-8.56657977	1.727504161