The Effects of Political Parties and Demographics on Economic Status of States

Authors: Alyssa Allums and Bill Li

Research Questions

1. **What relationships exist between each state’s political leanings and their economic status (GDP, minimum wage, unemployment rates)? Does control of a state government by a certain political party help or hinder economic growth?**
   1. We are trying to determine if there is a correlation between the economic success of a state and its political leaders. These calculations would be critical in our ability to predict future economic success or failure.
   2. **Result:** There is no obvious trend for average GDP of states of different political parties, though Democratic states generally tend to reside at a higher average GDP than states of other parties. Republican states have the most consistent average GDP, with only a small amount of variation. For average unemployment rate, states of all political parties tend to follow the same trend, although states with the political party “other” were able to resist to some degree the increase in unemployment rate in the 21st century that states of other political parties underwent. Finally, states of all political parties have experienced an increase in minimum wage over the years, and lie relatively close to each other on the graph.
2. **To what degree do demographics of a state (urban vs rural, population) affect economic growth?**
   1. We are trying to determine if there is a relationship between the economic growth in each state and the demographics of the state. If there is a relationship, we can predict the economic growth of each state with changes in the demographics.
   2. **Result:** Average GDP and minimum wage in rural and urban states tends to increase throughout time with urban average GDP always higher than urban states. There isn’t any general trend in comparing demographics vs unemployment, but we can still see peaks after specific events. In small and medium populated states, GDP tends to increase with decrease in population. For large populated states, GDP tends to increase with an inclining population.
3. **What relationship exists between each state’s demographics and their political leaning?**
   1. We are trying to determine if there is a relationship between demographics and political leaning. If we can find a clear correlation, this could give insight as to why some states affiliate with certain political parties, and depending upon changes in the demographics, we could predict a change in the political majority in the state.
   2. **Result:** The majority of urban states and large population states tend to be Democratic while the majority of rural states and small and medium population states are Republican.
4. **Can political leaning and demographics be used to predict the economic status of a state?**
   1. We are trying to determine whether political leaning is an accurate means to predict economic status. If it is, this would be useful to predict the economic status of states in the future based on their political party. It would also help to see which states, based on their political leaning, will flourish or decline economically.
   2. **Result:** demographics cannot build an accurate model to predict economic status of a state

Motivation and Background

With the presidential elections coming up, we were interested in whether there is a relationship between political leaning and the economic status of each state and how the demographics in each state might affect the relationship.

This research can be used to determine which party is more effective in its economic policies which can be used to inform people about the economic situation after the presidential elections and allow them to make better decisions at the time of the elections.

Data

Political Indicator: <https://www.kaggle.com/kiwiphrases/partystrengthbystate>

This dataset is downloaded from Kaggle, which provides a nicely formatted table of the data scraped from the Wikipedia pages for each state’s political bodies. The dataset provides the number of people in each governing and representative body that affiliate with each political party. The data covers all states from the years 1980 to 2015.

GDP data by state: <https://www.bea.gov/data/gdp/gdp-state>

This dataset is downloaded from The Bureau of Economic Analysis, which is part of the U.S. Department of Commerce. The site provides various release tables regarding the gross domestic product (GDP) of states in the US. The tables of interest to us are under interactive data > interactive data: GDP by state > annual gross domestic product by state > Gross Domestic Product (GDP summary (SAGDP1) where it can be filtered to give percent change in GDP and total GDP for each state.

Urban Percentages: <https://www.icip.iastate.edu/tables/population/urban-pct-states>

This dataset is downloaded from Iowa State University, which compiled the data form the Decennial Census conducted by the US Census Bureau. The data provides the percent of the population within each state that lived in an urban area. The data accounts for each decade from 1970 to 2010

State populations: [https://fred.stlouisfed.org/release/tables?rid=118&eid=259194&od=#](https://fred.stlouisfed.org/release/tables?rid=118&eid=259194&od=)

This dataset is downloaded from the Economic Research site produced by the Federal Reserve Bank of St. Louis. The site provides the population for each state in thousands of persons over the course of many years. We compiled the data into one file to include all states over the years 1990 to 2019.

Minimum Wage: <https://www.kaggle.com/lislejoem/us-minimum-wage-by-state-from-1968-to-2017>

This dataset is downloaded from Kaggle, which provides a nicely formatted table of data scraped from the US Department of Labor website. The data provides a high and low value for the minimum wage of each state between 1968 and 2017

Unemployment Rates: <https://www.kaggle.com/jayrav13/unemployment-by-county-us>

This dataset it downloaded from Kaggle, which provides a nicely formatted table of data scraped from the United States Department of Labor's Bureau of Labor Statistics. It provides the unemployment rate for each state from the year 1990 to 2016, broken down into month and county.

Methodology

In order to categorize the political leaning of each state, we will look at the house of representatives, senate, and the electoral college columns in the political indicator dataset and compute which party has the cumulative majority in two of the three bodies. The political party of each body will be determined by the party which has the majority. If there are an equal number of representatives of different political parties in a body, the governing body will be labeled with “tie”. This applies to the overall state as well—if there is no political party with a majority in two of the three bodies, the state will be labeled as “tie”. The label “other” is given to states that have a political leaning toward a party that is neither Republican or Democrat. Finally, states for which there is no data for at least two of the three bodies will be null, as we cannot decide the political majority with data for only one body.

In order to categorize a state as urban or rural, we will look at the urban percentages dataset. We will consider a state to be urban if it is above the average percentage for all states in a given year. If it is under this average value, we will consider it to be rural.

In order to categorize states based on population, we will look at the state population dataset. We will have 3 categories for state size—small, medium, and large. A state is considered small if it is smaller than the mean of all state populations of the given year, it is medium if it is above the mean but under double the mean, and it is large if it is greater than double the mean.

(Q1) To compare political leaning to the economies of each state, we will plot the average GDP for Democratic states vs Republican states over a range of years. We will also do this for wage and unemployment, plotting the average values for Democratic and Republic states over the given years. If there are obvious differences in the trends between Democratic and Republican states, we can say that these factors correlate to political leaning. The plots will reveal whether one party on average has higher GDPs, wages, or unemployment rates.

(Q2) To determine the relationship between demographics and economy we will plot the average GDP, wage, and unemployment rate for states that are categorized as rural and urban over a range of years. If there are strong trends with obvious differences, we can say that the demographics (population and urban/rural) play a role in its economy. The plot will reveal whether rural or urban states have higher average GDPs, wages, and unemployment rates. We will make a separate plot for each category that will display the average population of each year on the x-axis and the average GDP of each year on the y-axis. The plot will give our client a strong visualization of the correlation between the economic status and the populations of the states. If there is a strong positive/negative correlation (linear relationship) then we can conclude an increasing/decreasing population will lead to economic growth/decline. For other shapes of the scatterplot, we will make different conclusions based on the correlation coefficient.

(Q3) To determine whether there is a relationship between demographics and political leaning, we will calculate the percentage of urban and rural states that are Republican or Democrat. We will also calculate the percentage of states within each population group (as described above) that are Republican and Democrat. These percentages, depending on which party has the majority in each category and if there is a large difference in the percentages, will tell us whether demographics is a factor in the political leaning of a state.

(Q4) To assess whether demographics can be used to accurately predict economic status, we will train a model on the demographics data (urban vs rural and population size) to train a model to predict economic status. We will then calculate the mean squared error for a test set of data. If the error is relatively small, then we can conclude that the model that we build is effective for making predictions about the future economic status of the states. Otherwise, we will conclude what type of information can lead to a large error in this model.

Results

**Questions 1:** What relationships exist between each state’s political leanings and their economic status (GDP, minimum wage, unemployment rates)? Does control of a state government by a certain political party help or hinder economic growth?

**A close up of text on a white background

Description automatically generated**To compare a state’s political leaning to their economic status, we first looked at average GDP over the years 1997 to 2018 for each category of political parties. There appears to be no obvious trend for individual parties, but the trend for Democratic states is generally set at a higher average GDP than the other parties over the course of years. From this data it seems that states that are Democratic also generally have a higher GDP. There is however a small period of time from about 2013 to 2015 when states with no clear majority (labeled as ‘tie’) rose above Democratic states in terms of their average GDP. The large fluctuations could be due either to a change in a states GDP or it may reflect the change of a state’s political party. For instance, if state S had a high GDP and was labeled as a Democratic state in 2013, if it changed to a tied state in 2014 this could cause the average Democratic GDP to drop and the average tied GDP to rise. Another feature to note is that the trend line for Republican states is more consistent that the other parties—it looks almost horizontal with very small fluctuations in a small range. This implies that rather than economic growth or decline, Republican states have stayed in a relatively similar economic state defined by their GDP.

**A close up of a map

Description automatically generated**Next we considered the average unemployment rate for states of different political majorities. The trends for each political party are roughly a similar pattern, revealing that there is no obvious correlation between political party of a state and its unemployment rate. It may be worth noting that in the later years of the graph, states with the political party ‘other’ are relatively lower than states with other political majorities, while those with majority ‘Democrat’, ‘Republican’, or ‘other’ are nearly in the same range of unemployment rates. This could be due to the fact that, in more recent years, states with a majority in a political party that is not Republican or Democrat tend to be states with a large agricultural and farming industry. It would seem these jobs rely slightly less on the economy of the nation, such that their unemployment rates can withstand a recession slightly better than states where most jobs are more reliant on the economy and market. Overall, it appears that since all the political parties relatively follow similar trends, the unemployment rate is most likely affected by trends on a more national level. For instance, we can see the unemployment rate skyrocket around 2007-2009, the time of the recession.

For the final aspect of comparing political leaning of a state to economic status, we looked at the average minimum wage for states of each political party over the years 1980 to 2015. Looking at the graph, we can see that all the lines follow almost the same trend, with a small amount of variation. The general trend is that the average unemployment rate has increased over the years. Toward the later years represented on the graph it seems Democratic states have slightly risen above states of other parties, while Republican states are almost consistently the states with the lowest average minimum wage. This could be due to the types of states the tend to have a majority in each party—i.e. agricultural vs metropolitan—or could simply be due to differing political views in states that are either conservative or liberal. However, the fact that we can see all the trends increase and that they stay in a similar range implies that minimum wage is most heavily affected by features on a national level. It makes sense that the minimum wage in states of all political parties has risen over the years due to inflation and the general increase in prices, which have resulted in a necessary increase in the minimum amount of money required to sustain one’s self.

**A close up of a map

Description automatically generated**

**Question 2:** To what degree do demographics of a state (urban vs rural, population) affect economic growth?

**A screenshot of a cell phone

Description automatically generated**

From this plot, we can see a linear relationship (excluding the decline in GDP due to stock market crash around 2008) between the year and average GDP for both urban and rural states of each year. The comparison between rural and urban are reasonable and predictable since urban states usually have more population than rural states which leads to more products and services.

A close up of a logo

Description automatically generated

Similar to the last plot, there is a linear relationship between Minimum wage and year for both rural and urban states. This positive relationship is reasonable since if the government decreases the minimum wage, then it would probably cause strikes among the labor forces so in order to make the laborers happy, the government needs to increase min wage corresponding to money inflation and other factors. For example, after the stock market crash of 2008, the average min wage increased by 0.5 which is the largest change on the plot, and this is due to huge inflation after the crash.

A close up of a logo

Description automatically generatedUnlike the other two plots, there is a nonlinear relationship between the year and unemployment rate with the majority of the urban states having a higher unemployment rate than rural states. Beside the difference in relationship, we can still spot similarities between the three plots such as the increase in unemployment rate after the stock market crash in 2008 compared to decline in average GDP and increase in Min wage and a decline in unemployment rate during the 1990s due to the 1990s United States boom.

A close up of a piece of paper

Description automatically generated

Surprisingly, the total population in low population states from 1997 to 2018 decreased with a nonlinear change in average GDP. After 2008, we can see the population and average GDP both started to decline due to the house market crash of 2008 and I think this trend is reasonable since it is harder for smaller states to recover from economic crisis. Before this crisis, we can spot an increase in both GDP and population which is contributed by the 1990s United States boom.

A close up of a map

Description automatically generated

As you follow through the red line I draw on the plot, we can see that the medium populated states also suffered from the 2008 market crash but the main difference is that the these states is able to recover faster than smaller states since we can already see some trend of increase in population and average GDP represented by the curl at the end of the line(around the larger dots).

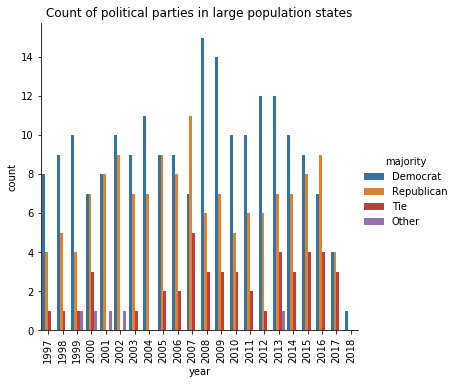
A close up of a piece of paper

Description automatically generatedUnlike the above plots, there was never a significant decline in population for large populated states. This trend can be explained by more job opportunities, better living environment in more populated states/cities that attracts many people from other states and countries which can also be a factor of declining population in small and medium populated states. The average GDP also suffered from the 2008 stock market crash just like the other states but this plot shows a staggering fast recovery from the recession. This can be explained by the fact that most government policies are more likely to favor large corporations (majority of them located in larger cities/states) so once those companies recover then their profit can trickle down to help the entire economy.

**Question 3:** What relationship exists between each state’s demographics and their political leaning?

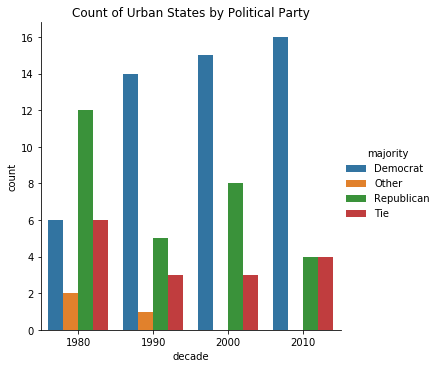
A picture containing writing implement, pencil, stationary

Description automatically generatedTo compare demographics to political leaning we first looked at the number of states in each political party in small, medium, and large population states (see methodology for description on how these populations are categorized). The graphs below show these counts from year 1997 to 2017. Comparing the graphs reveals that states categorized as low and medium population tend to have the majority of their states affiliated with the Republican party over the years. On the other hand, large population states tend to have the majority of states leaning toward the Democratic party. This reflects what we tend to see in the US, as states with large metropolitan areas that have large populations tend to be Democratic in a majority of elections (Washington, California, New York). That is not to say that there are no large states that are Republican—the graph shows there clearly are—but there seems to be a relationship between party and population. Based on real world knowledge of large states with metropolitan areas and the results of these graphs, I would theorize that metropolitan areas tend to be more liberal.

A picture containing pencil, writing implement, stationary

Description automatically generated

To continue comparing political party to demographics we plotted the number of states in rural vs. urban states that have a political leaning in each party. The results show that the majority of rural states tend to be Republican while urban states tend to be Democratic. This could be due to

the industries that dominate rural vs urban states. For instance, independent businesses and agricultural communities are probably less likely to want big government regulation policies (as typically endorsed by Democrats) while metropolitan areas that typically constitute an urban state are probably more likely to desire these policies so as to ensure a fair and competitive market. Another interpretation is that there is a link between metropolitan areas and liberal idea as explained in the previous section of question 3. States that are categorized as urban have a higher percentage of residents living in urban areas. It seems, based on our knowledge of the political climate in the US, cities tend to be more liberal (Democratic) than rural areas.

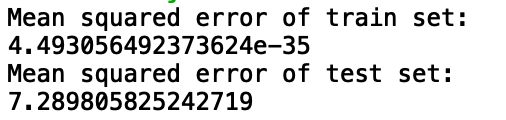
A screenshot of a cell phone

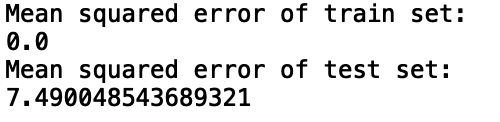
Description automatically generated

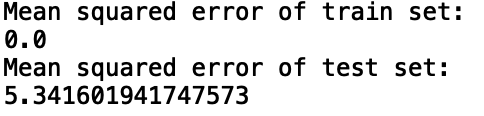
It is also worth noting that in the decade 1980 a majority of urban states were Republican while during the decades from 1990 to 2010 the majority for urban states is consistently Democrat. This could be due to the platforms of the parties at the time—i.e. the Republican party had a platform that would more so benefit the states—or it could be due to the shift of the Republican party toward a more conservative party. Throughout the 1900’s the Republican party began to ground its platform on more conservative ideas, similar to the ideas it advocates today. This change could have prompted urban states to transition to a more Democratic majority.

**Question 4:** Can political leaning and demographics be used to predict the economic status of a state?

From the model we built, we discovered that the mean squared error of the test set ranges from 5 to 8 which is considered to be really high since the data in percentage change in GDP typically range from 4 to 6. We concluded that this model can’t predict the future economic status of any state since the population and urban/rural of a state aren’t enough to determine the well-being in the economy. If we included more factors such as imports/exports, average incomes and etc then the model might be more accurate. Pictured below are different results when the program was run three separate times.







Reproducing the Results

The first step in reproducing the results is to download the data. Because there are so many datasets (the state populations data is only available as separate csv files for *each state,* meaning there are over 50 necessary csv files in all*)* we have provided a link to our GitHub repository, which will allow the user to clone the repository, providing all necessary csv files. Cloning the repository also provides the necessary python files. To do so, follow the link <https://github.com/libiyao/CSE-163-Project>. Once on the page, click the green “clone or download” button and copy the provided link to the clipboard. Next, navigate to your computer’s terminal and change your directory to where you wish to download the files. To change your directory, enter the command “cd \*directory path\*” where \*directory path\* is the path to the directory you wish to enter. For instance: ‘cd Documents/cse163’. When you have navigated to the directory, paste the link into the command ‘git clone \*link\*’. This will copy the repository into a folder in the current directory. Now the files can be run. To do so, change your directory once more so you are in the project folder (cd CSE-163-Project). Now simply enter the command ‘python CSE163\_Project.py’ to run the python file. This will save all plots you have seen above into the same folder as the CSE163\_Project.py. It will also print out two numbers separately stating the mean squared error of the trained set and mean squared error of the test set which are interpreted in the result and testing of question 4. To interpret the results the plots can be opened, compared, and analyzed. We simply compared the plots for each question and looked for trends.

Work Plan

**Work plan:** For the remaining three weeks, we will use two weeks to complete the coding part of the project and we will use the last week to complete the overall conclusion and write-up for this project. This weekend, we will meet together to process all the data we will need for each part. Then we will meet on specific dates to complete the coding portion for each question. We will also use GitHub to make it easier to share and access the code. After the completion of coding and testing, we will start to summarize our findings and work on the write-up.

We were able to roughly stick to the deadlines of the work plan. Processing the data took a bit longer than we thought as there were so many files that needed to be downloaded and cleaned. It was also slightly difficult to merge the data frames as the formats of the rows and columns varied among the data sets. This project was originally designed with three people but one of our group members dropped, so meeting up a couple times a week helped us manage the huge workload and collaborate. Overall, we were able to divide up the work and get the coding portion done in a couple of weeks, leaving us with about a week to do the write up.

Testing

Since the majority of this project’s results rely on plots, the only way to test it is to compare the plots to known data points. We also went through all of our conclusions and plots to see if it is reasonable and matched our expected output, otherwise, it probably means there is a bug in our code. For question 4, we calculated the mean squared error of the trained set and compared it to the mean squared error of the test set, testing it by comparing it to our expected results: that the MSE of the trained set should be around zero and smaller than the test set.

Collaboration

No one besides course staff and group mates helped us during this project.

Maybe people on stackoverflow ☺