**PSCI 250 Research Methods Final Paper**

**What is the effect of poverty levels on a presidential election?**

**Introduction**

In any given United States presidential election, voters are influenced by a host of factors when deciding for which candidate to cast their vote. Policy factors that influence voter decisions include foreign policy, economic policy, and social policy. Candidate personality factors may include charisma, toughness, professionalism, and control. On a more personal level, individuals are influenced by their own personal characteristics such as age, gender, race, region, and income level. This research project seeks to look into the effect of poverty in a county on presidential elections, specifically the 2016 election.

During the 2016 presidential campaigns, both major party candidates were questioned extensively on their economic policies. Republican candidate Donald Trump emphasized that his profound success as a businessman would give the economic policy knowledge necessary to benefit the national economy. Democratic candidate Hillary Clinton highlighted her years as the First Lady, a Senator, and the Secretary of State to show that she had plenty of experience working with policy of all kinds including economic policy. Because voters inherently take their personal economic and financial interests into consideration when going to the polls, each candidate’s economic policies were taken into consideration by voters in 2016. I was curious whether people in poverty would be influenced more by Trump’s businessman argument or Clinton’s political experience argument. This led me to the question, “*To what degree does a county’s percentage of people living in poverty affect which party will win the majority of the vote in that county in the 2016 presidential election?*”

In my research, people in poverty is defined as those individuals whose income is below the poverty threshold set by the United States Census Bureau. This threshold varies with the number of individuals that are in the household. I do hope that it will contribute to the already expansive literature on economic voting.

**Literature Review**

The relationship between household income and voting patterns in United States’ election, especially presidential elections, is a topic that has been researched extensively by many scholars in the post-World War II years. When one is seeking an answer to the question, “Why did people vote for a particular candidate in this particular election?” one is confronted by a variety of converging factors that influence the way an individual voted such as regionalism, ideological beliefs, party identification, and income. However, even though these other factors play a role, perhaps the most important cause of voting behavior is the voter’s personal household income.

Research in this field concentrates primarily on two main concepts: real personal income and income redistribution programs. Real income is income that has been adjusted for inflation, and income redistribution programs are government programs that move wealth from one segment of society to another, such as from rich to poor. Some researchers such as Clem Brooks, David Brady, and Douglas Hibbs believe that changes to real personal income do affect for which party a voter casts his or her vote. However, other researchers such as George Stigler, Christopher Faricy, and Christopher Ellis believe that voters’ differing views on upward or downward income redistribution programs affect their vote as well. Most researchers agree that both concepts affect voter choice, but debate arises over to what degree each concept affects voter choice and political party affiliation.

My review of previous studies begins with the groundbreaking work of George Stigler (1973) the recipient of the 1982 Nobel Prize for Economic Sciences. In his study, “General Economic Conditions and National Elections” he sought to test the commonly-held assumption that individuals’ personal economic experiences and situations influence their decision of which political party for which to vote. In his study, he used three main variables to gauge economic well-being: inflation rate, unemployment rate, and real personal income (Stigler, 1973). After putting these variables through a variety of complex mathematical equations, he concluded that these three variables are not the best indicators for how an individual will vote. Both political parties have such large coalitions comprised of many people from differing economic conditions. Therefore, it is in the interest of the party to benefit all of its constituents by keeping unemployment low so that people have incomes, and then keeping inflation low so that their real incomes do not decrease (Stigler, 1973). Even if one party does not seem to have accomplished this after a term, it is still in the best interest of the voter to vote for the party he perceives will have the best chance at increasing his economic well-being (Stigler, 1973). So, a person whose economic condition has declined slightly under a Democratic presidency will not immediately vote Republican in the next election because he still may believe that the Democrats are the best bet he has at increasing his economic standing. However, Stigler adds that the economic bases for a voter’s party affiliation in based most on income redistribution (Stigler, 1973) (Brooks, 1999). Stigler’s research does not come without limitations however. He does not address to what degree real personal incomes affects which party a voter chooses.

Although Stigler does not identify how personal income affects which party a voter chooses, a study by Clem Brooks and David Brady (1999) does identify which income levels correspond with electoral support for the Democratic or Republican Party. Furthermore, Brooks and Brady posit that voters will tend to vote for the opposite party as their economic situation changes, while Stigler stated that this is not, and should not be, the norm. In their study “Income, Economic Voting, and Long-Term Political Change in the U.S., 1952- 1996” they examined the degree to which household income affected election results and each party’s chance of winning a presidential election between the years of 1952 and 1996. They concluded that on an individual level, the estimated effect of household income is the same in all years with the exception of 1968 (Brooks, 1999). Voters who were one standard deviation poorer than average were 13% more likely to support a Democratic candidate than voters who were one standard deviation wealthier than average (Brooks, 1999). This conclusion is indicative of a “nontrivial impact on voting choice,” especially given the consideration that the model included controls for membership in social groups (Brooks, 1999). Therefore, the effect of income on voter choice were very stable in eleven of the twelve presidential elections from 1952-1996.

The first part of this study does not account for how changes in mean income affected outcomes in the postwar elections. In the next part of their study, Brooks and Brady found that shifts in mean income had a significant impact on elections, gifting the Republican Party with a growing electoral advantage (Brooks, 1999). For example, between 1952 and 1956, average household income increased by $5,000 (Brooks, 1999). The shift in income is predicted to have decreased the Democratic advantage by a significant 2% all by itself (Brooks, 1999). The cumulative shifts in income are more noticeable and impactful. For example, in 1972, NES data showed that voters’ average household income had increase by about $16,000 in a twenty year period. This alone created a 5% increase in the probability of a Republican candidate winning the election (Brooks, 1999). After 1972, voters’ average household income changed very little, leading to a relatively stable although higher level of Republican advantage. These results imply that if average household income increase in the future, the Republican advantage will increase even more (Brooks, 1999). This study is not without limitations, however. It does not address the idea that if economic circumstances are improving for people, they are more prone to reelect the party in power. This idea could conflict with the concept that the better people’s economic situations become, the more likely they are to vote Republican.

  A study by Douglas Hibbs (2000) agreed with the conclusions of the Brooks and Brady study, but it added a noteworthy corollary. In his study “Bread and Peace voting in U.S. presidential elections,” Hibbs concluded that the growth of “real disposable personal income per capita” during a presidential term accounted for 90% of the change in aggregate voting outcomes (Hibbs, 2000). By itself, real disposable income per capita changes best reflect the changing aggregate economic conditions of voters making it the best single-variable predictor of election results (Hibbs, 2000). This conclusion agrees completely with those reached by Brooks and Brady. However, Hibbs examines two anomalous years where Democrats perhaps should have won from a purely economic standpoint: 1952 and 1968. In these years, Hibbs asserts that United States’ intervention in the civil wars of Korea and Vietnam respectively robbed the Democratic Party of presidential wins they would have otherwise had (Hibbs, 2000). Surprisingly, Hibbs found that “no other variable appearing in the extensive literature on economic voting adds *anything* statistically to the explanation of aggregate presidential election outcomes when conditioned on weighted-average growth of per capita real disposable personal income and cumulative numbers of American military personnel killed-in-action in Korea and Vietnam” (Hibbs, 2000). This leads to voters punishing the party in power when they disapproved of the deaths being occurred in foreign conflicts. The limitation of this research is that is unclear from the results to what degree the “Peace” aspect of Bread and Peace affects the “Bread” or income aspect. Hibbs asserts that it does, but it is not clear to what degree.

In addition to finding that income levels affect voter choice, researchers have found that income in the form of income redistribution affects voter choice as well. Stigler concluded this in his study “General Economic Conditions and National Elections.” Stigler found that the economic bases for a voter’s party affiliation in based most on income redistribution (Stigler, 1973). Researchers Christopher Faricy and Christopher Ellis examine and revise this idea in their study “Public Attitudes toward Social Spending in the United States: The Differences between Direct Spending and Tax Expenditures” (2014). The study examines how individuals who identify with either the Democratic or Republican Party view three different types of social programs. The first is *mortgage interest tax deduction*,which allows individuals to deduct their mortgage interest payments from their federal taxes (Faricy, 2014). The second is *retirement savings tax deduction*,which allows individuals to contribute to employer-provided retirement plans before taxes (Faricy, 2014). The third program is *food stamps*,which provides assistance to needy families to purchase groceries and other necessities (Faricy, 2014). The first two programs are considered to be upward redistribution as they benefit the wealthy more than the poor. The third is considered to be downward redistribution as it benefits the poor more than the wealthy. The study concluded that Republicans do not consider the direction the program redistributes wealth when they are deciding whether to support or oppose a program (Faricy, 2014). Democrats, on the other hand, show significantly less support for the mortgage interest tax deduction and the retirement savings tax deduction when they are told that these programs benefit wealthier individuals the most (Faricy, 2014). Interestingly, Republicans are not less likely to support the food stamp program even though the program is obviously downward redistribution (Faricy, 2014). Furthermore, both Republicans and Democrats show greater support for the mortgage interest and retirement savings tax deductions when they are framed as tax credits and not direct payments (Faricy, 2014). I believe this study has a weakness because it did not test enough downward redistributive programs. If more downward redistributive programs had been included, I believe there may have been a difference between Democratic and Republican support. This would allow the parties to build coalitions around support for or resistance to income redistribution programs as they are doing today with issues such as the Affordable Care Act which downwardly redistributes income through healthcare subsidies.

Regarding party coalitions, political Scientist David Axelrod incorporated demographic information including income levels into his studies of electorate coalitions. However, he did not include information such as personal opinions on income redistribution programs. In his study “Presidential Election Coalitions in 1984” Axelrod identified the groups of people who are most likely to support each political party. Unsurprisingly, he concluded that the poor are more likely to support the Democratic Party; the poor were at least 25% more likely to vote Democratic than was the nation as a whole (Axelrod, 1986). This makes sense because the Democratic Party tends to support more downward redistributive social programs as the Faricy and Ellis study discovered. Additionally, the Democratic support of downward redistributive social programs seems to attract minority groups such as blacks (Axelrod, 1986). Unfortunately, this study lacks a comparison of certain variables together. For example, it does not specify whether the support of whites for the Republican Party decreases as the income of whites’ decreases. The addition of voters’ opinions on income redistribution programs also could have strengthened this study.

My research project will attempt to build upon the abundance of literature that has already been written on the topic of economic voting. My research question is “To what degree does a county’s percentage of people living in poverty affect which party will win the majority of the vote in that county in the 2016 presidential election?” I can assume that real personal income does influence a person’s voting behavior. Not only do people engage in voting behavior to “punish” the party in power during economic downturns, but they also vote for different parties based upon different income levels. These assumptions will help guide my research as I attempt to explain the effect income had on 2016 presidential election voting patterns. I will contribute to the ongoing conversation on economic voting by studying the relation between a county’s average household income and the county’s election results. My findings will help show to what degree household income was a reliable predictor of election results in the 2016 presidential election.

**Hypothesis**

*In a study of counties in the United States, those having a higher percentage of people in poverty will be more likely to have been won by Donald Trump in the 2016 presidential election than those that have a lower percentage of people living in poverty.*

**Methods**

For my research project, I had some difficulty finding data for the topic I wanted to research in my original conceptual research question: “How do political factors such as which party holds the governorship, the mayorship, and the majority in the city council, as well as policy factors, such as sales tax rates, property tax rates, income tax rates, and housing subsidies, between the years of 1955 and 2015 affect poverty rates in U.S. cities with populations greater than 600,000?” So, I changed the topic of my question, and I found two datasets that I will be able to utilize as I research my new operationalized research question: “To what degree does a county’s percentage of people living in poverty affect which party will win the majority of the vote in that county in the 2016 presidential election?” This question will show whether or not there is any correlation between poverty level and which political party wins an election. The coding scheme for this question will be interval because the categories will be numeric and I will know the exact numerical interval between each category.

The first dataset is from the United States Census Bureau, and it is titled “Small Area Income and Poverty Estimates.” The United States’ government is the entity that created this dataset. Beginning in 1964, during President Johnson’s “War on Poverty,” national poverty levels became a topic of interest for both researchers and policy-makers (Census Bureau, 2016). The next year, the Office of Economic Opportunity decided how it wanted to define poverty based on a methodology developed by Mollie Orshansky (Census Bureau, 2016). The Census Bureau was then charged with collecting this information, and in 1967 it released its first poverty estimates, which it has continued to release annually (Census Bureau, 2016).

The population under observation in this study is the whole United States. To study this population, the Census Bureau uses a set of income thresholds, expressed in dollar amounts, to figure out who is in poverty. These thresholds do not vary geographically across the United States, but they do vary based on the number of family members in the household, and the age and composition of these members. The thresholds are also adjusted for inflation using the Consumer Price Index. The Census Bureau then analyzes households by adding up the incomes of all related individuals in the household. If a family’s aggregate income level is below the threshold, then every single member of that family is considered to be in poverty (Census Bureau, 2016). This data is collected both through the census and through federal tax information.

This dataset will help me answer my new, operationalized research question by providing me with county-level poverty and income data. This will be an essential component of my research as I research how poverty levels in counties affect which political party won the county in the 2016 presidential election. Therefore, the most important variable for me will be the percentage of county population living in poverty. By comparing this to which party’s presidential candidate won the state, I will be able to see if there is any correlation between the two.

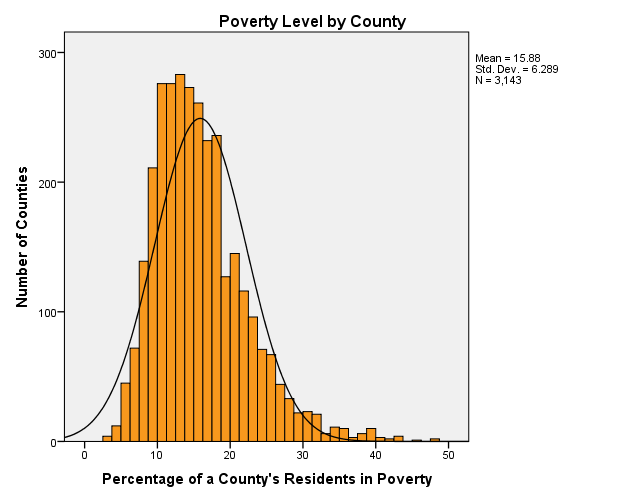
The second dataset that I will be analyzing is from data.world. This dataset provides county-level election results from the 2016 presidential election. This dataset was created by Gary Hoover and Ian Greenleigh, who both tracked and recorded the 2016 election results. They received most of their numbers from federal sources. This particular dataset does not have a long history as it was created immediately following the 2016 election, but data.world contributors have tracked and compiled election results for several years. This dataset would not be possible without both of the authors, the federal sources, or data.world which allows contributors to create datasets.

The population under survey in this dataset includes those individuals who voted in the 2016 presidential election. The number of individuals in this dataset totals 129,944,669 (Hoover, 2016). The dataset includes all of the counties in the United States, the candidate who won the county, the number of votes each candidate won in the county, the population of the county, and the total number of votes cast in each county (Hoover, 2016). The data was collected by the individuals who compiled the dataset using the federal sources. None of the variables are weighted.

This dataset will help me research my question by providing information about vote totals for each county in the United States. Therefore, the variables I am interested in are the number of votes cast for either Clinton or Trump in a given county. Then, I can compare the percentage of the vote each candidate received in a given county to the percentage of people living in poverty in that same county.

**Methods and Preliminary Results**

Using the combined data set that I created, I was able to study two variables. My independent variable was *What is the percentage of the county’s residents living in poverty?* (*Poverty\_Level*). I created this variable by recoding the variable *AllAgesinPovertyPercent.* First, I ran the frequencies of *AllAgesinPovertyPercent* which yielded the following results:



|  |  |  |
| --- | --- | --- |
| **Statistics** | | |
| All Ages in Poverty Percent | | |
| N | Valid | 3143 |
| Missing | 1 |
| Mean | | 15.88 |
| Median | | 14.90 |
| Mode | | 13 |

I recoded the nominal level percentages into the ordinal variable, *Poverty\_Level*, with the categories Low, Medium, and High (Low = 1/Medium = 2/High = 3). Medium consisted of all cases ± 1 standard deviation from the mean. Low consisted of all cases 2 or more standard deviations below the mean, while High was composed of all cases 2 or more standard deviations above the mean.

My dependent variable was *Who was the winner of the 2016 presidential election in the county? (CountyWinner)*. I created this variable by coding those counties with a higher vote total for Trump as “1” and those having a higher vote total for Clinton as “2.” Running frequencies on *CountyWinner* yielded the following results:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **County Winner** | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid |  | 2 | .1 | .1 | .1 |
| Clinton | 481 | 15.3 | 15.3 | 15.4 |
| No data | 29 | .9 | .9 | 16.3 |
| Trump | 2632 | 83.7 | 83.7 | 100.0 |
| Total | 3144 | 100.0 | 100.0 |  |

|  |
| --- |
|  |
|  |

I did not have many major challenges with my data. One minor challenge was the negative skew present in my dependent variable. This was due to the very small number of counties that had an unusually large number of residents living in poverty. Another challenge was that there was not already a variable present for which candidate won the county. The dataset had the vote totals each candidate received in each county, but it did not have a variable of just the winner’s name. Therefore, I went through the dataset and compared the vote totals for Clinton and Trump and created the new variable *CountyWinner.* I typed the name of the candidate who had the higher vote total in each county into a new column. This process was painstaking and tedious, but it was not a major challenge.

My preliminary results are as follows:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **County Winner \* Poverty Level Crosstabulation** | | | | | | |
|  | | | Poverty Level | | | Total |
| Low | Middle | High |
| County Winner | Clinton |  | 45 | 274 | 162 | 481 |
|  | 23.6% | 11.5% | 28.2% | 15.3% |
| No data |  | 7 | 17 | 5 | 29 |
|  | 3.7% | 0.7% | 0.9% | 0.9% |
| Trump |  | 139 | 2085 | 408 | 2632 |
|  | 72.8% | 87.7% | 71.0% | 83.7% |
| Total | |  | 191 | 2377 | 575 | 3143 |
|  | 100.0% | 100.0% | 100.0% | 100.0% |

As you can see, Trump performed much better than Clinton in counties of all poverty levels. Trump performed best in the Middle-income counties, while Clinton performed best in High-income counties. Trump performed worst in Low-income counties, but he still won more than triple the number of Low-income counties as Clinton did. Across the board, Trump won massive percentages of the counties at every poverty level. In order to determine whether or not the results were statistically significant, I made SPSS generate a Chi-Squared analysis of the *CountyWinner* and *Poverty\_Level* cross tabulation. The Chi-Square test produced a P value of 0.00 which indicates that the results of the CountyWinner and Poverty\_Level cross tabulation are significant. The Chi-Square value was 128.361a, while p=0.00.

**Discussion**

Following the statistical analysis, one could say that the original hypothesis “In a study of counties in the United States, those having a higher percentage of people in poverty will be more likely to have been won by Donald Trump in the 2016 presidential election than those that have a lower percentage of people living in poverty” was not supported. Those counties two or more standard deviations below the mean were 1.8 percentage points more likely to vote for Donald Trump than were counties two or more standard deviations above the mean. Also, the counties in the middle poverty group were most likely to vote for Trump at 87.7%.

It is unclear which factors influenced counties of all poverty levels to overwhelmingly vote for Trump. Was it his perceived success as a businessman? Was it the tax plan that he was proposing? Was it his promises to bring back jobs to America? Or was it something else such as Hillary Clinton’s high unfavorability ratings or his stance on ISIS? Whatever the case, it is clear that Trump had a broad appeal to counties of all poverty levels.

**Conclusion**

Although there was not a clear correlation between poverty level in a county and wins for Trump, it was clear that Trump won a massive majority of all counties across the board regardless of the poverty level. Future studies could focus more on the poverty aspect of the election or they could take into account other variables such as region, age, gender, and race. Seeing as how studies have already shown that Clinton was favored in metropolitan areas, it would be interesting to research how the number of counties that voted for Trump and the number of counties that voted for Clinton compare when counties over a certain population number are removed from the study. Additionally, future research should focus on how the two different candidates’ views on redistributive economic policies impacted people living in poverty. Hopefully, the research of this study on the relationship between poverty levels and county winner will add to the overall discussion on economic voting.