Income Inequality and 2016 Voting Trends

Gabriel Wachowski, Matt Dahle, Mitch Rosiejka, Nikki DeCleene

What we did with our feedback:

We researched the gini coefficient for income inequality. However, this data was not offered at a county-level. Therefore, we accepted this as a limitation of our project and continued to used the ratio of income at the 80th percentile to the 20th percentile as a measure of income inequality.

We included median income as a confounding variable in our design.

We researched a possible metric for major college campuses in counties. Unfortunately, because this data was not available, and limitations such as college students not being registered to vote where they go to school and choosing to send votes in from their home county instead of the county where their college is located, we could not include major college campuses in counties as a confounder.

We adjusted the figure labels to make them larger.

Introduction

Many factors affect the political party voting trends of counties in the United States. One of these factors could be income inequality, which has been increasing in America since the 1960s (Faricy 2016). Income inequality is influenced by economic policies, which are very different for Democratic and Republican parties. Research has even found that "Republican Party control results in economic policies that increase income inequality while Democratic rule assuages inequality" (Faricy 2016). Therefore, it is likely that voters may try to influence economic policies (specifically, change income distribution) by voting for one political party over another. It seems plausible that people living in areas with high income inequality may be more likely to vote Democratic, since Democratic party control would mean economic policies that would lessen income inequality, which would be a key issue for these voters. Controlling for population density, education, race, income and state effects, we found **the more unequal the income distribution of a county, the more the county tended to vote democrat in the 2016 U.S. presidential election**.

It would be useful to examine the relationship between income inequality and voting patterns because understanding this trend could affect the ways politicians promote their campaigns in future elections. In the future, Republicans may want to target areas with high income inequality where fewer people voted for the Republican party in order to change this trend by promoting more policies to change income distribution. Democrats could target these areas and further emphasize their mission to limit income inequality in order to keep up this trend and secure the county.

Possible confounders that we examined were population density, race, education, income, and state effects. Population density could affect political trends because there is a relationship between urban versus rural areas and who they tend to vote for. Urban areas tend to be more democratic and rural tend to be more republican (Gamio 2016). Education also has been found to affect how people vote. The more educated a person is, the more they tend to vote democratic and vice versa, especially in the 2016 election (Suls 2016). State effects are also very important to control for, since many states have different voting methods and factors such as whether the state is a “swing state” may affect how people are likely to vote -- or if they vote at all. Race can also affect voting patterns. Race has long been a prevalent part of American politics, but it was especially important in the 2016 election, since Donald Trump’s “rhetoric around race, ethnicity, and nationality” made winning over minority voters difficult (Henderson 2016). Therefore, it is important to include race as a confounding variable. It is also important to consider median income itself when examining income inequality, since counties where all residents are wealthy and counties where all residents are poor would both have low income inequalities, but would most likely vote very differently, since income affects voting trends (Thompson 2012).

Methods

http://www.countyhealthrankings.org/

http://www.countyhealthrankings.org/explore-health-rankings/our-methods/data-sources-and-measures

Information regarding income inequality was obtained from County Health Rankings & Roadmaps, which was compiled with Robert Wood Johnson Foundation and the University of Wisconsin Public Health Institute. The measure of income inequality we used was the ratio of household income at the 80th percentile to income at the 20th percentile. Although this number was included in the data set, it was listed as a rounded value. Therefore, we re-calculated the ratio ourselves in order to make income inequality a continuous variable.

This data set was also used to obtain information about race. We used the percentage of African American, Hispanic, Asian, American-Indian/Alaskan Native, Native Hawaiian or other Pacific Islander, and Non-hispanic white in a county. These were used as a confounding variable in order to account for the fact that race affects voting trends and income disparities.

Another confounding variable that we used from this data set was education. We created a variable of how many people in a county had some college education by dividing the number of adults age 25-44 with some post-secondary education by the number of adults age 25-44. This was used as a confounding variable to account for the fact the amount of education affects voting and income inequality.

This data set also included median income for each county, which was another confounding variable that we accounted for in our linear model.

https://github.com/tonmcg/County_Level_Election_Results_12-16/blob/master/2016_US_County_Level_Presidential_Results.csv

2016 election results were obtained from GitHub, which compiled election results published on Townhall.com: https://townhall.com/election/. This dataset listed the number of votes for the Democratic Party and the number of votes for the Republican Party. We created another variable to measure the percentage that voted Democratic by dividing the number of votes for the Democratic Party by the total number of votes for the Democratic Party added to the total number of votes for the Republican Party, in order to exclude votes for Third Party candidates.

https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

We used population density data as a confounding variable. This data was obtained from the U.S. census. The variable calculated population density by dividing total population by the land area in the county. We included this data in our regression to account for the fact that large cities usually have more income disparities and are usually more liberal voting wise. Using population density allows us to account for big cities compared to rural areas on the county level.

Data sets for containing the ratio of income inequality, percentage of votes for the Democratic Party, percentage of non-white hispanics, percentage of adults with some college education, median income, and population density were combined into an excel file. Counties with any missing data were erased from the data set.

We ran a linear regression on our income variable and democratic voting ration, as well as all of confounding variables. After noting a non-random pattern on our residual map, we added a dummy variable to account for a state effect. We used an ANOVA test to decide the validity of including state effects as a confounding variable. This linear regression model allows us to determine the significance of our variables as a way to predict the ratio of democratic votes of a county from income inequality.

Results

We examined the relationship between income inequality and proportion of democratic votes (Fig. 1). We found income inequality was significant, as evident by a low p-value, and almost every state was significant, when compared to Alabama as baseline (Fig. 2). The confounding variables (race, education, population density, and income inequality) were significant as well (Fig. 2). Adding confounding variables decreased the residuals, however are still non-random patterns in our residual map (Fig. 3). Using ANOVA, we compared our linear model with state effects to a model without state effects. The results showed a low p-value, meaning that the state effect is significant (Fig. 4). This shows that the state effects model should be used along with other confounders. In conclusion, our results showed that there is a positive relationship between income inequality and democratic voting in the 2016 election.

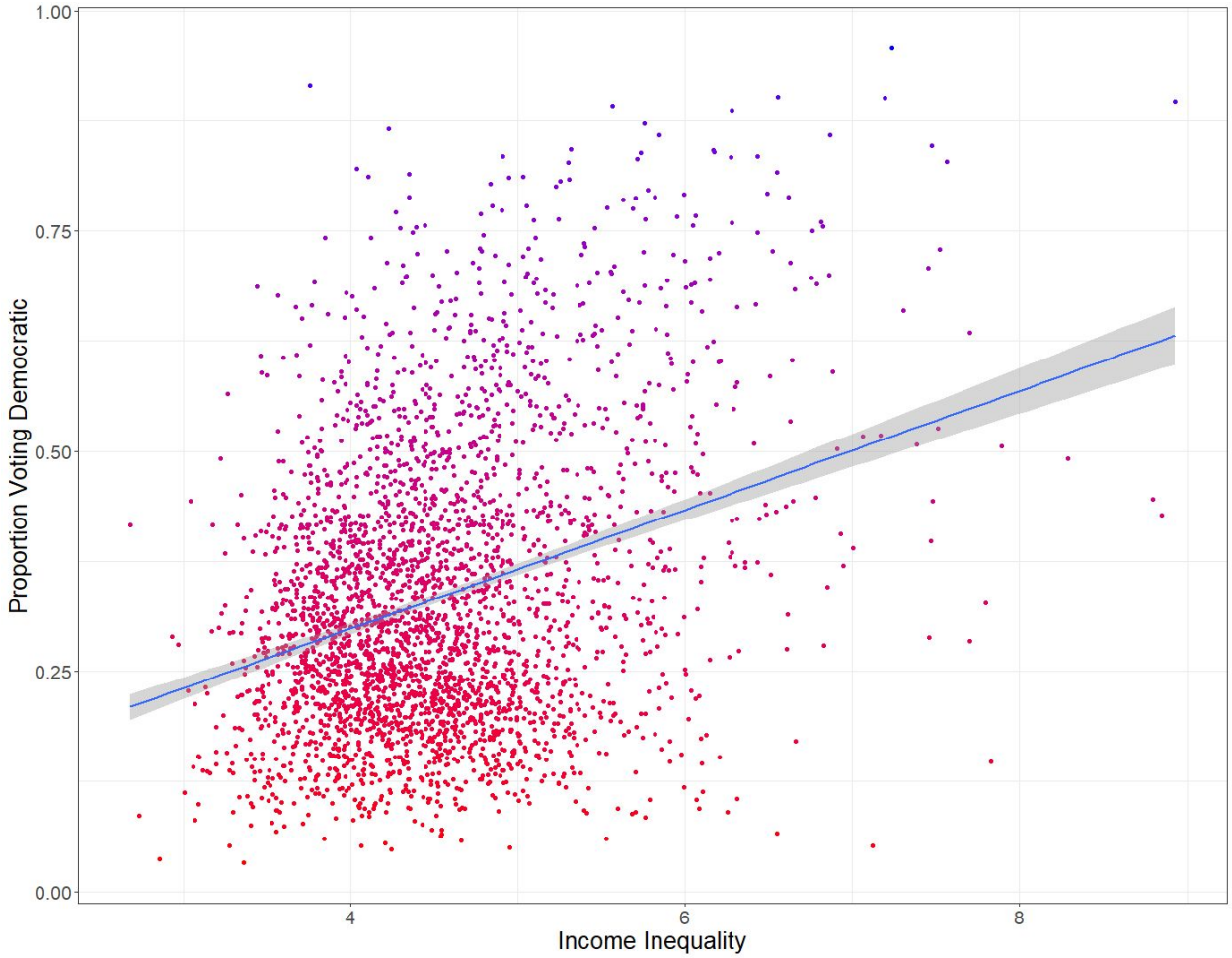

Fig. 1. Income inequality vs proportion voting democratic with simple linear regression line

```
Residuals:
     Min       1Q   Median       3Q      Max
-0.28766 -0.04752 -0.00308  0.04179  0.36357

Coefficients:
                                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                              1.551e+00  1.851e-01   8.378  < 2e-16 ***
Income.Ratio.2                           2.837e-02  2.617e-03  10.839  < 2e-16 ***
Pop.Density.per.sq.mile                  1.476e-06  8.912e-07   1.656 0.097898 .
Percent_college                          3.408e-01  1.877e-02  18.158  < 2e-16 ***
StateAlaska                             -1.639e-03  2.145e-02  -0.076 0.939085
StateArizona                             1.104e-01  2.278e-02   4.843 1.34e-06 ***
StateArkansas                            8.085e-02  1.298e-02   6.230 5.30e-10 ***
StateCalifornia                          1.496e-01  1.530e-02   9.779  < 2e-16 ***
StateColorado                            1.567e-01  1.425e-02  11.000  < 2e-16 ***
StateConnecticut                         2.220e-01  2.897e-02   7.665 2.37e-14 ***
StateDelaware                            1.574e-01  4.509e-02   3.490 0.000489 ***
StateDistrict of Columbia                2.017e-01  7.742e-02   2.605 0.009245 **
StateFlorida                             7.974e-02  1.349e-02   5.911 3.77e-09 ***
StateGeorgia                            -2.139e-03  1.114e-02  -0.192 0.847788
StateHawaii                             -7.166e-02  9.825e-02  -0.729 0.465804
StateIdaho                               7.128e-02  1.546e-02   4.611 4.18e-06 ***
StateIllinois                            1.456e-01  1.255e-02  11.598  < 2e-16 ***
StateIndiana                             1.618e-01  1.291e-02  12.535  < 2e-16 ***
StateIowa                                1.974e-01  1.291e-02  15.284  < 2e-16 ***
StateKansas                              2.734e-02  1.284e-02   2.130 0.033259 *
StateKentucky                            1.138e-01  1.217e-02   9.352  < 2e-16 ***
StateLouisiana                          -1.625e-02  1.342e-02  -1.211 0.225941
StateMaine                               3.460e-01  2.161e-02  16.007  < 2e-16 ***
StateMaryland                            1.039e-01  1.858e-02   5.593 2.43e-08 ***
StateMassachusetts                       3.469e-01  2.293e-02  15.127  < 2e-16 ***
StateMichigan                            2.137e-01  1.317e-02  16.231  < 2e-16 ***
StateMinnesota                           1.879e-01  1.323e-02  14.209  < 2e-16 ***
StateMississippi                        -3.002e-03  1.276e-02  -0.235 0.814064
StateMissouri                            1.085e-01  1.232e-02   8.812  < 2e-16 ***
StateMontana                             1.119e-01  1.462e-02   7.657 2.53e-14 ***
StateNebraska                            2.510e-02  1.328e-02   1.889 0.058921 .
StateNevada                              4.564e-02  2.180e-02   2.094 0.036385 *
StateNew Hampshire                       3.379e-01  2.624e-02  12.876  < 2e-16 ***
StateNew Jersey                          1.338e-01  1.991e-02   6.721 2.14e-11 ***
StateNew Mexico                          7.244e-02  1.813e-02   3.995 6.61e-05 ***
StateNew York                            2.078e-01  1.407e-02  14.768  < 2e-16 ***
StateNorth Carolina                      9.214e-02  1.219e-02   7.556 5.45e-14 ***
StateNorth Dakota                        5.306e-02  1.486e-02   3.572 0.000360 ***
StateOhio                                1.696e-01  1.290e-02  13.146  < 2e-16 ***
StateOklahoma                           -5.789e-02  1.478e-02  -3.916 9.20e-05 ***
StateOregon                              1.808e-01  1.649e-02  10.959  < 2e-16 ***
StatePennsylvania                        1.869e-01  1.368e-02  13.666  < 2e-16 ***
StateRhode Island                        3.071e-01  3.567e-02   8.608  < 2e-16 ***
StateSouth Carolina                      5.216e-02  1.468e-02   3.553 0.000386 ***
StateSouth Dakota                        1.009e-01  1.433e-02   7.040 2.37e-12 ***
StateTennessee                           1.039e-01  1.259e-02   8.254 2.24e-16 ***
StateTexas                              -6.288e-02  1.188e-02  -5.292 1.29e-07 ***
StateUtah                                5.270e-02  1.777e-02   2.967 0.003033 **
StateVermont                             4.663e-01  2.280e-02  20.456  < 2e-16 ***
StateVirginia                            1.158e-01  1.171e-02   9.890  < 2e-16 ***
StateWashington                          1.922e-01  1.618e-02  11.877  < 2e-16 ***
StateWest Virginia                       1.447e-01  1.441e-02  10.038  < 2e-16 ***
StateWisconsin                           2.672e-01  1.371e-02  19.491  < 2e-16 ***
StateWyoming                             1.870e-02  1.898e-02   0.985 0.324586
Household.Income                         7.722e-07  1.795e-07   4.302 1.75e-05 ***
African.American                        -1.040e+00  1.859e-01  -5.593 2.42e-08 ***
Hispanic                                -1.263e+00  1.790e-01  -7.056 2.10e-12 ***
Non.Hispanic.White                      -1.845e+00  1.848e-01  -9.981  < 2e-16 ***
Asian                                   -1.067e+00  2.010e-01  -5.309 1.18e-07 ***
Native.Hawaiian.Other.Pacific.Islander  -1.745e+00  8.745e-01  -1.996 0.046036 *
American.Indian.Alaskan.Native          -1.394e+00  1.870e-01  -7.455 1.16e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.07619 on 3077 degrees of freedom
Multiple R-squared:  0.7785,    Adjusted R-squared:  0.7741
F-statistic: 180.2 on 60 and 3077 DF,  p-value: < 2.2e-16
```

Fig. 2. Linear model summary examining the relationship between proportion voting democratic and income inequality, controlling for race, education, median income, population density, and state effects

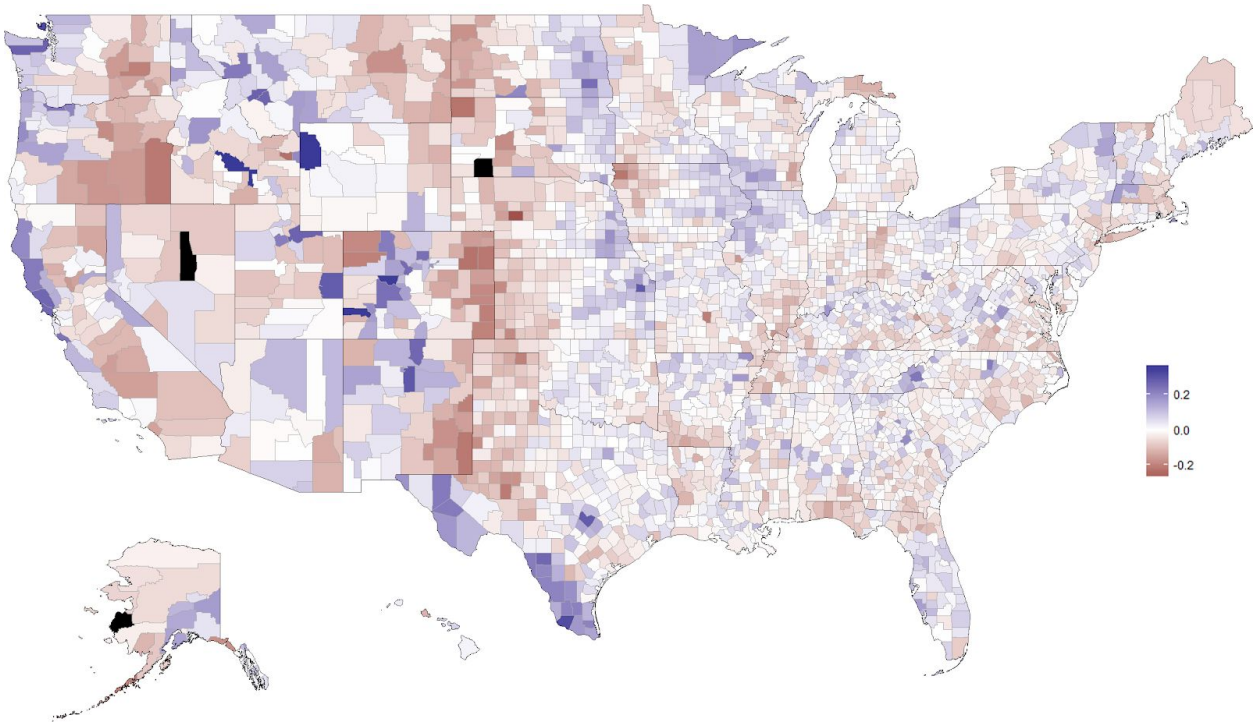

Fig. 3. Map of residuals from linear model with the confounders of race, education, median income, population density, and state effects

```
Analysis of Variance Table

Model 1: ((votes_dem)/(votes_gop_dem)) ~ Income.Ratio.2 + Pop.Density.per.sq.mile +
    Percent_college + State + Household.Income + African.American +
    Hispanic + Non.Hispanic.White + Asian + Native.Hawaiian.Other.Pacific.Islander +
    American.Indian.Alaskan.Native
Model 2: ((votes_dem)/(votes_gop_dem)) ~ Income.Ratio.2 + Pop.Density.per.sq.mile +
    Percent_college + Household.Income + African.American + Hispanic +
    Non.Hispanic.White + Asian + Native.Hawaiian.Other.Pacific.Islander +
    American.Indian.Alaskan.Native
  Res.Df    RSS  Df Sum of Sq      F    Pr(>F)
1   3077 17.862
2   3127 37.757 -50   -19.895 68.543 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Fig. 4. ANOVA results comparing linear models with and without state effects

Conclusion

Returning to our thesis statement, we see that our prediction holds true when looking at our results, since income inequality was a significant predictor for the ratio of democratic votes in a county when controlling for education, race, median income, population density, and state effects. The linear model shows significance for almost every factors, as seen by the low p-values, including the income ratio, which is our main focus. Our p-value for the model is significant as well as our ANOVA for the state effects model. Our coefficient for income ratio is positive suggesting that as the income ratio increases, the proportion of democratic votes in a county also increases.

The causal story of this project lies in the stereotypical economic policies of the two main political parties. The Republicans are known for cutting taxes for the rich and also putting a focus on the middle class. Democrats typically focus on economic reforms to help those in poverty. These preconceived notions are supported by research into economic policies of both parties, which find that under Republican Party control, economic policies increase income inequality, while Democratic economic policies reduce income inequality (Faricy 2016). Thus, if someone is in poverty but sees a lot of wealthy people in their area, you would think that they would vote for candidates that focus on helping them by reducing income inequality when they get into office. This idea is supported by our data, in how the counties with highly unequal incomes tended to vote Democrat in the 2016 election.

The limitations of our study are that we only looked at one election, so looking at multiple elections could give a better idea of historical trends. Furthermore, there were

some trends in our residual map in the Great Plains area that is left unexplained, even after adding confounding variables to our model. We found that many college towns were outliers as well, we were unable to control for this factor because of difficulties finding data on the subject. In conclusion, we see that income inequality and the voting patterns of 2016 are significantly correlated.

Works Cited

Faricy, Chris. 21 January 2016. “The Politics of Income Inequality in the United States.” *Political Science*. http://www.oxfordbibliographies.com/view/document/obo-9780199756223/obo-9780199756223-0147.xml

Gamino, Lazaro. 17 November 2016. “Urban and Rural America are becoming increasingly polarized.” *The Washington Post.* https://www.washingtonpost.com/graphics/politics/2016-election/urban-rural-vote-swing/

Henderson, Nia-Malika. 1 September 2016. “Race and Racism in the 2016 Campaign”. *CNN News.* https://www.cnn.com/2016/08/31/politics/2016-election-donald-trump-hillary-clinton-race/index.html

Suls, Rob. 15 September 2016. “Educational divide in vote preferences on track to be wider than in recent elections.” *Pew Research Center.* http://www.pewresearch.org/fact-tank/2016/09/15/educational-divide-in-vote-preferences-on-track-to-be-wider-than-in-recent-elections/

Thompson, Derek. 5 November 2012. “Does Your Wage Predict Your Vote?” *The Atlantic*. https://www.theatlantic.com/business/archive/2012/11/does-your-wage-predict-your-vote/264541/