Inequality, voter turnout, and congressional net worth

By Max Eber
Advised by Michael Rizzo

Abstract:

Economics and politics are inherently related, given that the decisions of policymakers have a direct impact on economic outcomes. With the trend of rising economic inequality in the United States, how do politicians influence the perpetuation or reduction of such a divide? The relationship between income inequality, voter turnout and income bias in voter turnout are empirically investigated utilizing voter data distinct from the sources used in most similar analyses. Subsequently, the relationship between income inequality and the net worth of U.S. Representatives is evaluated in order to test the plausibility of a positive feedback loop between economic inequality and unequal political representation. Income inequality is seen to be negatively associated with overall voter turnout, with the wealthiest individuals of a district being the most likely to vote. These findings are subject to significant concerns over lack of robustness. A positive relationship between income inequality and net worth of representatives relative to the median income value of their district is found when testing in each direction and when incorporating lag. This result indicates a quantitatively supported positive feedback loop.

Section 1: Introduction

Income inequality continues to rise in the United States, as it does in much of the world. While some individuals are not concerned about this trend, it has the capacity to affect almost all aspects of society, from nationwide economic growth to education and healthcare outcomes. Policy is arguably the most macro method of addressing and affecting distribution of income, so it is valuable to consider the political landscape when evaluating the impact of income inequality. If economic inequality leads to unequal political representation, it can create a positive feedback loop which could further perpetuate an economic divide. In a representative democracy like the United States, political inequality initially stems from the voters. The existence of forces such as special interest groups and lobbying can certainly influence political decision making after officials are elected. However, voting foundationally determines who represents the constituents of each district. Therefore, the active voter population holds substantial power in contributing to the political climate via the selection of elected officials.

Likelihood of voting is a measure of a vote's perceived value. For instance, if one forgoes voting, it is implied that their time is valued higher than the vote itself. Several questions must be considered in analyzing this concept. As income inequality rises, how does the perceived value of a vote fluctuate? Furthermore, does this perception fluctuate unevenly among the income distribution? Observing evidence of income bias among voters would imply that wealthy supported candidates are likely overrepresented in politics and poor supported candidates are likely underrepresented.

It may be beneficial to investigate the profiles of elected representatives in areas of higher inequality. For instance, are these officials most representative of the median voter or either extreme? Furthermore, if income bias is identified in the voters, and if individuals vote in their own self-interest, would wealthy candidates be largely supported by the affluent? If there is in fact a positive feedback loop between income inequality and unequal political representation, this would perpetuate increasing income inequality nationwide. Observing the validity of this relationship, as well as understanding particular mechanisms within the relationship, would provide insight into the context of current trends.

As such, this paper analyzes how income inequality affects voter turnout and income bias in voter turnout. Furthermore, the relationship between income inequality and wealth of representatives is explored empirically. To investigate the second half of this possible feedback loop, this paper measures whether wealthier representatives are associated with districts of higher inequality. Lastly, this paper explores, through existing literature, variables impacting likelihood of voting, whether people vote in their own self-interest, and how representative politicians are of their constituents. The focuses of this paper provides empirical insight into the dynamics of income inequality and political inequality, illuminating specific mechanisms of that relationship.

Section 2: Literature review

2.1 - Introduction to income inequality

Inequality in general is the primary focus of much existing literature. Researchers have extensively investigated the causes of, impacts of, and potential solutions to different forms of inequality. To provide sufficient context for this paper, it is useful to review the implications of income inequality outside those pursued by my own investigation.

Inequality is frequently discussed as it relates to economic growth and individual wealth. Kuznets (1955) hypothesized the Kuznets curve, which postulates an inverted U-shaped relationship between income per capita and inequality. Namely, once people become wealthy enough, the economy will reach a turning point and inequality will begin to decline. Though many argue against the legitimacy of this theory by claiming it has a pick-and-choose nature to its case studies (Deininger & Squire, 1998) and empirical evidence to the contrary (Fields, 2001), it still provides an early example of inquiry into this dynamic relationship.

There is not a consensus among researchers as to the exact relationship between income inequality and economic growth. The existing literature's focus differs based on whether a country is developed or developing, with developing countries subsequently segmented between high-income and low-income.Utilizing panel data, some research indicates that in the short and medium run, there is a positive relationship between increasing income inequality and economic growth (Forbes, 2000). Other papers present rather contradictory results. For example, in high income countries this same result is achieved, but low income countries were found to have a statistically significant negative relationship between income inequality and economic growth (Barro, 2000; Fawaz, Rahnama, & Valcarcel, 2014).

Income inequality has been associated with an array of other outcomes. Some of these variables are increased mortality rates, residential concentration of wealth, increased crime rates, as well as reduced educational success (Wilkinson & Pickett, 2009; Kawachi & Kennedy, 1997). Without delving into detail on the specific outcome metrics, it can be observed that a variety of undesirable results have been shown to derive from higher rates of income inequality.

2.2 - Likelihood of voting

As this paper seeks to observe rates of bias in voter turnout, it is valuable to have a background concerning what existing studies have seen to affect likelihood of voting. The variables that may or may not affect voter turnout can be divided into two subdivisions: institutional factors and socioeconomic factors. A recent meta-analysis of voter turnout research by Daniel Stockemer (2016) challenged the conclusivity of previous findings. His analysis of 135 articles finds the institutional factor of compulsory voting, the importance of an election, and the size of a country to be the most impactful variables in increasing voter turnout (Stockemer, 2016). The conflict with existing research arises over the influence of other possible predictor variables. Stockemer (2016) claims these can be inconclusive at best, whereas previous studies made more resolute claims.

Some institutional factors typically investigated are compulsory voting, type of electoral system, type of legislative body, importance of election, number of parties, voting age, and registration laws. As one might expect, compulsory voting is found to have a significant positive effect on voter turnout. While this is less relevant to the focus of this paper as the United States does not mandate voting, it is useful in providing an example of an institutional factor that increases voter participation rates.

The type of electoral system is constant in the United States so variation in voter turnout between systems does not directly impact the inquiry posed by this paper. Nevertheless, understanding the impact of a given electoral system allows for more realistic baseline expectations. Further, much of the existing literature addresses the relationship with proportional representation (PR), a system which the U.S. House of Representatives, whose elections are the focus of this paper, subscribes to. While a variety of more dated papers indicate a positive relationship between PR and voter turnout (Blais and Carty, 1990; Jackman & Miller, 1995; Radcliff & Davis, 2000), a large portion of the more recent literature do not find PR to have an influence on macro-level voter turnout (Brockington, 2004; Lehoucq & Wall, 2004). These discrepancies seem to indicate that the magnitude of impact of PR, or the electoral system in general, has reduced over time (Stockemer, 2016).

The type of political system is relevant in that it varies the significance of a given election. The political system outlines the level of power and influence of a particular legislative body and subsequently, how potential voters value their participation in the respective election. Evidently elections deemed more important have been found to produce higher rates of voter turnout (Johnston, Matthews, & Bittner, 2007; Matsubayashi & Wu, 2012; Reif & Schmitt 1980). By both comparing first- and second-order elections and utilizing proxy variables for election importance, these increased turnout rates were corroborated (Stockemer, 2016). This paper is focused solely on elections in the U.S. House of Representatives so election importance is not varying over chamber type, but this does not mean that election importance is invariable. Without much difficulty, one could hypothesize two different elections to the same chamber having significantly different stakes for the respective electorates due to differences in policies proposed by the candidates. The theories discussed in this paper incorporate notions of the policies offered by competing candidates which could serve as proxies for election importance. Elections where the views and policies of candidates differ to a larger degree are likely seen as more important by the electorate than an election where candidates are quite similar. Thus, these findings indicate that when candidates are more distinct, a higher voter turnout would be expected.

The number of parties is not entirely relevant to this study as congressional elections in the United States consistently have two major parties and sometimes minor third or fourth party candidates who do not garner much support. Even so, the existing research is inconclusive as to the impact of the number of political parties with some citing positive influence on voter turnout, others arguing negative influence, and most finding no relationship (Stockemer, 2016). Voting age is consistent across the United States (and much of the world) so it will have no impact on this study's result. General voting age is distinct from the age of a given voter, which will be discussed within the socioeconomic factors.

Concerning voter registration laws, data supports logical expectation as we can observe a statistically significant negative relationship between strict registration laws and voter turnout (Stockemer, 2016). This result is significant in the scope of this paper as voter registration laws can vary rather significantly from state to state which could influence the observed voter participation rates and, if not controlled for, bias the results.

Beyond the array of institutional factors that may influence voter turnout, the impact of sociodemographic factors are of particular interest to the focus of this investigation. Research indicates that an individual's age and education tend to have the most significant influence on likelihood of voting (Wolfinger & Rosenstone, 1980; Blais, 2000). Higher religiosity, marital status, and higher incomes were all seen to have a positive, though relatively smaller, correlation with voter turnout (Blais, 2000). That said, income and turnout were found to be unrelated once a threshold of financial security had been reached (Wolfinger & Rosenstone, 1980). There is significant support for the claim that economically advanced countries have higher rates of voter turnout (Blais & Dobrzynska, 1998; Norris, 2002; Fornos, Power, & Garand, 2004). As this study compares congressional districts, it is important to consider that the average income level itself may impact turnout. There are also large differences in relative voter turnout by race, ethnicity, and gender over time, making these necessary control variables for this paper which seeks to isolate the effect of income on likelihood of voting (Leighley & Nagler, 2014). Populations of smaller size are found to have higher likelihood of voting as the relationship between citizens and their representatives is more personal (Stockemer, 2016). As a result, this analysis will also include a control for size of a given congressional district.

Policy is more closely based on income than other socioeconomic variables as it is often used as a threshold for both spending and taxation (Leighley & Nagler, 2014). As such, bias in income is of particular interest as it can rather directly create biased policies and subsequent impacts. Various studies have hypothesized that income inequality either increases or decreases voter turnout. As this paper focuses predominantly on the variable of income inequality, it is contextually relevant to observe the philosophies behind these opposing expectations. Two main hypotheses are the power theory and conflict theory. Power theory posits that higher rates of income inequality will yield lower voter turnout. The mechanism by which this functions is that income level is tied to perceived power and those who feel without power (i.e., the poor) become disengaged from the political process and, as a result, participate in said process at a lower rate (Goodin & Dryzek, 1980; Lister, 2007). In contrast, conflict theory stipulates that with greater income inequality, the wealthy and the poor both have a more significant stake in the outcome of an election as there is a greater divide in priorities (Stockemer, 2016). As a result, it is expected that overall voter turnout will increase where income inequality is high as both the rich and the poor have more to gain or lose.

The relationship between income inequality and overall voter turnout is by no means the only point of interest, however. Beyond the sheer number of people voting, it is vital to consider the composition of the voting population — that is, whether the voter turnout rates of different socioeconomic groups affected unevenly by income inequality. If that is the case, income inequality could lead to a biased voting body, whether it be by economic means, education level, or another demographic. This is arguably a more important inquiry as the overall voter turnout does little to tell us what mechanisms are acting but rather gives a more macro-level observation of the impact of an array of potential mechanisms. A variety of theories attempting to explain this relationship have been proposed and tested

Solt (2010) evaluated the legitimacy of the Schattschneider Hypothesis, which posits that where there is greater income inequality the wealthy have more power to decide which issues do or do not make the political agenda (Schattschneider, 1960). The preliminary notion of this hypothesis is rather well supported, as demonstrated by the positive relationship between the preferences of the rich and policy outcomes and the nonexistent relationship between preferences of the poor and policy outcomes (Bartels, 2008; Gilens, 2012). In essence, wealthy Americans have substantially more political influence than their poor counterparts (Jacobs & Page, 2005). Based on this rationale, we would expect to see likelihood to vote increase for wealthier individuals as the issues they value highly are the very ones being voted over. Conversely, the poor likely would not have their prioritized issues on the political agenda which would make their votes seem less valuable, leading to lower voter turnout within that economic stratum.

Yet these disparate theories do not paint an entirely clear picture. For instance, conflict theory posits that income inequality will lead to a greater clash between rich and poor over political issues, so overall voter turnout should increase. Further, as the baseline voter turnout is significantly lower for the poor than it is for the wealthy (Leighley & Nagler, 2014), the aggregate effect will likely be reduced income bias in the voting body as there is a larger number of previously non-voting poor than rich (Freeman, 2003). This prediction is in direct confrontation with the Schattschneider hypothesis, demonstrating a lack of theoretical consensus on the issue.

Another slightly different theory concerning voter turnout is the resource theory. This postulates that participation is significantly influenced by the fixed costs of taking part in the voting process such as transportation costs or taking off work (Verba, Schlozman, & Brady, 1995). As such, we would expect to see income bias in the voting population as wealthier people can bear the fixed cost of participation while said costs may be a more inhibiting obstacle for poorer citizens. With that said, the rather constant nature of fixed costs implies that at higher income levels the cost burden of participation should be equally negligible.

These contradictory ideas do not remain solely in theory, but the data provides conflicting results as well. Data on gubernatorial elections found that individuals from states with higher income inequality were seen to have a lower likelihood of voting and the population that does vote has higher rates of income bias (Solt, 2010). Similarly, utilizing data from the American Community Survey (ACS) and National American Election Studies (NES), evidence indicates that from from 1964 to 2000 the observed overall decrease in voter turnout was disproportionately drawn from the demographic groups with lower education, income, and age (Freeman, 2003). The same result of increasing income bias was found beginning in 1960 arguably caused by disproportionately lower participatory attitudes, social attachments, and demographic shifts (Darmofal, 2005).

Despite the seeming conclusivity of these results, studies utilizing the very same data sources and investigating the same questions have found conflicting results. Evidence indicates that that over time periods of increasing economic inequality, the wealthy vote consistently at much higher rates than the poor, but the difference between these voting rates remains relatively unchanged (Leighley & Nagler, 2014; Teixeira, 1992). Stability of income bias in voters was also found for midterm elections between 1958 and 1994 indicating the overall reduction in voter turnout over those years are drawn relatively equally from the different socioeconomic groups (Shields & Goidel, 1997).

These dramatically differing conclusions over the same questions, in many cases even using the same data sources, raise questions over the legitimacy of results and what result should be believed. The disparity in results are likely caused by different methods of measurement, time periods included, and the type of statistical approach and analysis (Leighley & Nagler, 2014). Due to this lack of consensus, it will be valuable to observe on which side of the existing literature the results of this paper land and how my own methods and approach compare.

2.3 - Voting and self-interest

The findings concerning voter turnout rates just discussed are illuminating insofar as different demographic groups vote differently. That is, in the impractical hypothetical where all people vote along the same lines, shifting bias in the voting body would have no impact. We can shift this hypothetical toward practicality by asking whether people vote in their own self-interest. If people vote instead on symbolic attitudes like ideology, party, or prejudicial lines it may be more difficult to predict outcomes of elections or resulting policies based on socioeconomic factors. Of course, there are some relationships between socioeconomic standing and symbolic attitudes (Janowitz & Segal, 1967; Weiner & Eckland, 1979), but not to the extent that symbolic attitudes are a sufficient predictor of socioeconomic standing or vice versa.

Therefore, the way people vote is important in evaluating the impact of bias in the voters. Existing research indicates that symbolic attitudes have a significant effect on policy preferences and voting behavior, whereas self-interest had little to no effect (Sears, Lau, Tyler, & Allen, 1980). This finding remains true when observing trends over specific issues like busing in the 1970s (Sears, Hensler, & Speer, 1979). These results imply that self-interest is frequently overestimated as the variable of influence in determining voting behavior.

Individuals who rank particular issue items highly on their personal agendas are found to be more likely to vote out of self-interest (Young, Borgida, Sullivan, & Aldrich, 1987). In relation to income inequality, we might expect poorer individuals to rank certain issues highly on their personal agendas, such as redistribution, welfare programs, or affordable education and healthcare. Despite disagreement over the specific form of the relationship between inequality and support for redistribution, the literature generally predicts such a relationship does exist and acts in the expected direction (Benabou, 2000; Meltzer & Richard, 1981).

When the implications of policies are well demonstrated, data indicates that lower socioeconomic groups express strong support for redistributive practices (Franko, Tolbert, & Witko, 2013). This indicates the importance of both what policies are proposed by candidates and the voting population's clarity of the policy implications. Note that the policies proposed are integral components of the Schattschneider hypothesis, conflict theory, and other hypotheses for voter turnout. Support for redistribution is dependent on inequality, but also on the structure of the redistributive targeting (Moene & Wallerstein, 2001).

Some research indicates these are few attitudinal differences between voters and non-voters which may imply increases in overall turnout would not affect policy outcomes (Wolfinger and Rosenstone, 1980; Verba et al., 1995). That said, we expect there are actually differences in attitude between socioeconomic groups. Thus, if there are disproportionate changes in the voting population by socioeconomic group, reducing bias in the voting body may actually impact policy outcomes. Furthermore, similarities between voters and non-voters does not extend to all policy areas. Research demonstrates that when it comes to redistribution policies, voters are consistently more conservative than non-voters (Leighley & Nagler, 2014).

The notion of shifting bias affecting policy outcomes tends to hold when looking particularly at welfare and redistribution policies. Voting bodies that have greater income bias favoring the wealthy result in governments that show less support for redistributive policies and spend less on welfare than their counterparts with more equal voting bodies (Hill & Leighley, 1992). This demonstrates that bias in voter turnout does seem to have tangible effects on policy outcomes. Therefore, even if people are not necessarily voting in their own interests, relevant policy outcomes still trend in the direction of the interests of the wealthy and away from the interests of the poor when there are biased voting populations. Due to this result, the level of bias among the voters remains relevant.

2.4 - Representativeness of politicians

The subsequent inquiry is what impact bias in the voting body has on the makeup of U.S. Representatives. That is, despite the level income inequality in the United States, do we see politicians that accurately represent the socioeconomic composition of their constituents?

Existing literature indicates that the U.S. federal bureaucracy as a whole is representative of the American people on the variables of age, income, education, and parental occupation (Long, 1952; Meier, 1975). That said, when breaking down the federal bureaucracy by pay grade and position a different trend emerges. The more elite members of the U.S. bureaucracy are far less representative of the American people (Meier, 1975), indicating that biases in representation are present in the United States. This is particularly impactful as the more elite members of a bureaucracy will be those who generally have more power and policy-making influence.

Figure 1: Representativeness of Bureaucracy

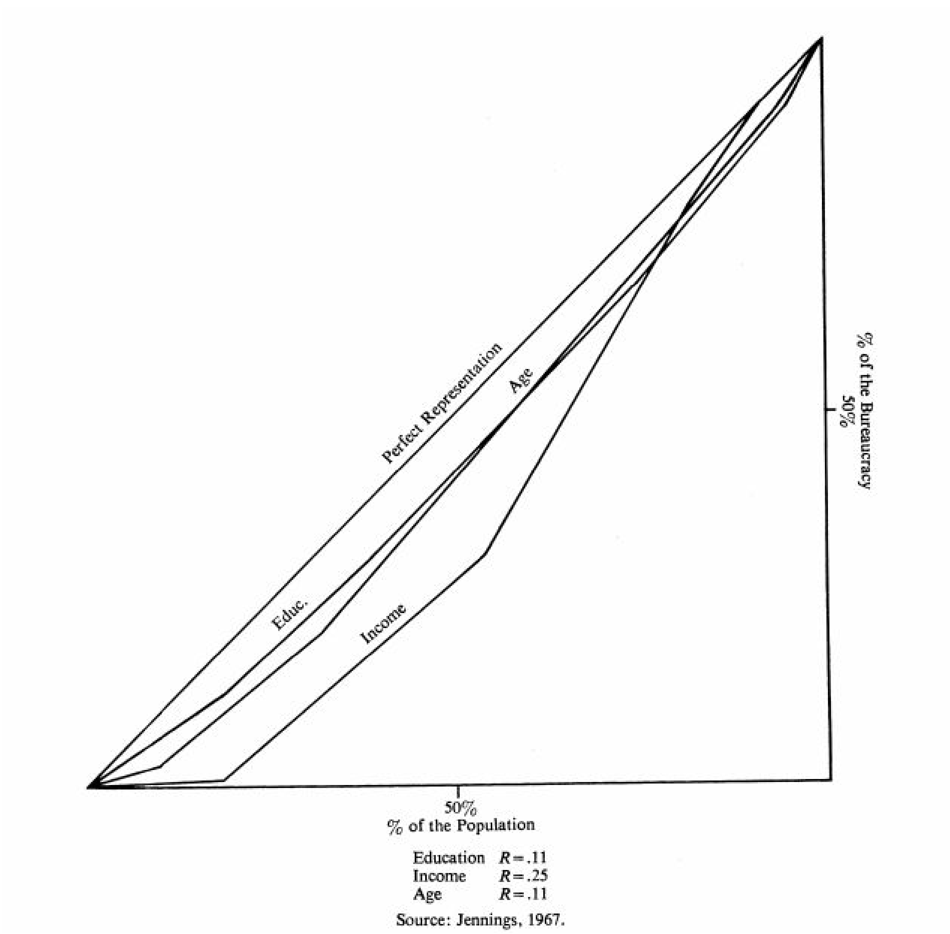
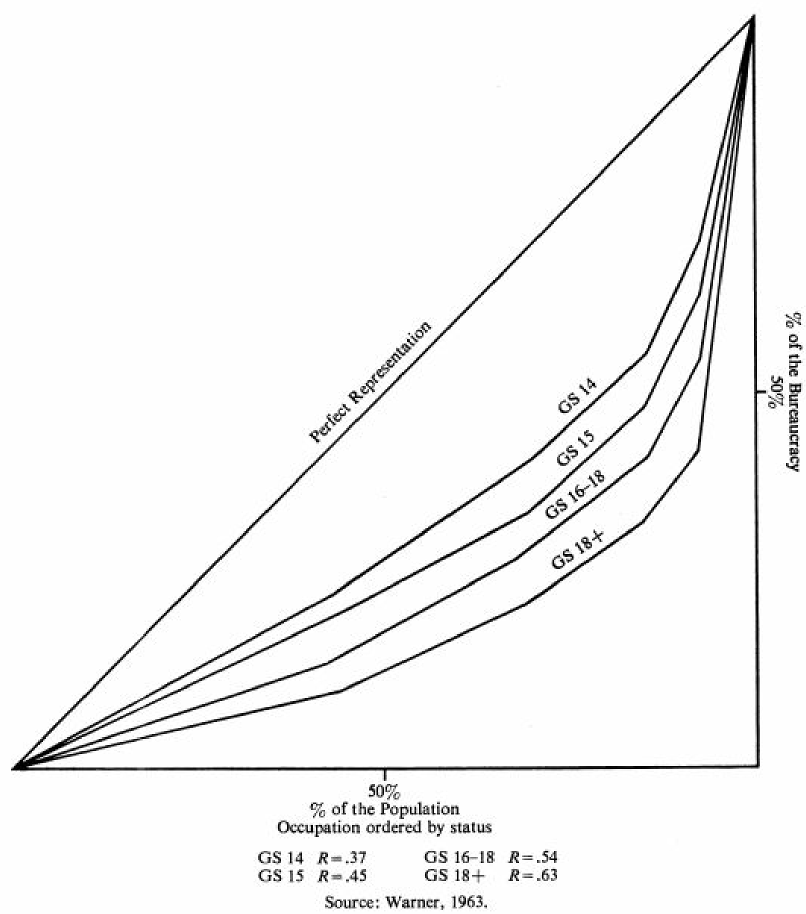


Figure 2: Representativeness of High Grade Bureaucracy



Elite members make decisions regarding general support and proposed policy initiatives with their own interests in mind. The personal wealth interests of politicians were found to be positively associated with financial sector support in general, as well as both the amount and timing of support (Tahoun & van Lent, 2013). These implications provide added motive behind the questions investigated in this paper as the wealth level of a politician has further reaching ramifications regarding policy and support.

2.5 - Policies concerning income inequality

After observing the demonstrated ramifications of income inequality and the later findings of this paper, it is useful to look into what policies and initiatives have been seen to actually reduce income inequality. This paper does not argue for the implementation of any particular method, but rather presents several such proposals that may address the underlying variable.

Policies of redistribution are some of the first proposed methods for reducing the disparity of income between members of a country. These redistributions are generally conducted either through taxation or transfers. Though focused primarily on Latin America, some literature indicates that the effectiveness of redistribution relies on the structure of that country's fiscal system and that transfers are far more effective than taxes (Goni, Lopez, & Servén, 2008). The latter of these findings is corroborated by research which indicates that taxes have an unclear effect on inequality (Glomm & Ravikumar, 2003).

The structure of a country's financial system is reliant on many factors, but government corruption can be a significant barrier to progress toward equality. A 2002 study found that increasing corruption by one standard deviation increased the respective gini coefficient by 11 points (Gupta, Davoodi, & Alonso-Terme, 2002). This large magnitude of this finding indicates that working to reduce corruption within a government would likely reduce income inequality.

Microfinance is another financially based model for reducing income inequality. Highly debated and mostly utilized in developing nations, microfinance initiatives have been linked with contradictory outcomes. A study observing 70 developing countries did find a reduction in inequality with increased levels of microfinance, but the magnitude of the reduction was not significant and should therefore not be considered a universal solution (Hermes, 2014).

Rather logically, education is considered to be a force that can either perpetuate or reduce inequalities within a population. Existing research tends to indicate improving educational quality and access through increased public education expenditures reduces income inequality (Sylwester, 2002; De Gregorio & Lee, 2002). The support for public education is particularly important as research finds larger reductions in inequality under public education as opposed to private (Glomm & Ravikumar, 1992). Despite the large support for education as a reducer of inequality, findings indicate this is a long term method of addressing inequality rather than short term fix (Glomm & Ravikumar, 2003).

All of the discussed avenues of addressing income inequality are neither definitive nor sufficient alone. Furthermore, far from all possible initiatives have been discussed, merely giving a glimpse into the possible avenues of impact. Far more evaluation should be conducted when measuring how best to affect income inequality, but as that is not the direct focus of this paper, I will not go more in depth.

Section 3: Model

In order to investigate the relationship between income and voter turnout, the following model was constructed. The unit of observation is at the individual level.

$$Voted = \alpha(Income\ Vector) + \beta_1(Age) + \beta_2(Age)^2 + \beta_3(Education) + \beta_4(Race) + \beta_5(Gender) + \beta_6(Marital\ Status) + \beta_7(Employment) + \beta_8(Home\ Ownership) + \beta_9(Religion) + \beta_{10}(Religion\ Importance) + \beta_{11}(Party) + \beta_{12}((Log)\ District\ Mean\ Income) + \beta_{13}(District\ Family\ Count) + \beta_{14}(Gini\ Coefficient) + \beta_{15}(Year) + \gamma$$

With this equation I model two distinct non-linear contributions of income on likelihood of voting. This is as the relationship between income and the propensity to vote is not expected to be consistent among all income levels. The variable "Income Vector" serves as the placeholder for these two non-linear income measures. In model 1, income vector will take the form of variables for an observation's family income and family income squared. In model 2, income vector will appear as dummy variables for the following income tiers to observe inter-tier differences.

Income Tier	Family Income Range
1	< $10,000
2	$10,000-14,999
3	$15,000-24,999
4	$25,000-34,999
5	$35,000-49,999
6	$50,000-74,999
7	$75,000-99,999
8	$100,000-149,999
9	$150,000-199,999

Model 2 includes indicator variables for income tiers 1 through 9. This means that the regression coefficients will capture the difference in propensity to vote for people in that income bucket as compared to people with family incomes of over $200,000. A family income of

$200,000 and over represents the top 4.01% of a district's population on average. The array of controls included are intended to account for any differences in composition between the voters and non-voters within my dataset. Variations in these control variables are very likely to affect the outcome variable, thus are likely different in the voting and non-voting population and are needed as controls to avoid omitted variable bias.

In order to investigate the relationship between income inequality (in a congressional district) and the net worth of representatives elected by voters in that district, the following two models were constructed:

Model 3:

$$(Log)\ Net\ Worth\ Scaled\ =\ \alpha(Gini\ Coefficient)\ +\ \beta_1(Party) + \beta_2(Year) +\ \gamma$$

Model 4:

$$Gini\ Coefficient = \alpha((Log)\ Net\ Worth\ Scaled)\ +\ \beta_1(Party)\ + \beta_2(Year) +\ \gamma$$

Model 3 investigates the relationship between the gini coefficient in the district and the (log) net worth of the respective representative. The model seeks to answer the question of whether we expect wealthier representatives in office in more unequal districts. As we are interested in how representative the elected representative is, their net worth was scaled by the mean and median family incomes in that district. The same scaling was done for model 4. A control for the representative's party was included as there may be a relationship between party and net worth as wealthier representatives may align unequally with the different political parties. Furthermore, there may be a relationship between district inequality and sitting political party through their policy decisions. As such, leaving a control for representative party out of the models would introduce the potential of omitted variable bias on gini coefficient in model 3 and (log) net worth scaled in model 4. Due to the interest in and plausibility of a positive feedback loop between income inequality and representative net worth, I anticipate these outcome variables to be correlated in both directions. To observe both directions of that feedback loop, model 4 switches the outcome metric to look at the impact of (log) net worth of the representative on the gini coefficient. Once again, this seeks to answer whether we expect more unequal districts when wealthier representatives are in office. The same party indicators were included as controls as there may be a relationship between district inequality and sitting party through their policy decisions.

There are, of course, a variety of other covariates that likely influence (log) net worth and gini coefficient. Quality control data for these covariates could not be acquired under the constraints of time and resources. These omitted covariates and the likely direction of their bias are discussed in detail within the following data section.

Lagged regressions were conducted for both Models 3 and 4. For Model 3, regressions with gini coefficient lagging 1 and 2 years behind (log) net worth are included. Similarly, for

Model 4, regressions with (log) net worth lagging 1 and 2 years behind gini coefficient are included. These are present to observe if there is a delayed relationship between these two variables of interest. That is, a higher gini coefficient may result in a wealthier representative in the future but not in the directly matched year. This could be due to a delay in the turnaround between income inequality and unequal political mobilization. Similarly, a wealthier representative could result in more unequal districts in the future. But the mechanism by which this is expected is the enactment of policy which takes time to create, implement, and observe the results. Therefore, lagged regressions may provide more insight into the true relationships concerning the positive feedback loop between income inequality and representative net worth.

Section 4: Data

The data on net worth of congressmen was acquired from the Center for Responsive Politics. The net worth values of congressmen were calculated by summing a member's assets and subtracting their liabilities. Both assets and liabilities are reported in ranges of values in the financial disclosure reports from the Office of the Clerk of the House from which these values are drawn. Due to this, the net worth value in the dataset is a range of possible values. These personal financial disclosure forms are filed by May 15th each year corresponding to the previous calendar year. The data covers the years from 2004 to 2014. Average values were used to test this paper's models. For particularly large values, the dataset caps individual assets at over $50 million and spouse or dependent child assets at over $1 million. These value caps were left unchanged, which may limit the accuracy concerning extremely wealthy individuals. These caps have the potential to skew average net worth values downward.

Net worth as an outcome metric is far more reflective of wealth than income as it incorporates value over time. It is also steadier than income over time, as income varies more when changing jobs. If income were used instead of net worth as the outcome metric, note that as an individual is elected their income would change to the standard income for a U.S. Representative. While there may still be variation due to income earned from other sources, such as investments, it would still be less representative of true wealth differences between representatives. The stability of net worth has disadvantages as well. There were some congressmen in the dataset that had no variation in their net worth data over a range of years. For these instances only one unique entry was kept. As this paper seeks to examine how changes in inequality influence changes in net worth, without any variation in net worth no results could be observed.

For some representatives, net worth data was included in the dataset for years prior to their being in office. All such observations were dropped. There were a handful of other miscellaneous issues with the net worth data which resulted in some observations not matching with the economic variables by district. These observations were also dropped. In each of the various merges (direct year and lags) these issues accounted for less than 200 observations out of over 3,800 total.

Ideally, to observe the true relationship between gini coefficient and politician net worth, an array of controls would be included. Unfortunately, adequate control data for politician net worth could not be acquired except for political party. This limitation implies that the findings from model 3 must be taken with some hesitation. That is not to say the results are invalid, but merely that there is most definitely omitted variable bias influencing the findings. In an attempt to parse out some of these biases, I will now discuss control variables I would have liked to include and the expected direction of their influence on net worth, and subsequently on gini coefficient to anticipate the impact of their omission.

Years of education, parent's income, age, and years of work experience would all be expected to have a positive impact on a person's wealth. The longer an individual is educated and the more experience they have, the more expansive their human capital. The level of human capital impacts net worth through income as higher paying jobs are available to those with more education and experience, and over time these higher incomes accrue into higher levels of net worth. Similarly, age is positively associated with the development of human capital through means like education and experience. While conceptually inclusion of age may bring collinearity concerns with education and experience, these concerns are small relative to the amount of variation it accounts for. Parent's income has a rather direct positive relationship with net worth as much of net worth is acquired from parents. Also, parent's income will impact the amount of capital an individual has to build on and the more they have to start the faster the wealth can grow. The impact of race on net worth is less straightforward as it is a categorical variable. We would expect certain races to have a positive effect on net worth and other races to have a negative effect.

To understand how these omitted variables will bias the findings of this paper, the relationship between each and gini coefficient must be explored. As these control variables are constructed with net worth of an individual in mind, they are measures at the individual level. To see their relationship with gini coefficient, which is not a metric over individuals, I will consider the average of these control variables over a given district.

There is a cap on how high educational attainment can go, so wealthier individuals can only increase their educational standing so far whereas poorer individuals have much room to improve. As a result, an increase in average education in a district will likely see more educational improvement from lower income tiers than from higher ones, reducing inequality. Thus the relationship between education and gini coefficient is expected to be negative, which means the bias results in underestimating the impact of gini coefficient on net worth of representatives.

Parent's income does not have a similar cap at the high end. This means that an increase in average parent's income for a district could either mean that values increased for the low income individuals or an increase for the wealthier individuals is simply skewing the mean higher. If the former is true, increasing parent's income is an equalizing force, so it would have a negative bias on the impact of gini coefficient. If the latter is true, it would indicate the rich getting richer, which would be positively associated with gini coefficient. Then the omission would present positive bias on the impact of gini coefficient.

The relationship between age and gini coefficient is rather ambiguous. Age is expected to be positively related with net worth as stated above, but has a less linear relationship with income. Age, as well as years of work experience, will have a critical point in its relationship with income. This is as after a certain age productivity declines and most individuals retire. For work experience, after a certain value of years it begins to indicate an individual started working earlier, sacrificing years of education, which relates to lower levels of income. So if we see an increase in average age of a district, we do not know where the value is in relation to the critical point or whether the increase is from poorer individuals, richer individuals, or both equally. If the increase is below the critical point, it indicates income level increases for whatever population subset is causing the increase in average. If the average is above the critical point, the opposite is true. Thus depending on where the average is in relation to the critical point and what population subset is driving the increase, the bias of omitting age on the impact of gini coefficient can be negative, positive, or non-existent. The same is expected for years of experience.

The relationship between race and gini coefficient is quite ambiguous. First, the relationship between income and race varies dramatically based on the race. Second, the impact on inequality of increasing the representation of a given race in a population relies on the previous composition of the population. To try to make sense of this relationship we can think of the race control as a metric ranging from homogeneity to heterogeneity. Therefore an increase in the race control variable would be an increase in racial diversity, likely causing an increase in income disparity in the population. This control and gini coefficient would have a positive relationship, so its omission presents positively biased results.

Similarly, I have insufficient control data for model 4. To that end, I will now discuss the variables that likely impact gini coefficient, and their relationship with net worth. As gini coefficient is a measure of income inequality, various policies are going to have the largest influence over it. Policies regarding taxation, transfers, healthcare, education, welfare, and employment would likely have the most significant relationship with inequality. Taxes, transfers, employment, and welfare policy all directly affect distribution of income. Education policies have the potential to affect levels of human capital which impact the types of jobs attainable and subsequently the level of income received. Healthcare policies will affect how healthy and able to work individuals are, which in turn influences the number of days they can work as well as their productivity, both of which lead to changes in income. If all of these policy variables were controlled for, it would likely lead to some collinearity issues with the party controls, which could probably then be removed.

In order to understand how these omitted variables are biasing the results, we must look to their interaction with net worth. All of these policies can vary on a scale from what I will call "pro-poor" to "pro-wealthy" in a slight oversimplification. Pro-poor policies will focus efforts on providing service to low-income communities at little to no cost. Pro-wealthy policies will instead aim to promote pay-your-way strategies which inherently favor the wealthy and do not allow the poorer members of the community to afford quality services. As our variable of interest in this model is net worth of representatives, not simply net worth, we must consider where in the distribution of income representatives lie. It is a somewhat fair assumption to state that representatives will be on the higher end of the income distribution as they have the means and education to reach visibility and then public office. This is not an empirically based assumption, but could definitely be investigated more thoroughly to better understand the directions of bias here.

Under this assumption, we would expect pro-wealthy policies to have a positive effect on the net worth of representatives through a positive effect on their income. Conversely, pro-poor policies generally are more of a financial burden for the more wealthy members of a community as they finance a portion of the services provided to poorer individuals. As such, implementation of pro-poor policies would have a negative relationship with net worth of representatives. We can think of policies on a scale from pro-poor to pro-wealthy, meaning an increase in the policy control is a shift toward pro-wealthy and away from pro-poor. Then omitting controls for the policies discussed, which have a positive relationship with net worth of representatives, would mean the coefficient of net worth is likely overestimated.

The various economic data was gathered from the American Community Survey's (ACS) 1-year estimates. The 1-year estimates were used as opposed to the 3- or 5-year estimates to promote accuracy and comparability over years. All of the data acquired was broken down at the level of U.S. congressional district. The variables utilized from the ACS are gini index, mean income, median income, and income tier by both household and family. The data acquired ranges from 2006 to 2015. Data for 2015 was included despite the net worth data only reaching 2014 because of the inclusion of lagged regressions in this paper.

From the following figures some basic trends in the net worth and ACS data can be observed. They are each averaged over all congressional districts by year. From Figures 3, 4, and 5 we can observe that net worth of representatives, mean and median family incomes, and gini coefficient are all increasing rather steadily over the time span of 2006 to 2014. The values for mean and median family income are inflated, likely due to a compositional effect of averaging values over all congressional districts.

Figure 3: Net Worth of Representatives Over Time

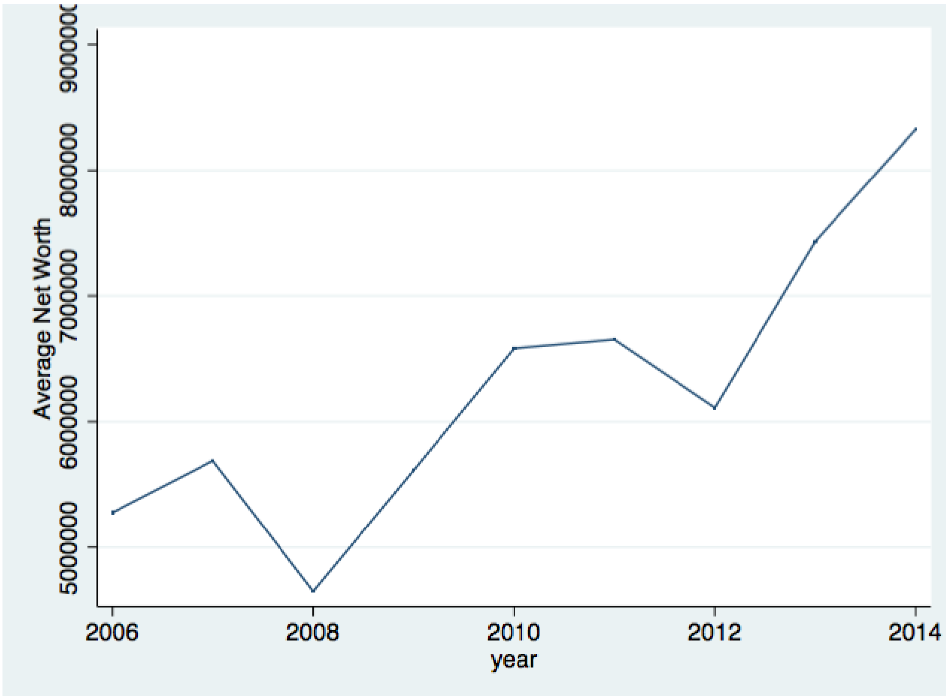


Figure 4: Mean and Median Family Incomes Over Time

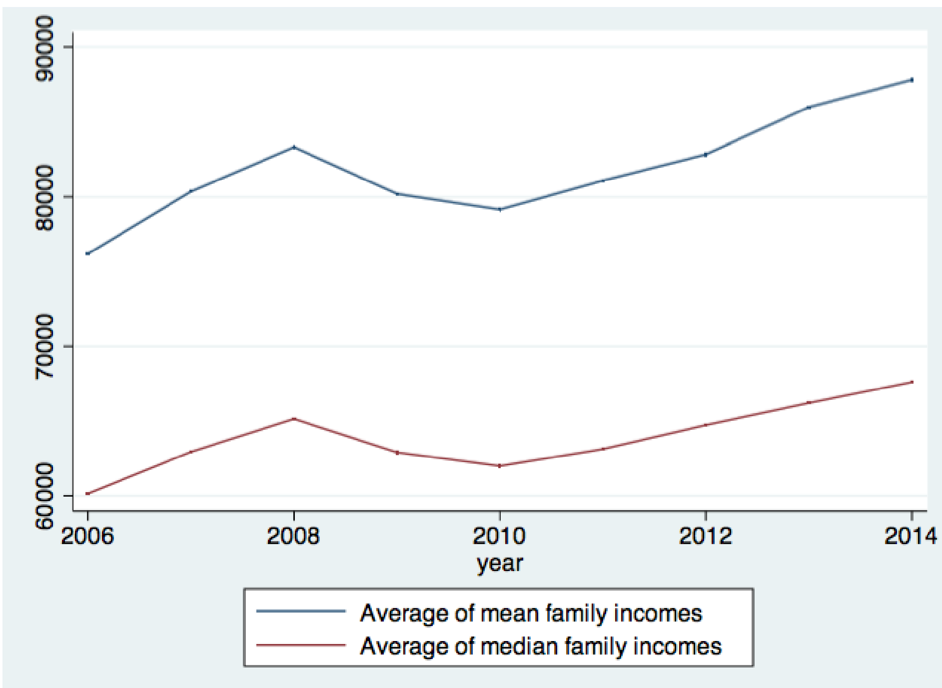
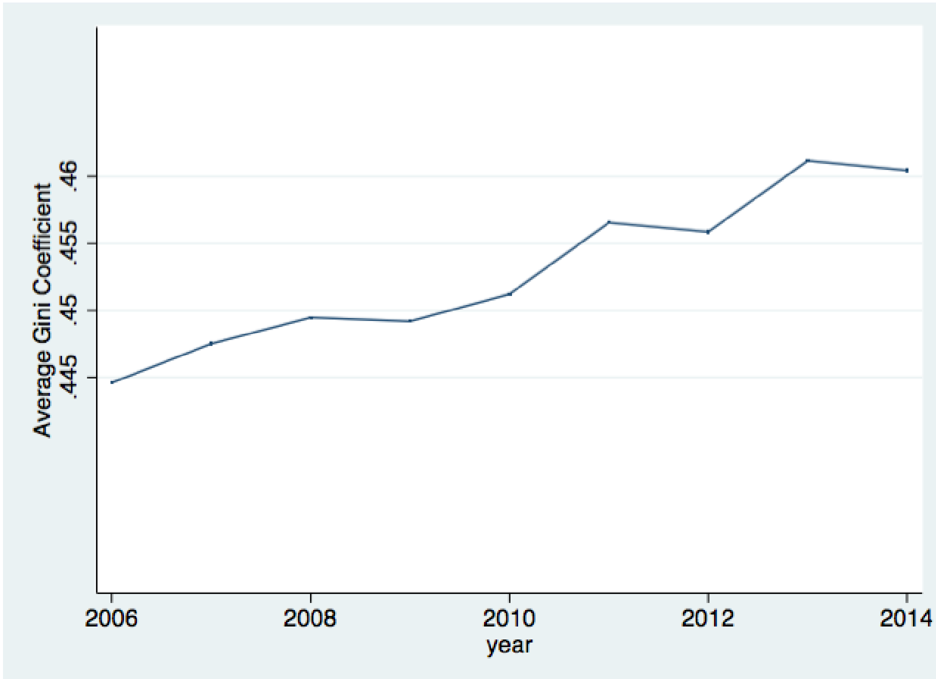


Figure 5: Gini Coefficient Over Time



The following figures demonstrate the distribution of the variables net worth, mean family income, and gini coefficient. The net worth data seen in Figure 6 appears to have quite a few observations well above the bulk of observations. These are not real concerns as outliers are due to the nature of net worth. That is, extremely wealthy people are not necessarily the result of data issues. The same idea holds true with the values of mean family income observed in Figure 7. Also included is a breakdown of the political parties of representatives by year.

Figure 6: Distribution of Net Worth of Representatives Over Time

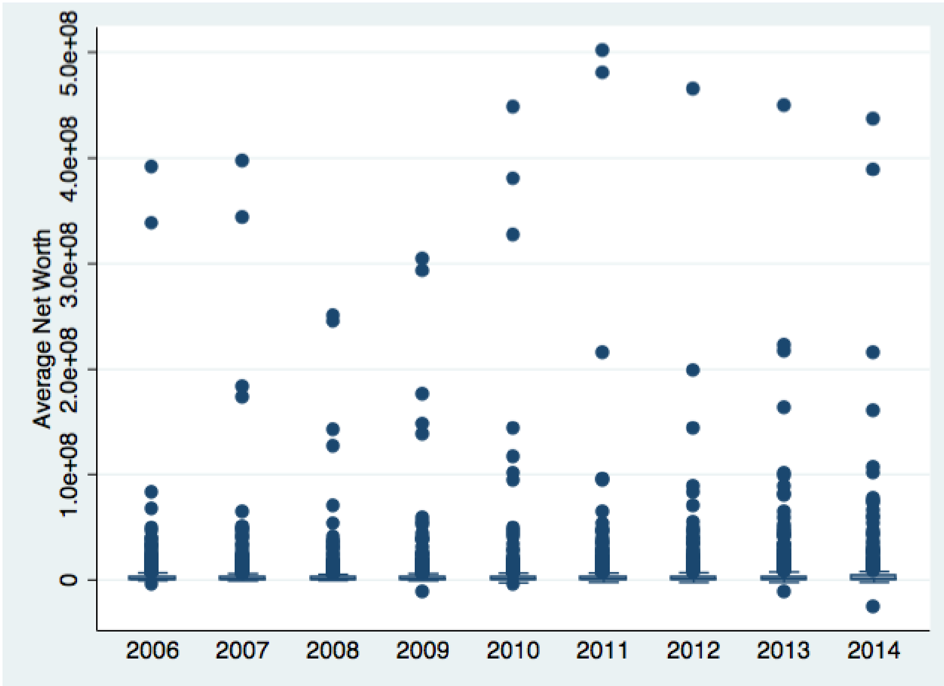


Figure 7: Distribution of Mean Family Income Over Time

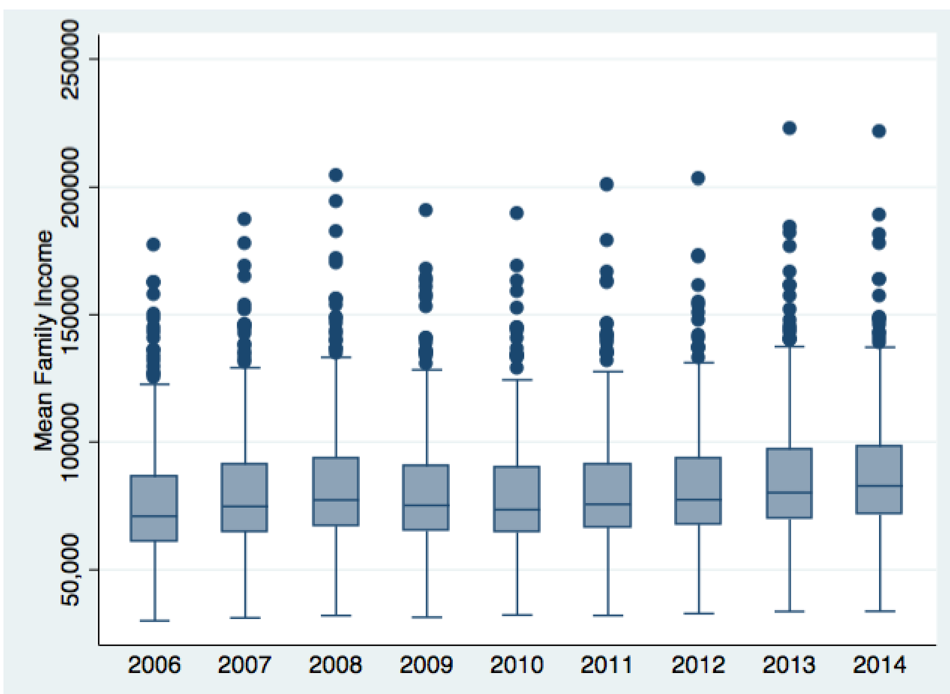


Figure 8: Distribution of Gini Coefficient Over Time

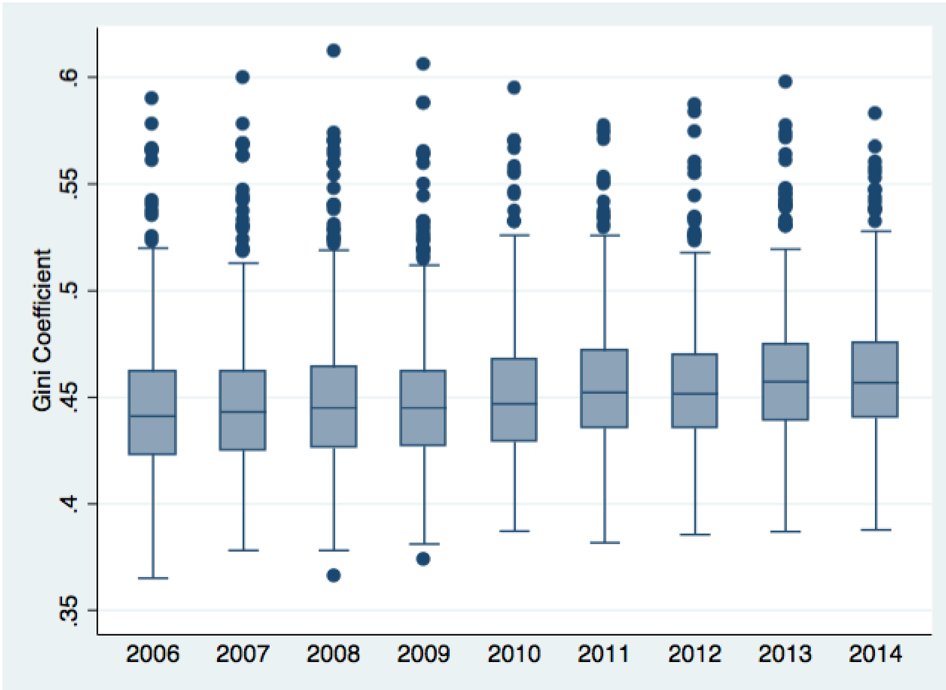
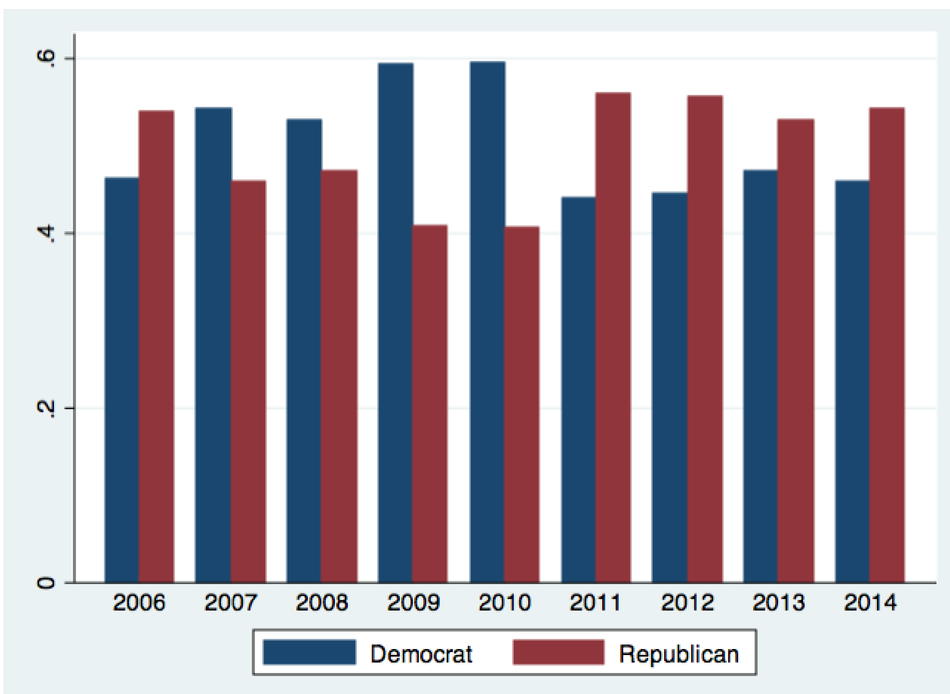


Figure 9: Share of Elected Representatives by Party



We can observe general summary statistics for this data in the following table. These summary statistics are pulled from the direct year merge, so lagging regressions will have slightly different summary statistics. In the direct merge there are 3,637 observations. As some of the variables themselves are means and medians, observing their distribution must be done with care. For example, interpreting the mean of median family income is the average value of median family incomes across all districts and years in the dataset. We can see that the mean of mean family income is roughly $18,000 higher than that of median family income which indicates there are wealthy families which pull the mean higher than the median. Heat maps for average net worth, gini coefficient, and mean family income for the years 2006 and 2014 can be found in Appendix A.

Table 1: Net Worth and ACS Data Summary Statistics

Variable	N	Mean	Standard Deviation	Median	Minimum	Maximum
Average net worth	3,637	$6,222,405	3.01e+07	$827,512	-$2.52e+07	$5.01e+08
Gini coefficient	3,637	.45272	.03280	.4482	.365	.612
Mean family income	3,637	$81,797.9	23299.6	$76,458	$29,982	$222,793
Median family income	3,637	$63,815.6	16987.6	$60,522	$20,425	$135,276

As the net worth data did not identify the district of the representative, intermediary match data was required to link the net worth data with the ACS economic data. Congressional profiles were drawn from History, Art & Archives for the United States House of Representatives for the 108th through the 114th Congress.

The voter data used was drawn from the Cooperative Congressional Election Study (CCES). This study creates a national sample which is stratified by both state and type of district for optimal comparability at the congressional district level. As I was interested in individuals who did or did not vote in congressional elections, only data from even numbered years were utilized. Questions concerning vote choice in the congressional election as well as an array of controls were pulled from the CCES data.

The majority of existing studies which utilize bias of voters as the outcome variable have used the American Community Survey (ACS) and National American Election Studies (NES) as their data sources. As discussed in the literature review, despite utilizing the same sources and looking at the same variables, different papers found directly conflicting results. Due to this, I chose to utilize a data source distinct from those commonly used so my results are not merely adding another paper with the same data to either side of the debate. Rather, the results take advantage of a different data source to add weight to the argument of one side through new empirical support.

There are no non-voting observations in the dataset for years 2012 and 2014. This presents an issue in that the regressions in this paper are utilizing a voting indicator variable as the outcome variable. As such, there is no variation in this indicator for the years 2012 and 2014 as all observations voted and will have a value of 1 on the indicator variable. So, any fluctuations in the variables of interest and controls will have no fluctuation in voting likelihood to account for. To observe the impact of this, a separate model will be included which excludes the years 2012 and 2014 from observation.

From the following figures, we can observe breakdowns of key variables in the voter dataset. These breakdowns are of the entire dataset, not just observations who have voted in their congressional election. The average observation is between ages 50 and 60 and has a family income of around $50,000. Gender is relatively evenly represented in the dataset. The dataset overrepresents whites with 80.2% of observations reporting their race as white while the true U.S. population value is closer to 63%. Both hispanics and blacks are underrepresented, composing 5.7% and 8.8% of the voter dataset, respectively. Their true population representation is around 17% hispanic and 12% black.

Education level is somewhat varied with most observations either having a high school degree, some college, or a 4-year degree, though postgraduate is increasingly represented in the later years of the dataset. Most observations are married and employed full-time, with the employment status of retired accounting for the second most observations. The religion most represented is Protestantism, followed by Roman Catholicism. Most individuals in the dataset are homeowners. In terms of party affiliation, the distribution between Democrats and Republicans is relatively even, as most individuals self-categorize as strong members of their party. The middle income tiers are most represented in the dataset, with higher middle income tiers most represented.

Figure 10: Distribution of Age

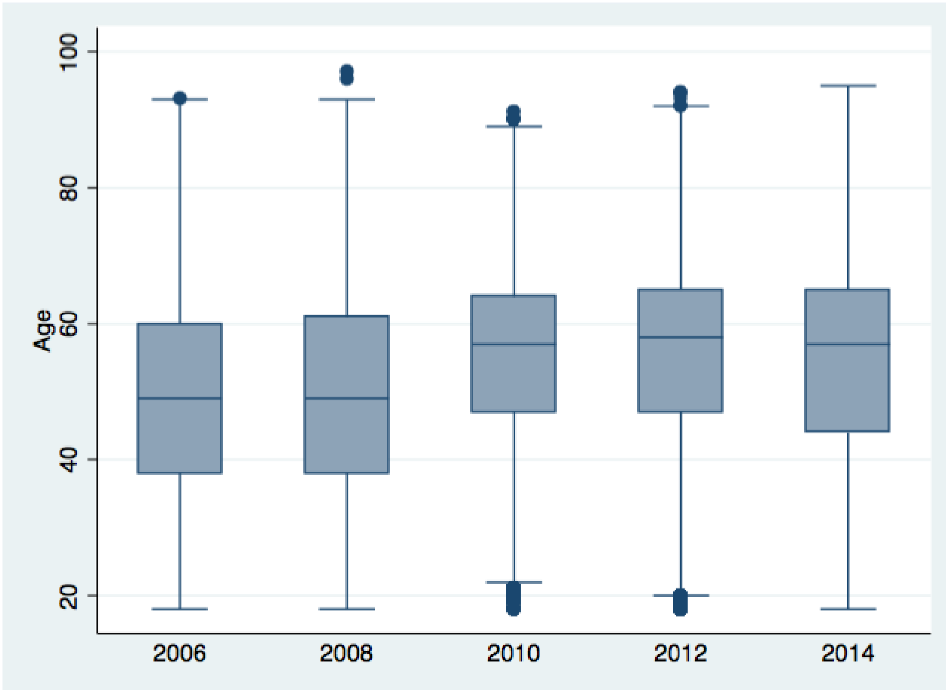


Figure 11: Distribution of Family Income

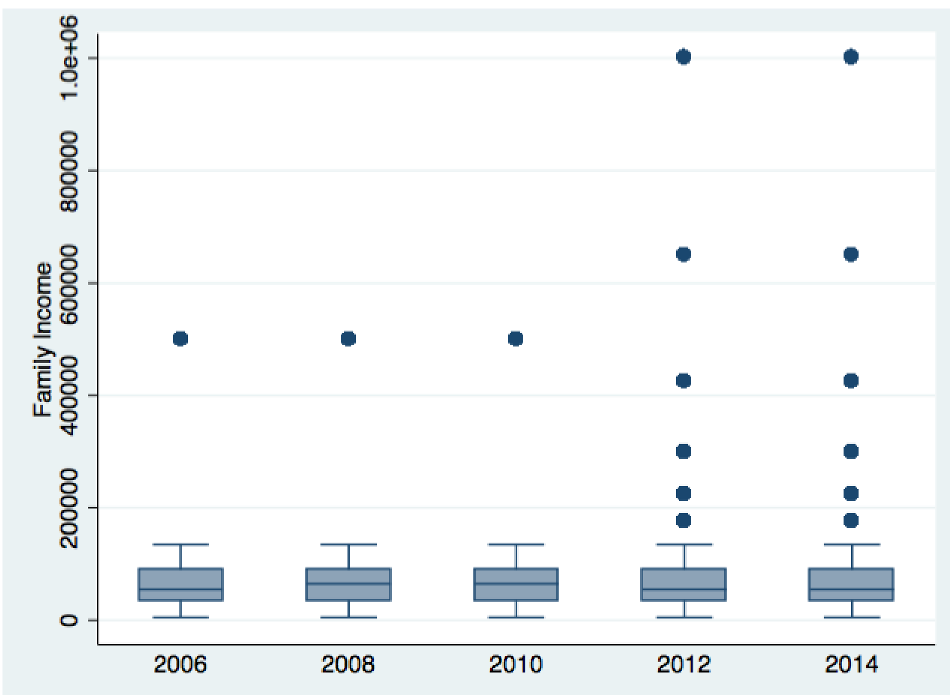


Figure 12: Gender

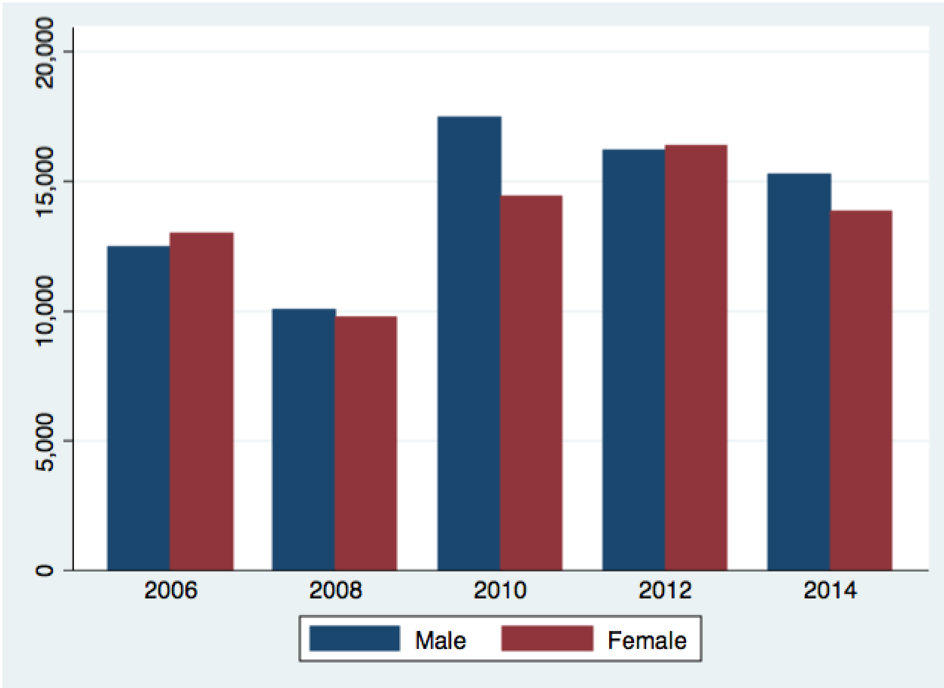


Figure 13: Race

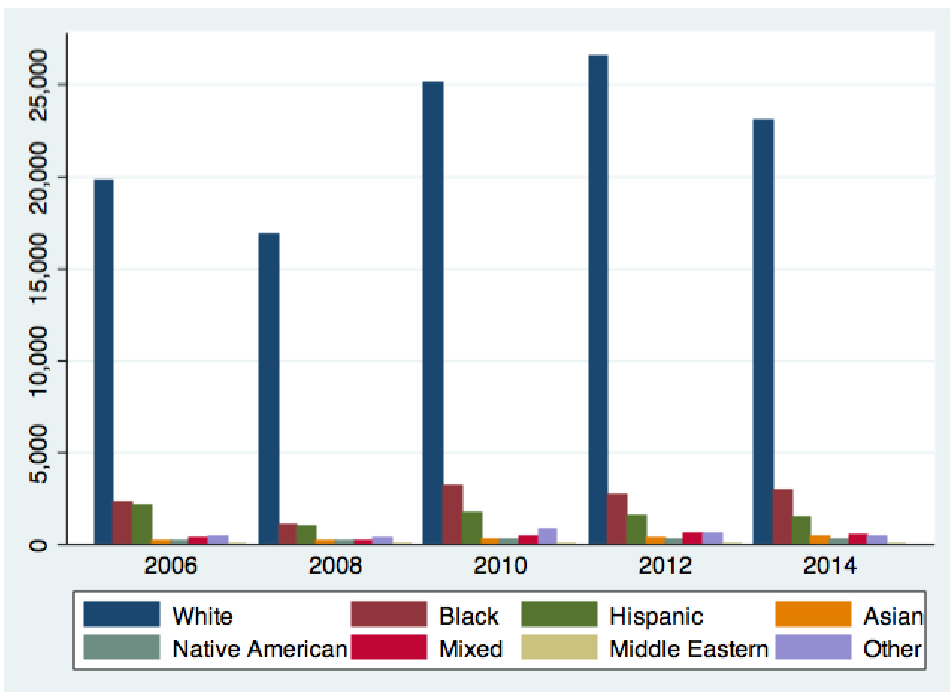


Figure 14: Education

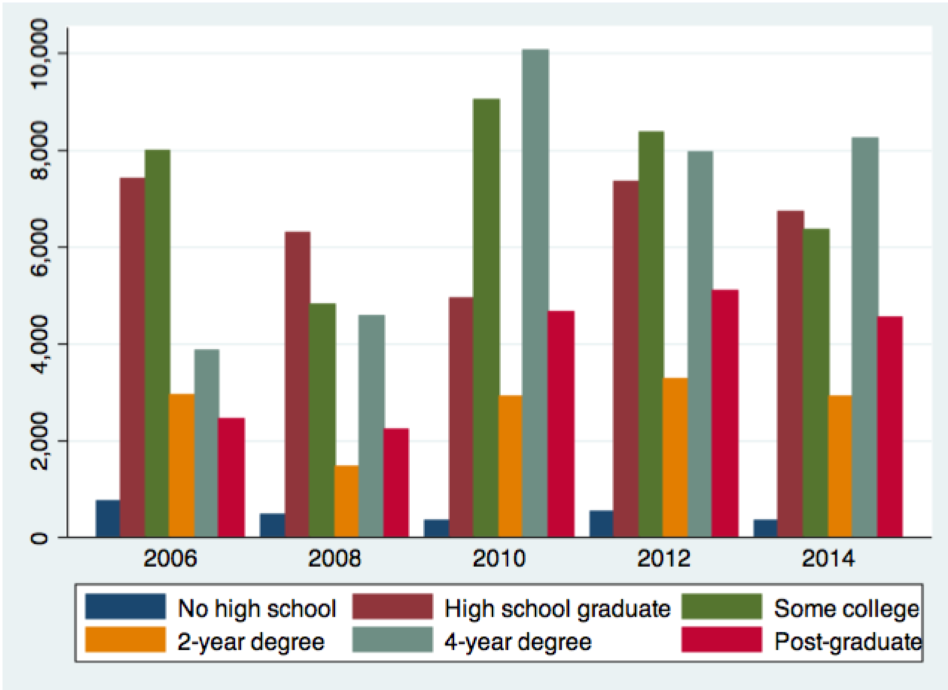


Figure 15: Marital Status

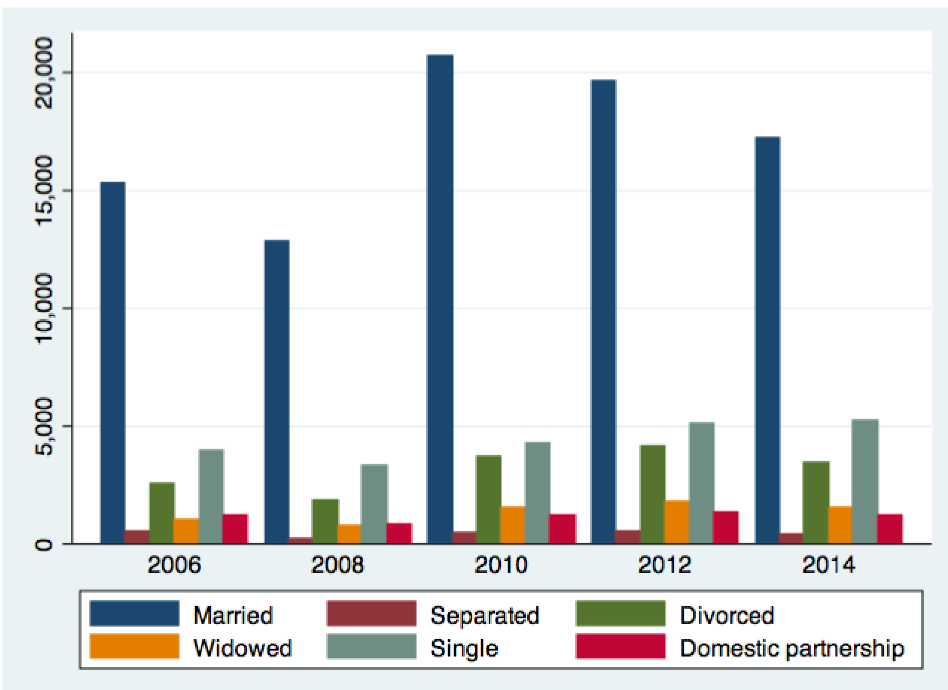


Figure 16: Employment Status

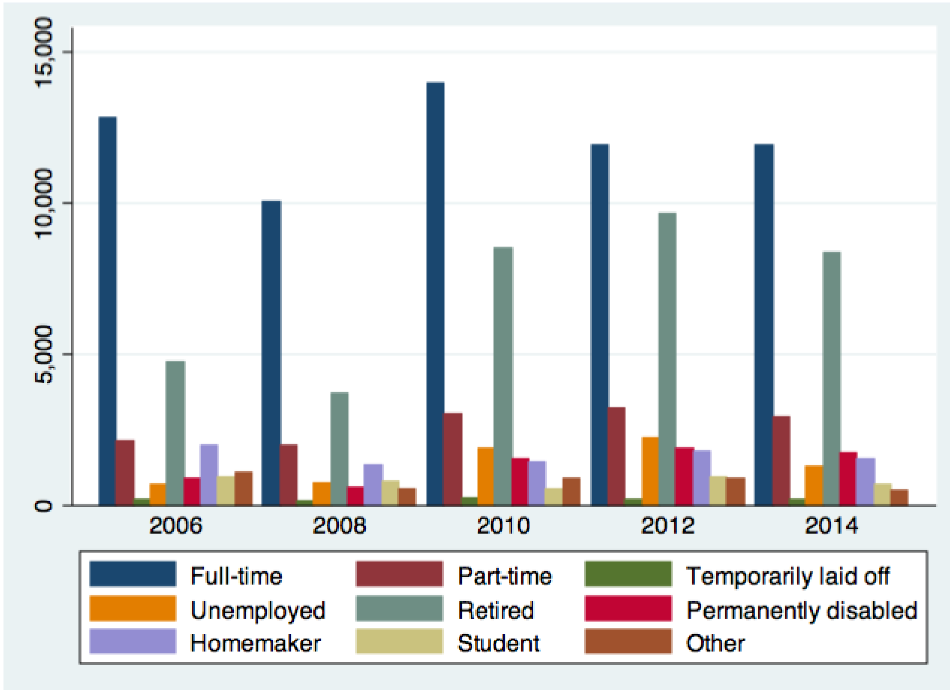


Figure 17: Home Ownership

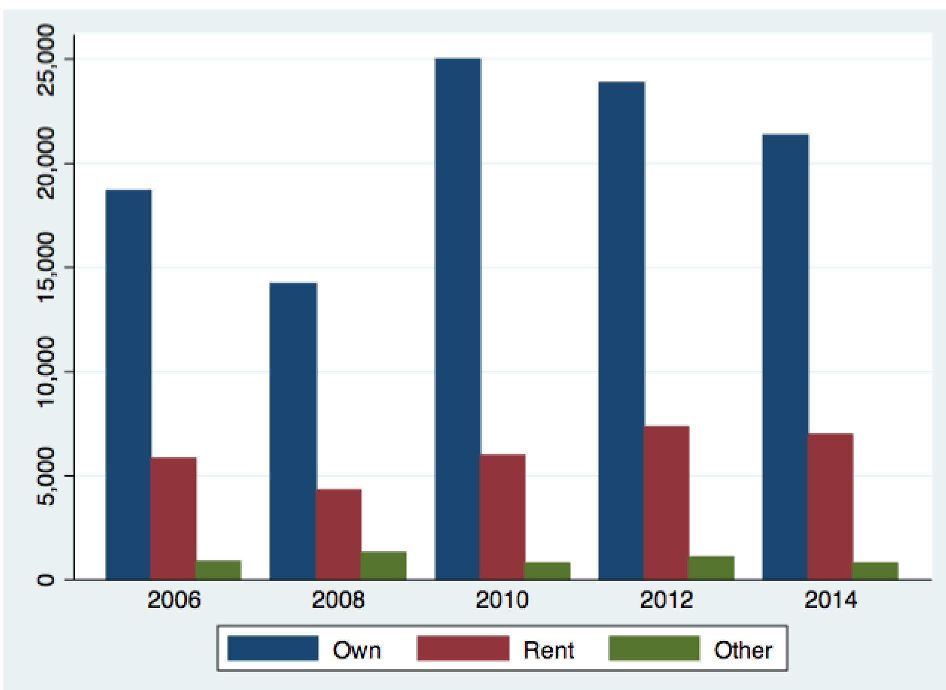


Figure 18: Religion

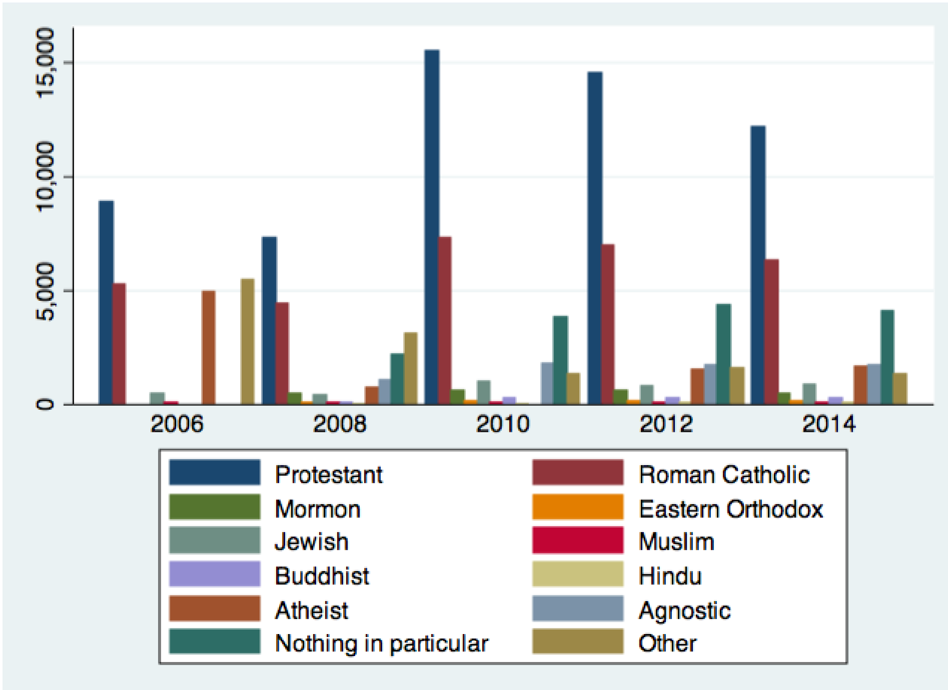


Figure 19: Religion Importance

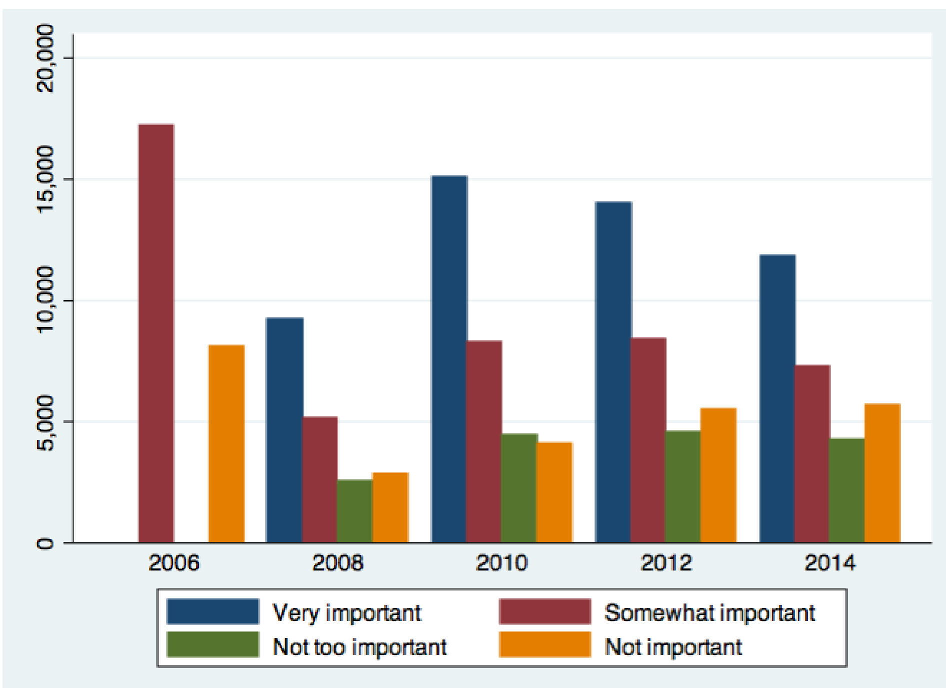


Figure 20: Party Affiliation

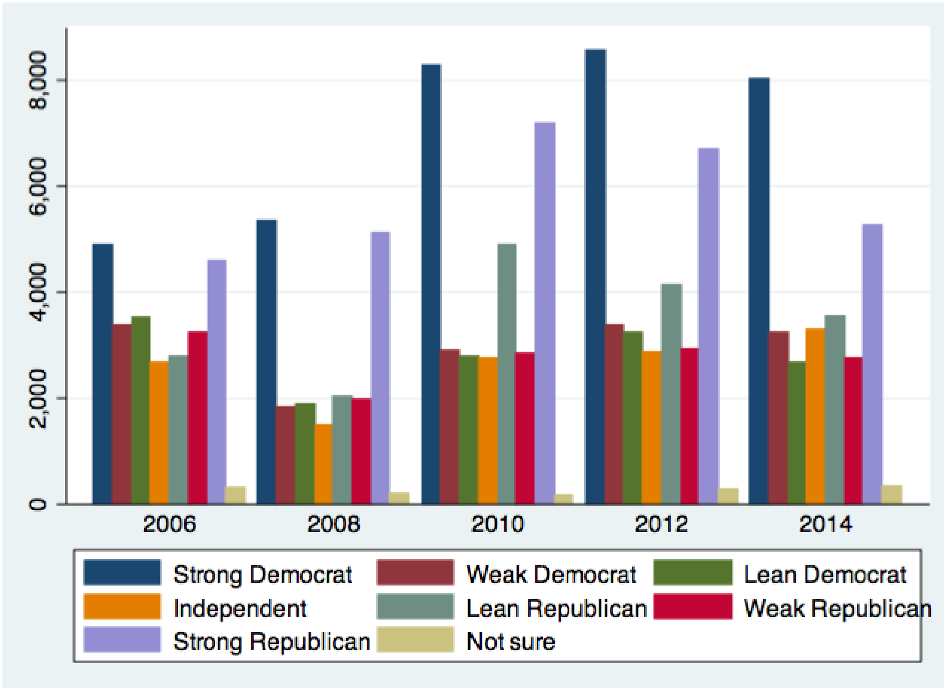


Figure 21: Income Tier (1)

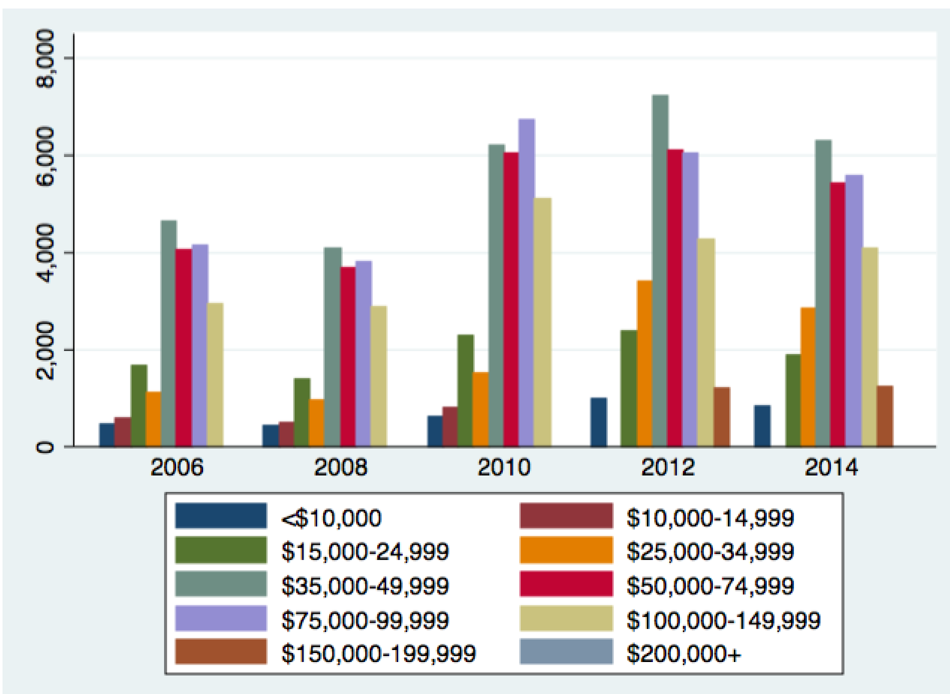
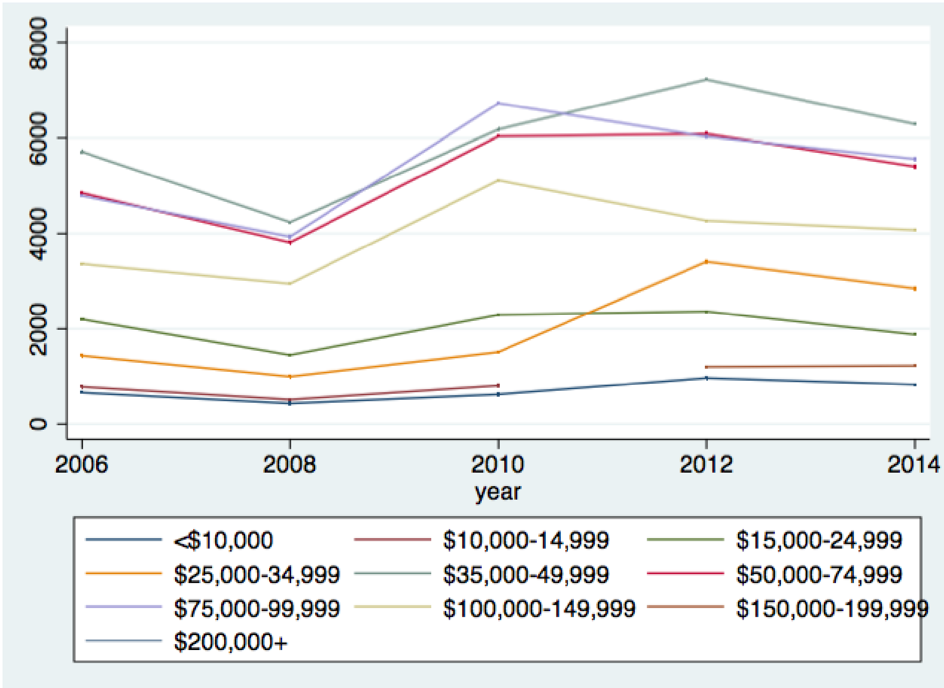


Figure 22: Income Tier (2)



The following table has the distribution for the quantitative variables in this dataset. Gini coefficient and mean family income from the ACS data are included in this table. This is as districts are unequally represented in the voter dataset so their distributions will be different than those reported in Table 1. We can observe that the distribution of family income in the voter data is quite close to distribution of mean family income in the ACS data indicating that the observations in the voter data are relatively representative of the population at large. That said, the standard deviation is significantly higher in the voter data. This is likely because the family income of observations within the voter dataset were reported in ranges, so their actual income was recorded as the mean of their represented range. For the lowest tiers (less than $x) and highest tiers (above $x) values of $5,000 and $1,000,000 were used, respectively. These values are likely skewing the distribution toward the extremes which inflates the standard deviation as observed.

Table 2: Voter Data Summary Statistics

Variable	**N**	**Mean**	**Standard Deviation**	**Median**	**Minimum**	**Maximum**
Family income	138,731	$85,401.7	103107.2	$55,000	$5,000	$1,000,000

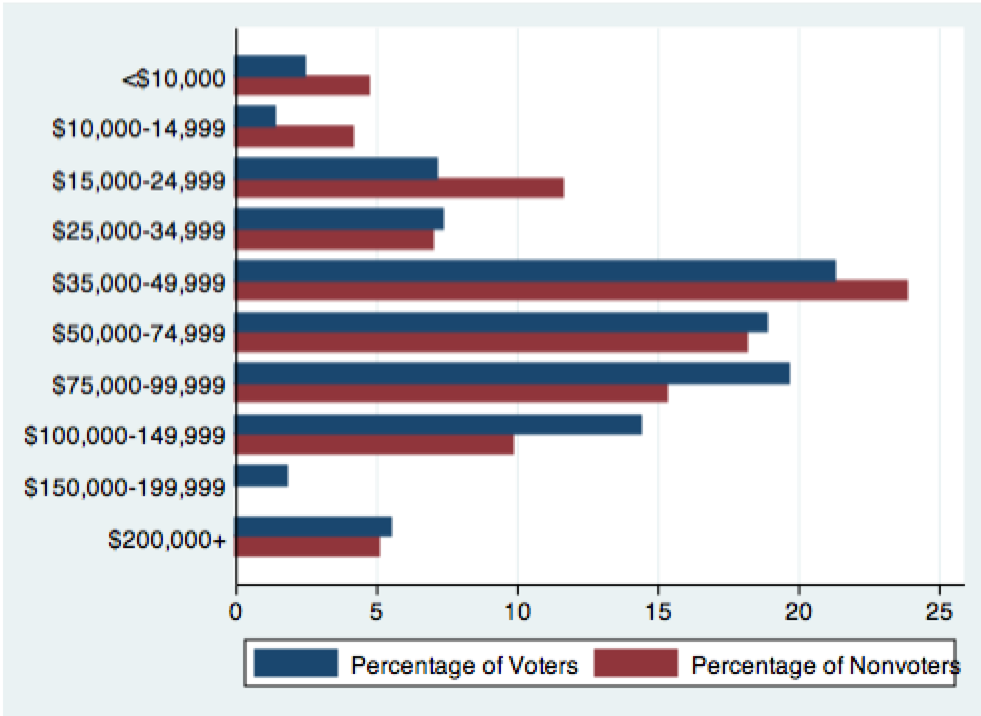
(Voter data)						
Age	138,731	52.884	14.914	55	18	97
Gini coefficient	138,731	.45011	.03055	.4461	.365	.612
Mean family income (ACS data)	138,921	$82,853.3	22203.3	$77,592	$32,015	$221,157

It is not necessary to compare the voting and non-voting subsets of the CCES data, given that the models are controlling for the variables that may show compositional differences. Thus, any inequalities between these population subsets will be captured in the coefficients of the covariates in the regressions with an indicator for voting as the outcome metric.

That said, it may be of interest to know what these initial differences look like. The voting and non-voting subsets of the dataset were compared. Males make up a larger percentage of the voting subset than females. Whites also make up a larger percentage of voters than they do of non-voters, whereas all races except 'other' have the opposite hold. In terms of education, individuals with a 4-year degree or postgraduate education levels are overrepresented in the voting body, while observations with no high school experience, high school graduates, and those with some college are underrepresented.

For religion, Protestants, Mormons, Jews, Agnostics, and those who classify as "Nothing in particular" saw the largest positive difference between their representation among voters and non-voters. Atheists and "Other" are overrepresented among the non-voters. With employment, observations who are retired are the most overrepresented in the voting subset of the dataset, while homemakers and students are most underrepresented. Both of these results are quite logical. Observations that identify as either a strong Democrat or Republican compose larger percentages of the voting subset than the non-voting subset.

Figure 23: Income Tier Representation Among Voters and Non-voters



From Figure 23 a breakdown of income tier representation can be observed. These are, in a sense, uncontrolled preliminary results of model 2. The lowest income tiers are most underrepresented in the voting body, while the higher end of the middle income tiers are most overrepresented. Note that there are no non-voting observations for income tier 9 ($150,000-199,999) which must be monitored as results are further discussed.

Section 5: Results

To begin with, observe the individual level regressions examining the effect of income on likelihood of voting (ie. models 1 and 2). Note that as the outcome variable for these models is an indicator variable, 1 is the maximum possible value and 0 the minimum. That is quite important in understanding the magnitudes of the coefficients.

Variables	**Model 1**	**Model 2**
Outcome metric:	Voted Indicator Variable	Voted Indicator Variable
(log) Income	0.0300** (0.00927)	—
(log) Income squared	-0.00109** (0.000416)	—
Income tier 1	—	-0.0268*** (0.00462)
Income tier 2	—	-0.0215** (0.00674)
Income tier 3	—	-0.0168*** (0.00317)
Income tier 4	—	-0.0152*** (0.00288)
Income tier 5	—	-0.00927*** (0.00243)
Income tier 6	—	-0.00658** (0.00237)
Income tier 7	—	-0.00321 (0.00228)
Income tier 8	—	-0.00289 (0.00229)
Income tier 9	—	-0.0293*** (0.00211)
Age	0.00332***	0.00332***

	(0.000233)	(0.000234)
Age squared	-0.0000235*** (0.00000214)	-0.0000235*** (0.00000214)
Education	0.00395*** (0.000340)	0.00405*** (0.000342)
Race	-0.000989* (0.000411)	-0.000998* (0.000411)
Gender	-0.00959*** (0.000990)	-0.00956*** (0.000991)
Marital status	0.000878* (0.000346)	0.000864* (0.000347)
Employment	-0.000754** (0.000251)	-0.000769** (0.000251)
Home ownership	-0.00588*** (0.00118)	-0.00586*** (0.00118)
Religion	-0.000962*** (0.000158)	-0.000955*** (0.000158)
Religion importance	-0.00795*** (0.000567)	-0.00788*** (0.000567)
Party	-0.00277*** (0.000214)	-0.00277*** (0.000214)
(log) District mean income	0.0163*** (0.00214)	0.0168*** (0.00215)
District family count	0.000000164*** (2.76e-08)	0.000000162*** (2.76e-08)
Gini coefficient	-0.107*** (0.0185)	-0.106*** (0.0185)
Year	0.0171*** (0.000241)	0.0172*** (0.000244)
Constant	-33.76*** (0.485)	-33.80*** (0.490)

N	134,773	134,773
Adjusted R-squared	0.095	0.096

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Both models have rather low R-squared values, which indicates that much of the variation in the voting indicator variable is not explained by variation in the covariates included in these models. It is quite possible that the importance of an election, which existing literature has seen to positively impact voter turnout, is not well represented by these covariates. Most of the institutional variables that have been tied to voter turnout either do not vary or are controlled for. Dummy variables for state will be included in robustness checks in the following section. Most of the socioeconomic variables that have been tied to voter turnout are controlled for. It is possible the data quality of those controls or the way they are coded are limiting some of their explanatory power. The models will be tested with dummy variables for each categorical variable suboption to explore this plausibility.

For model 1, we can observe that all coefficients are statistically significant at the 5% level. Most variables remain statistically significant at 1% and even 0.1%. The (log) income variable has a positive values coefficient of 0.03 while (log) income squared has a negative coefficient with magnitude 0.00109. This indicates that if there is an increase in (log) income by 1 we would expect likelihood of voting to increase by 0.03, but at a decreasing rate. By taking the derivatives of the model with respect to (log) income, we can find the income value where the positive relationship between income on voter turnout ends. The derivative yields a (log) income value of 13.76, which corresponds to a family income value of $947,391, after which a negative impact of increasing income on likelihood to vote is expected.

For model 2, all income tiers have a statistically significant result at the 1% level except for tiers 7 and 8, whose coefficients are statistically insignificant at the 5% level. The lack of significance means the results of those tiers are inconclusive, though they will be discussed nonetheless. All nine of the income tiers have negative coefficients indicating observations within these tiers are all less likely to vote than individuals with a family income over $200,000. Income tiers 1, 2, and 9 have the greatest magnitudes of their coefficients. This indicates that switching from a person with a family income of over $200,000 to income tiers 1, 2, or 9 would reduce likelihood to vote by 0.0268, 0.0215, or 0.0293, respectively. This means tier 9 has the lowest likelihood of voting. Recall from Figure 23 that there are no non-voting observations in this tier which makes this quite an usual result. It is possible that the demographics of income tier 9 and observations with family income over $200,000 are quite similar so the inclusion of demographic controls are skewing the observed likelihood to vote. In the robustness checks, tier 9 will be removed to observe if the results are consistent. Income tiers 5, 6, 7, and 8 have the smallest magnitudes of their coefficients, which means they are the tiers with the closest likelihood to vote as individuals with over $200,000 for family income. Switching into tiers 5, 6, 7, or 8 results in a decrease in likelihood to vote of magnitude 0.00927, 0.00658, 0.00321, and 0.00289, respectively. Thus, tiers 7 and 8 have the highest likelihood of voting.

As the direction of the coefficients for all controls are the same for models 1 and 2, they will be discussed once for both models. Some categorical control variables are not arranged in progressive fashions, which makes interpretation of their coefficient direction and magnitude nonsensical. As such, they are not discussed here, but dummy variables for each subsection will be included in robustness checks to observe any interesting results

Age has a positive coefficient value of 0.00332, while age squared has a negative coefficient of 0.0000235. Again, these coefficient directions indicate that with an increase in age we would predict increased likelihood of voting, but this effect is diminished as age gets into higher values. By taking the derivative of the models with respect to age, the value where age's positive effect on likelihood to vote ends at 70.64. These results are quite logical as we expect higher turnout in more middle aged individuals as they tend to be more politically conscious and feel politics affect them more directly than youth. The negative impact of increasing age on likelihood to vote after age 70.64 is also logical as there become more barriers to voting such as immobility.

Education has a positive coefficients in both models, meaning educated individuals are predicted to vote at a higher percentage. That is, increasing education level by one subcategory (e.g., 2-year degree to 4-year degree) to increase likelihood of voting by 0.00395 and 0.00405 in models 1 and 2, respectively. This result shows that education level, which can vary largely across districts, can be a rather significant influence over voter turnout due to the coefficient magnitude and the number of subcategories in this variable (6). That is, if one district's average highest level of education is a high school diploma and another is a 4-year degree, model 2 predicts the latter should be more likely to vote with a magnitude of 0.01215. Gender is coded with female as 1 and male as 0. As a result, its negative coefficient values of 0.00959 in model 1 and 0.00956 in model 2 indicates that, ceteris paribus, males are expected to be more likely to vote than females.

The religion importance variable is coded on a scale of "very important" to "not important" so higher values indicate an individual finds religion less important. Therefore, the coefficient of that variable indicates when a person cares less about religion by one importance classification model 1 predicts they are less likely to vote by 0.00795, and model 2 predicts they are less likely to vote by 0.00788. These coefficient values are fairly large in comparison to other control coefficients. If most individuals of a district rank religion as "not important" and another district ranks religion as "very important," the latter would be predicted by model 1 to vote at a higher rate by a rather significant magnitude of .02385. As the value of party increases, it indicates moving away from strong Democrat and toward strong Republican by subcategory values. So the coefficient demonstrates that as people identify more strongly with Republicans by one subcategory, both models expect them to turn out at lower rates by magnitude 0.00277.

With an increase of the (log) average income of a district by one, models 1 and 2 would predict an increased likelihood to vote of magnitude 0.0163 and 0.0169, respectively. These values are significant, but lower than the coefficient of the variable of interest in model 1, (log) income. This indicates that variation over individual family incomes have a larger impact on likelihood of voting than variation over the mean family income of that district. This is a logical result, as we would expect people to alter their behavior based on their individual situation more than the mean of their district.

The coefficients of district family count are quite small relative to the other coefficients with a value of 0.000000164 for model 1 and 0.000000162 for model 2. These small values are rational as the unit of increase is a single family, so adding one family to a congressional district would not be expected to have a large impact on much of anything. That said, variation in family size over districts can be quite large. The minimum value of the dataset is 80,208 whereas the maximum is 262,233. Comparing family counts for districts along the 75th and 25th percentiles would result in a difference of 22,157 families. An increase in family count of this scale would be expected by model 1 to increase likelihood of voting by a magnitude of 0.0036, which remains rather small.

The negative coefficient values of magnitude 0.107 and 0.106 for gini coefficient for models 1 and 2, respectively, indicate that increasing gini coefficient by 1 reduces likelihood to vote by those magnitudes. As gini coefficient operates on a scale from 0 to 1, these magnitudes lack meaning. They can be scaled for a more accurate view. That is, with an increase of gini coefficient by 0.1, models 1 and 2 would predict a reduction in likelihood to vote of 0.0107 and 0.0106, respectively. From Table 1 it can be observed that gini coefficient values range from 0.365 to 0.612, so a difference of 0.2 in gini coefficient between districts is quite plausible. Such a difference would be associated with a roughly 0.0212 magnitude reduction in likelihood of voting which is a considerable impact.

The control variable for year had coefficients of 0.0171 for model 1 and 0.0172 for model 2, demonstrating that likelihood of voting is increasing by those magnitudes each year. It is very plausible that likelihood of voting and year do not have a strictly linear relationship, so dummy variables for year are included in robustness check regressions.

After observing these results, which indicate a statistically significant relationship between both income and likelihood of voting ands between gini coefficient and likelihood of voting, we remain interested to observe the wealth level of representatives in districts with varying levels of inequality. The unit of observation for this model is U.S. congressional districts.

Variable	**Model 3**	**Model 3**
Outcome metric:	(Log) Representative net worth scaled by mean family	(Log) Representative net worth scaled by median

	income	family income
Gini coefficient	0.815 (1.014)	2.976** (1.008)
Party (1 = R)	0.314*** (0.0625)	0.316*** (0.0625)
Year	0.0288* (0.0123)	0.0272* (0.0123)
Constant	-55.88* (24.56)	-53.42* (24.55)
N	3,449	3,449
Adjusted R-squared	0.009	0.010

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

It can be observed that the adjusted R-squared values are extremely low, which was anticipated due to the lack of sufficient controls. Such controls were discussed extensively the data section and their variation would likely explain a good portion of the variation in (log) scaled representative net worth.

Gini coefficient is seen to have a positive coefficient when scaling by mean and median family income. That said, when scaling by mean, the coefficient is statistically insignificant at the 5% level, while scaling by median results in a significant coefficient of gini coefficient at the 1% level. The significant coefficient indicates that model 3 predicts increasing gini coefficient by 0.1 to be related with an increase in (log) net worth of representative scaled by median of magnitude 0.2976. To understand this impact in more consumable terms, we can observe the expected impact of increasing gini coefficient on the mean value for (log) net worth of representative scaled by median of 2.785. The model predicts with that a 0.2 increase in gini coefficient the net worth of a representative to increase from 16.20 times median family income to 29.38 times, which is a very substantial increase.

The slightly higher adjusted R-squared value when scaling by median family income rather than mean family income indicates that variation in gini coefficient can account for variation in representative net worth slightly better when using that scaling metric. As median family income is lower than mean family income, as seen in Figure 4, this result likely indicates that in districts with higher inequality, the representatives elected are wealthier, but not so much so. To increase net worth scaled by mean family income by 1 requires a larger increase in net worth than the same increase in net worth scaled by median family income. Therefore, if gini coefficient had an incredibly large positive impact on net worth of representatives, scaling by mean family income would be more likely to be statistically significant. So with an increase in inequality in a district, we expect net worth of a representative to increase, but not to a massive degree.

The coefficient of the party control variable is statistically significant at the 0.01% level and has positive coefficients of 0.314 and 0.316 when scaling representative net worth by mean and median, respectively. This variable is coded as an indicator for Republican. This means that the model predicts Democratic representatives to have lower net worth than their Republican counterparts when scaling by either mean or median family income.

The coefficient for year is significant at the 5% level with a value of 0.0272 when scaling (log) net worth of representative by median family income. This is in accordance with the overall trend of increasing net worth of politicians over time as seen in Figure 3, but further that the net worth of representatives scaled by their constituent's median family income is also increasing over time. Again, this relationship may not be purely linear and dummy variables for year will be included in robustness checks.

These results are consistent when lagging gini coefficient behind net worth of representative by both one and two years. The magnitude of the coefficient of gini coefficient is largest when lagging one year, followed closely by the direct year. When scaling (log) net worth of representative by median family income, the adjusted R-squared value with one and two years of lag is 0.011, which is very slightly higher than the value for the direct year.

With the significance of model 3's results, the first half of the hypothesized feedback loop between income inequality and political inequality is supported by the data. The second half, in which unequal political representation leads to further income inequality will be tested by the following model 4. Once again, the unit of observation is U.S. congressional districts.

Variable	**Model 4**	**Model 4**
Outcome metric:	Gini coefficient	Gini coefficient
(Log) Representative net worth scaled by mean family income	0.000252 (0.000313)	—
(Log) Representative net worth scaled by median family income	—	0.000919** (0.000308)
Party (1 = R)	-0.0133*** (0.00108)	-0.0135*** (0.00108)
Year	0.00224***	0.00221***

	(0.000211)	(0.000211)
Constant	-4.035*** (0.423)	-3.990*** (0.423)
N	3,449	3,449
Adjusted R-squared	0.069	0.071

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

The adjusted R-squared values are low for model 4 when scaling representative net worth by either mean or median family income. This is an expected result due to the lack of controls. The variables expected to explain much of the variation in gini coefficient were discussed in detail within the data section. Note that the R-squared values are higher in this model than in model 3. This indicates that the lacking controls account for more of the variation in net worth scaled than the variation in gini coefficient.

Similar to the results for model 3, when scaling representative net worth by mean family income its coefficient is statistically insignificant result at a 0.05 alpha level. At the 1% level, there is a statistically significant positive coefficient for net worth when scaling by median family income. Increasing (log) net worth of representative scaled by median income by 1 is expected by model 4 to positively impact gini coefficient by 0.000919. positive direction indicates that when scaled net worth of a representative is increased, the model predicts an increase in gini coefficient. That is, as a representative gets wealthier, we expect their district to have higher levels of income inequality.

Once again, the significance of the coefficient of net worth when scaling by median but not for mean indicates that wealthier representatives predict higher gini coefficients, but massive positive changes between the net worth of representatives are not required for this result. As it takes a smaller increase of net worth to increase the median scaled variable than the mean scaled variable, we predict smaller increases in net worth to relate to increases in gini coefficient. There can be quite large differences in the net worth of representatives caused by variations in omitted variables. The absence of these variables might be why we are seeing smaller increases be statistically significant. That is, smaller differences in net worth may compare representatives that are far more similar in terms of the absent control variables whereas scaling my mean may be comparing representatives that are more dramatically distinct.

The 0.1% level statistically significant negative coefficient of the party control variable suggests that, holding all else constant, the model predicts switching from a Republican representative to a Democratic one increases the gini coefficient of the represented district by 0.0135. The coefficients for year are also statistically significant at the 0.1% level. When scaling representative net worth by median, increasing year by one anticipates a 0.00221 increase in gini coefficient. This is consistent with the trend of increasing income inequality that can be observed in Figure 5.

As above, I re-estimate these models by using lagged values of net worth (both one and two year lags) to capture the possibility that it takes some time for policy preferences of representatives to be revealed in the aggregate district level statistics. As above, the results using the lags are very similar to the models using contemporaneous values. The coefficient of (log) net worth scaled is similar in each regression, but largest in when lagging by two years with a coefficient of 0.000926 for (log) net worth of representative scaled by median family income. This result is logical as policy creation and implementation are both time consuming processes, and seeing the impact of those policies will take even longer. Thus, the significance, direction, and magnitude of the lagged coefficients lend support for political inequality leading to further economic inequality. The unequal representation in politics The R-squared values are highest when lagging by one year, but only slightly higher than those of the direct year.

Section 6: Evaluation

Note that all of the regressions throughout this paper were run with the robust command to correct for heteroskedasticity in the standard error terms.

Checks for outliers were conducted for variables of potential concern such as net worth, family income, and gini coefficient. While there were some potential outliers, the nature of these variables lend themselves to the existence of extreme observations. Further, various tests for outliers did not raise concern over these values. Dropping the top and bottom 5% of values for scaled net worth resulted in consistent results for both models 3 and 4.

Due to concerns about income tier 9's lack of non-voting observations, model 2 was tested with 8 income tiers rather than the 9 used in the original model. That is, the coefficients of each tier reflect that tier's likelihood of voting relative to observations with family income values of over $150,000. That income cutoff is reflective of the top 8.82% of a district's population on average. This resulted in statistically significant negative coefficients for tiers 1, 2, 3, and 4 at significance levels of 5% or lower, with tiers 1 and 2 having the largest magnitude. The coefficient of tier 5 was also negative, but not statistically significant. Tier 6 had a statistically insignificant positive coefficient. Tiers 7 and 8 had positive statistically significant coefficients at the 5% level. These results demonstrate that the lower income tiers are expected to vote at lower rates than individuals with family incomes of over $150,000, while the higher middle income tiers are expected to vote at higher rates. This is not consistent with model 2's initial results.

This observed difference is likely driven by the large negative coefficient of tier 9 in the initial results. Members of tier 9 were seen to be significantly less likely to vote than individuals with family incomes of over $250,000. This lower value for likelihood of voting was of significantly larger magnitude than for tiers 7 or 8. So while tiers 7 and 8 were less likely to than individuals with family income over $250,000, they were predicted as more likely to vote than individuals in tier 9. Therefore, the exclusion of tier 9 from the regression controls lumps those observations into the over $150,000 comparison category, causing the changing coefficient directions for tiers 7 and 8. While these results can be explained, the ease of shifting the regression results by merely fluctuating the cutoff line for tiers causes some concern. While the findings are not necessarily invalid, the somewhat arbitrary nature of the designated income tiers inhibits robust conclusions.

A control for state was added for models 1 and 2 and found to be statistically insignificant, meaning that the impact of varying state on expectation of voting is not differentiable from zero. With this addition, the R-squared value of each model remained unchanged. These results were somewhat surprising because of the fact that we might expect changing state to affect likelihood of voting, as a result of variations in factors such as voter registration and voter identification laws.

To further explore these variables, models 1 and 2 were run with state controls included as a full set of dummy variables. Most of the dummies are statistically significant at the 0.1%

level and the R-squared value of both models increase. This result implies that the initial statistical insignificance of state as a control is caused either by the lack of any linearity in these categorical variables or was thrown off by particularly insignificant states. The inclusion of state dummy variables have a minimal effects on most of the coefficients of the other controls of models 1 and 2. That said, a few impacts were more substantial. The coefficient of marital status saw increased magnitude and became significant at the 1% level, whereas it was previously only significant at the 5% level. The negative coefficient of religion importance had a larger magnitude, indicating an even more impactful relationship between stated religious importance and voter turnout. Most interestingly, with the inclusion of state control dummies, the coefficient of gini coefficient was no longer significant for both models 1 and 2. This implies that the previously observed significance was likely a result of differences between states, manifesting in variations of gini coefficient. The finding that higher income inequality leads to higher voter turnout is therefore not robust.

From the dummy variables for state we can observe that Alaska predicts the highest likelihood of voting, closely followed by Washington D.C. and Missouri. Alabama is seen to have the lowest likelihood of voting. Alaska and Washington D.C. having the highest likelihood of voting could be caused by the low number of observations for both, yet the state of Missouri has a large number of observations, which provides evidence to the contrary. This result could instead be a reflection of the relationship between population size and voter turnout. If this is the case, it would be counter to the results found for the coefficient of district family count in models 1 and 2.

Models 1 and 2 were run with dummy variables included for years to include the possibility of a not entirely linear relationship between year and likelihood of voting. With this inclusion in model 1, the coefficient of (log) income increases from 0.030 to 0.041 and (log) income squared reduces from -0.00109 to -0.0016779. This indicates a change in the income level where the positive relationship between income and likelihood of voting ends from \$947,391 to \$202,330, which is a rather dramatic difference. The coefficients of race and religion importance were no longer statistically significant, which is unexpected considering the previously impactful findings concerning religion importance. These variables were similarly insignificant when the year dummies were included in model 2. Furthermore, tiers 4, 5, 6, 7, and 8 have coefficients that are no longer statistically significant. Tier 9, which was previously seen to be one of the least likely to vote with a coefficient of -0.029 saw its magnitude reduced dramatically to -0.0087. For both models the coefficient of the year dummy variables were in a nearly perfect linear relationship, with the only exception of the positive coefficient of 2012 being very slightly larger than that of 2014, relative to 2006. With all of these dramatic changes, the R-squared values for both models increased. These results indicate that the findings are far from robust, as the coefficients change quite significantly when controlling dummies for years. This likely indicates the existence of data quality issues.

As 2012 and 2014 had no non-voting observations, models 1 and 2 were run with observations from those years dropped. Model 1 had statistically significant coefficients for log income and log income squared in the same direction, but with magnitudes of 0.101 and -0.00422, respectively. For model 2, income tiers 1 through 5 had negative coefficients, while tiers 6, 7, and 8 had positive. Tiers 5 and 6 were statistically insignificant while the rest were significant at the 5% level or lower. Tier 9 had no observations not in years 2012 or 2014, so had a coefficient of zero.

Substituting income tier values for (log) income and income tier squared for (log) income squared for model 1 sees a result consistent with the initial findings. Removing all controls from model 1 sees consistent coefficient direction and statistical significance. As expected, it is accompanied by a dramatic reduction in R-squared value. Performing the same action for model 2 sees negative coefficients for income tiers 1 through 6, and positive for tiers 7, 8, and 9. All tier coefficients are significant at the 5% level except tiers 4 and 6. The changes among the income tier coefficients are likely caused by omitted variable bias. Again, the R-squared value decreases significantly.

Models 3 and 4 were all run using values of household income instead of family income. The results are consistent with the findings previously reported when scaling representative net worth by either mean or median family income and when regressing in the direct year or with lag. When removing the controls for party and year, the results of all of these models are consistent. Though, as one would expect, the R-squared values of both models are somewhat lower after removing the party and year controls. Testing models 3 and 4 with dummy variables for year yields very similar results as the initial outcome. With slightly higher values for R-squared, the coefficients of gini coefficient in model 3 and representative net worth scaled are of nearly identical magnitude, direction, and significance. The coefficients of the year dummies indicate a nonlinear relationship between (log) representative net worth when scaled by mean or median. Relative to 2006, the coefficients of 2007, 2008, and 2011 all had negative coefficients, while the rest of years were positive. The coefficient of the year dummies in model 4 showed a far more linear relationship between gini coefficient and time, despite some slight nonlinear fluctuations.

From the variety of regressions tested in this section, it can be observed that this paper's results concerning voter turnout are highly volatile. Including or excluding different dummy variables for models 1 and 2 can dramatically alter the results in terms of significance, direction, and magnitude. This variability could be caused by insufficient data quality within the voter data, very impactful omitted variable bias, or a combination of both. If it was simply an issue over the cutoffs of income tier chosen, model 1 would not have the lack of robustness it does. Due to the result of these checks, the main findings of this paper concerning voter turnout and income bias in voter turnout are quite sensitive to changes and should be considered with caution. The results concerning the relationship between representative net worth and gini coefficient appear to be more robust. Due to the lack of controls, there are less checks to implement, but with those conducted, the results appear rather consistent.

Section 7: Conclusion

Despite concerns over lack of robustness in models 1 and 2, their initial results will be taken in their current form in order to investigate how they fit into existing theories. That said, the findings are not conclusive and are highly subject to change. As such, these results should be considered plausible rather than definitive.

As discussed in Section 5, the negative coefficient on gini coefficient in models 1 and 2 demonstrate that higher levels of income inequality are expected to result in lower levels of voter turnout. The coefficients of the all income tiers in model 2 indicate that they are less likely to vote than the wealthiest individuals in a district who have family income values of over $200,000. These results are interesting in that they allow us to investigate how various theories of inequality and voter turnout seem to be supported or contradicted by these data sources. The overall lower turnout in more unequal districts is consistent with the expectation of power theory and contradictory with that of conflict theory. That said, conflict theory expects issues presented by the candidates to diverge to a more pro-poor and pro-wealthy divide. This would make the value of a vote increase for both the poorest and wealthiest income tiers, indicating that voter turnout should increase in both tiers. While the results do indicate high voter turnout among the wealthiest individuals, the poorest income tiers (1 and 2) were among the least likely to vote overall. This outcome seems to be more aligned with power theory. The negative expectation on voter turnout when increasing income inequality is consistent with the prediction of power theory. Furthermore, the lower voter turnout in the poorer income tiers can be linked to perceived lack of power and subsequent disengagement as the theory postulates.

The Schattschneider hypothesis suggests the wealthy will vote at higher rates than the poor because their issues make it on the political agendas of candidates. This hypothesis appears to be supported by the results of model 2. The lowest rates of voting in the poorest income tiers and the highest voter turnout among the wealthiest individuals of a district suggests this dynamic holds under the test of this data. Resource theory is plausible under the present results, but not likely. If the barrier to voting was a fixed financial cost, we would expect to see turnout rates be similar for all income tiers above a certain level. While it is possible that the barrier after which one would see equal turnout rates is after a family income of $200,000. That is unlikely as that would be quite a large fixed cost of voting. The results indicate the top income tier all vote at a lower rate than the wealthiest individuals, and various top tiers significantly more likely than others. Thus, resource theory does not seem to be well supported by this data.

With the support of the data, the forces at play can be hypothesized. The median voter theorem postulates that politicians will cater to the median voter's issues. If this were the case, one would expect the middle income tiers have the most to win or lose over the small differences in candidates; whereas, the extremes on both ends of the income spectrum likely feel either candidate affects them similarly. As the higher middle income tiers are the most likely to vote after the most wealthy, this is technically plausible, but less likely than other hypotheses.

Conflict theory posits a greater political clash between the poor and the wealthy indicating candidates represent more extreme interests on both sides. Yet, the low voter turnout among the lowest income tiers suggest this is either not the case or there are more significant burdens to voting that outweigh the increased perceived value of their vote. The theorem that seems to be best supported by the data is the Schattschneider hypothesis, which argues that wealthy interests dominate the political agenda. This would explain the high voter turnout among the wealthiest individuals and the relative disengagement of the other income tiers.

As discussed in the results section, the coefficient of district family count is positive, meaning in districts with more families we expect higher voter turnout. This is contradictory to what is expected from the existing literature. The relationship between small nations and high voter turnout are generally based on the idea that each person's vote has greater influence. This expectation is based on comparisons between the sizes of nations, but when comparing the family count of congressional districts the opposite is found. The average family size is not taken into account and likely varies over congressional districts. As a result, controlling for the raw population of a district would be an interesting and useful addition.

The findings from models 3 and 4 indicate that wealthier individuals are expected to represent less equal districts and less equal districts are expected to have wealthier representatives. These results reveal a positive relationship between income inequality and net worth of representatives. The significance of the models run with both 1 and 2 years of lag, suggesting a self-perpetuating and compounding relationship between the two variables.

As research is a process of continual improvement, it is useful to look at ways this paper can be improved and expanded upon. Adequate control data should be found and included for models 3 and 4 to ensure legitimacy of results. Some of these desired controls were discussed in the data section of this paper, but those are not comprehensive. Including controls for variables that affect representative net worth and gini coefficient will support the model's reliability and create more conclusive results. Similarly, the low R-squared values in models 1 and 2 indicate that the inclusion of more controls would greatly benefit their reliability. The variability in the robustness checks compound this need.

Additionally, there is significant room for improvement in data quality. For the voter data, additional observations would be ideal for non-voters. The dataset is predominantly individuals who voted in their congressional election, resulting in little variation in the voted indicator variable. To best capture effects on likelihood of voting, more variation in the outcome variable would provide more opportunity for the regressors to explain likelihood of voting, resulting in more robust results. This is a plausible cause for the lack of robustness observed in the evaluation section. For the results of this paper to hold weight, data quality and, subsequently, result robustness must be improved.

In order to establish more robust results, income tiers should be converted into deciles, or even smaller subdivisions, so the cutoffs between income groups are less arbitrary. That does not guarantee similar issues with inclusion and exclusion of observations as seen in the evaluation section will not arise. Though, it would allow the cutoffs for inclusion to be more logically based on proportional income breakdowns.

In regards to the net worth of representatives data, the values used are calculated from the average of a reported range. This makes the results less accurate as they are mere approximations of net worth rather than true values. In addition, while there may be significant differences between the net worths of two representatives, their average values may be identical revealing less precise findings. The reporting in ranges also leads to some individuals having no variation in average net worth over the years despite their true net worth actually fluctuating. While perhaps it would be very challenging to get precise measures of representative net worth, that would improve the quality of this paper's results.

Further questions could be explored with the inclusion of data on the specific policies offered by candidates. While this data may prove difficult to acquire, its use can provide particularly illuminating results. Much of the existing research on voting in one's self-interest is quite dated. If the policies promised by candidates are recorded in the dataset, a measure of self-interest could be constructed based on those policies and the socioeconomic composition of an individual. Though not utilized in this paper, the CCES voter data includes which candidate each observation voted for. From this, a model could then be constructed to test with more recent data whether individuals vote for the candidate that better represents their self-interest or if they vote along other lines.

Furthermore, inclusion of this data could help explain the reasoning for the observed results. As previously discussed, likelihood of voting correlates with an individual's perception of their vote's value. The value of a vote is directly related to the policies offered by the candidate and how those policies affect the voter. Therefore, in combination with the investigation into voting and self-interest, it would be interesting to parse out the methodology of the various voter turnout theories. That is, whether in districts with higher inequality we see candidates catering to the median voter as median voter theory suggests, or all candidates cater to the wealthy as Schattschneider hypothesis posits, or they cater to the extremes as conflict theory posits. I have speculated based on the results of this paper's models the forces at play, but obtaining data for the actual policies offered by candidates would allow for a more extensive analysis into these various hypotheses and the interaction between income inequality and voting. This data would also be useful in determining the importance of a given election and how that relates to voter turnout.

Lastly, data on the policies supported by representatives once in office would be useful in observing the legitimacy of the feedback loop between net worth of representative and gini coefficient. The results of the lagged models in each direction seem to indicate the existence of a positive feedback loop. However, data on the actual policies supported could help parse out whether this hypothesis has policy-based support or is moreso reflective of already existing trends over time.

References

Allen, F., Greiner, R., & Wishart, D. (2017). CFM-ID Applied to CASMI 2014. *Current Metabolomics, 5*(1), 35-39. doi:10.2174/2213235x04666160620100011

Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth, 5*(1), 5-32. doi:10.1023/A:1009850119329

Bartels, L. M. (2008). *Unequal democracy: the political economy of the new Gilded Age*. Princeton, NJ: Princeton University Press.

Benabou, R. (2000). Unequal Societies: Income Distribution and the Social Contract. *American Economic Review, 90*(1), 96-129. doi:10.1257/aer.90.1.96

Blais, A. (2000). *To vote or not to vote?: the merits and limits of rational choice theory*. Pittsburgh: University of Pittsburgh Press.

Blais, A. (2006). What Affects Voter Turnout? *Annual Review of Political Science,9*(1), 111-125.
doi:10.1146/annurev.polisci.9.070204.105121

Blais, A., & Carty, K. (1990). Does proportional representation foster voter turnout? *European Journal of Political Research*

Blais, A. and Dobrzynska, A. (1998), *'Turnout in Electoral Democracies', European Journal of Political Research, 33(2): 239–61.*

Brockington, D. (2004), 'The Paradox of Proportional Representation: The Effect of Party Systems and Coalitions on Individuals' Electoral Participation', Political Studies, 52(3): 469–90.

Darmofal, D. (2005). *Socioeconomic Bias, Turnout Decline, and the Puzzle of Participation* (Unpublished master's thesis). University of South Carolina. Retrieved from http://people.cas.sc.edu/darmofal/TurnoutBiasPaper.pdf

Darmofal, D. (2010). Reexamining the Calculus of Voting. *Political Psychology,31*(2), 149-174. doi:10.1111/j.1467-9221.2009.00754.x

De Gregorio, J., & Lee, J. (2002). Education and Income Inequality: New Evidence From Cross-Country Data. *Review of Income and Wealth, 48*(3), 395-416. doi:10.1111/1475-4991.00060

Deininger, K., & Squire, L. (1996). A New Data Set Measuring Income Inequality. *The World Bank Economic Review,10*(3), 565-591. doi:10.1093/wber/10.3.565

Fawaz, F., Rahnama, M., & Valcarcel, V. J. (2014). A refinement of the relationship between economic growth and income inequality. *Applied Economics, 46*(27), 3351-3361. doi:10.1080/00036846.2014.929624

Fields, G. S. (2001). *Distribution and development: a new look at the developing world*. New York, N.Y. : Cambridge, Mass. ; London: Russell Sage foundation ; MIT press.

Forbes, K. J. (2000). A Reassessment of the Relationship Between Inequality and Growth. *American Economic Review, 90*(4), 869-887. doi:10.1257/aer.90.4.869

Fornos, C. A., Power, T. J., & Garand, J. C. (2004). Explaining Voter Turnout in Latin America, 1980 to 2000. *Comparative Political Studies, 37*(8), 909-940. doi:10.1177/0010414004267981

Franko, W., Tolbert, C. J., & Witko, C. (2013). Inequality, Self-Interest, and Public Support for “Robin Hood” Tax Policies. *Political Research Quarterly, 66*(4), 923-937. doi:10.1177/1065912913485441

Freeman, R. (2003). What, Me Vote? doi:10.3386/w9896

Gilens, M. (2012). *Affluence and influence economic inequality and political power in America*. Princeton, N.J: Princeton University Press.

Gilens, M. (2014). *Affluence and influence: economic inequality and political power in America*. Princeton: Princeton University Press. Retrieved from http://press.princeton.edu/titles/9836.html

Glomm, G., & Ravikumar, B. (1992). Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality. *Journal of Political Economy, 100*(4), 818-834. doi:10.1086/261841

Glomm, G., & Ravikumar, B. (2003). Public education and income inequality. *European Journal of Political Economy, 19*(2), 289-300. doi:10.1016/s0176-2680(02)00178-7

Goni, E., Lopez, J. H., & Serven, L. (2008). Fiscal Redistribution And Income Inequality In Latin America. *Policy Research Working Papers*. doi:10.1596/1813-9450-4487

Goodin, R., & Dryzek, J. (1980). Rational Participation: The Politics of Relative Power. *British Journal of Political Science, 10*(03), 273-292. doi:10.1017/s0007123400002209

Gregorio, J. D., & Lee, J. (2002). Education and Income Inequality: New Evidence From Cross-Country Data. *Review of Income and Wealth, 48*(3), 395-416. doi:10.1111/1475-4991.00060

Gupta, S., Davoodi, H., & Alonso-Terme, R. (2002). Does Corruption Affect Income Inequality and Poverty? *Economics of Governance, 3*(1), 23-45. doi:10.1007/s101010100039

Hermes, N. (2014). Does microfinance affect income inequality? *Applied Economics, 46*(9), 1021-1034. doi:10.1080/00036846.2013.864039

Hill, K. Q., & Leighley, J. E. (1992). The Policy Consequences of Class Bias in State Electorates. *American Journal of Political Science, 36*(2), 351. doi:10.2307/2111481

Jackman, R. and Miller, R.A. (1995), ‘Voter Turnout in the Industrial Democracies during the 1980s’, Comparative Political Studies, 27(4): 467–92.

Jacobs, L. R., & Page, B. I. (2005). Who Influences U.S. Foreign Policy? *American Political Science Review, 99*(01). doi:10.1017/s000305540505152x

Janowitz, M., & Segal, D. (1967). Social Cleavage and Party Affiliation: Germany, Great Britain,
and the United States. *American Journal of Sociology, 72*(6), 601-618. doi:10.1086/224398

Johnston, R., Matthews, J. S., & Bittner, A. (2007). Turnout and the party system in Canada, 1988–2004. *Electoral Studies, 26*(4), 735-745. doi:10.1016/j.electstud.2007.08.002

Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review,45.*

Lehoucq, F. and Wall, D.L. (2004), 'Explaining Voter Turnout Rates in New Democracies: Guatemala', Electoral Studies, 23(3): 485–500.

Leighley, J. E., & Nagler, J. (2007). Unions, Voter Turnout, and Class Bias in the U.S. Electorate,
1964–2004. *The Journal of Politics, 69*(2), 430-441. doi:10.1111/j.1468-2508.2007.00541.x

Leighley, J. E., & Nagler, J. (2014). *Who votes now?: demographics, issues, inequality, and turnout in the United States*. Princeton: Princeton University Pres.

Lister, M. (2007). Institutions, Inequality and Social Norms: Explaining Variations in Participation. *The British Journal of Politics and International Relations, 9*(1), 20-35. doi:10.1111/j.1467-856x.2007.00246.x

Long, N. E. (1952). Bureaucracy and Constitutionalism. *American Political Science Review, 46*(03), 808-818. doi:10.2307/1952286

Maddocks, E. (2010). [Review of the book *The Spirit Level: Why More Equal Societies Almost Always Do Better by Richard Wilkinson; Kate Pickett*]. *Canadian Public Policy, 36*(1), 133-135.

Matsubayashi, T. and Wu, J.D. (2012), 'Distributive Politics and Voter Turnout', Journal of Elections, Public Opinion and Parties, 22(2): 167–86.

Meier, K. J. (1975). Representative Bureaucracy: An Empirical Analysis. *The American Political Science Review, 69*(2), 526-542. Retrieved from http://www.jstor.org/stable/1959084

Meltzer, A. H., & Richard, S. F. (1981). A Rational Theory of the Size of Government. *Journal of Political Economy, 89*(5), 914-927. doi:10.1086/261013

Moene, K. O., & Wallerstein, M. (2001). Inequality, Social Insurance, and Redistribution. *The American Political Science Review, 95*(4), 859-874. Retrieved from http://www.jstor.org/stable/3117718?seq=1#page_scan_tab_contents

Norris, P. (2002), Electoral Engineering: Voting Rules and Political Behavior (New York: Cambridge University Press).

Radcliff, B., & Davis, P. (2000). Labor organization and electoral participation in industrial democracies. *American Journal of Political Science, 44*(1).

Reif, K. and Schmitt, H. (1980), 'Nine Second-order National Elections: A Conceptual Framework for the Analysis of European Election Results', European Journal of Political Research, 8(1): 3–44.

Schattschneider, E. (1960). *The Semisovereign People: A Realist's View of Democracy in America*.

Sears, D. O., Hensler, C. P., & Speer, L. K. (1979). Whites opposition to busing: Self-interest or symbolic politics. American Political Science Review, 73, 369–384.

Sears, D. O., Lau, R. R., Tyler, T. R., & Allen, H. M. (1980). Self-interest vs. symbolic politics in
policy attitudes and presidential voting. American Political Science Review, 74, 670–684.

Shields, T. G., & Goidel, R. K. (1997). Participation Rates, Socioeconomic Class Biases, and Congressional Elections: A Crossvalidation. *American Journal of Political Science, 41*(2), 683. doi:10.2307/2111783

Solt, F. (2010). Does Economic Inequality Depress Electoral Participation? Testing the Schattschneider Hypothesis. *Political Behavior, 32*(2), 285-301. doi:10.1007/s11109-010-9106-0

Stockemer, D. (2016). What Affects Voter Turnout? A Review Article/Meta-Analysis of Aggregate Research. *Government and Opposition,* 1-25. doi:10.1017/gov.2016.30

Sylwester, K. (2002). Can education expenditures reduce income inequality? *Economics of Education Review, 21*(1), 43-52. doi:10.1016/s0272-7757(00)00038-8

Tahoun, A., & Lent, L. V. (2013). Personal Wealth Interests of Politicians and Government Intervention in the Economy: The Bailout of the US Financial Sector. *SSRN Electronic Journal*. doi:10.2139/ssrn.1570219

Teixeira, R. A. (1992, October 01). The Disappearing American Voter.

Verba, S., Brady, H. E., & Schlozman, K. L. (2002). *Voice and equality: civic voluntarism in American politics*. Cambridge, MA: Harvard University Press. Retrieved from http://www.hup.harvard.edu/catalog.php?isbn=9780674942936

Weiner, T. S., & Eckland, B. K. (1979). Education and Political Party: The Effects of College or Social Class? *American Journal of Sociology, 84*(4), 911-928. doi:10.1086/226866

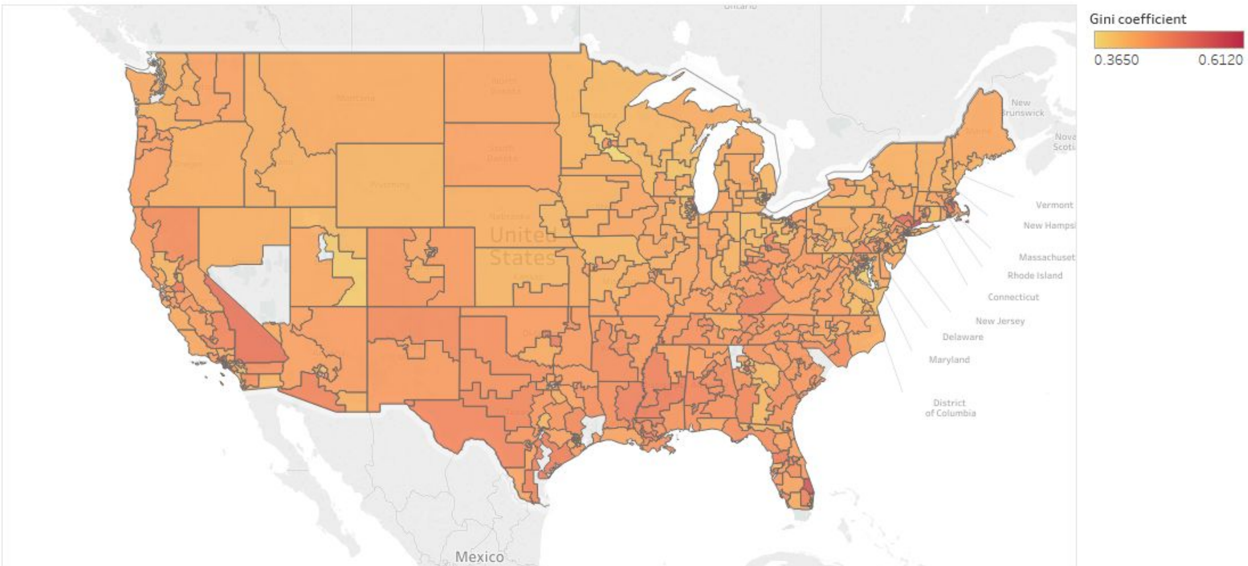
Wolfinger, R. E., & Rosenstone, S. J. (1980). *Who Votes?* New Haven (Conn.): Yale University Press.

Young, J., Borgida, E., Sullivan, J., & Aldrich, J. (1987). Personal Agendas and the Relationship Between Self-Interest and Voting Behavior. *Social Psychology Quarterly, 50*(1), 64. doi:10.2307/2786891

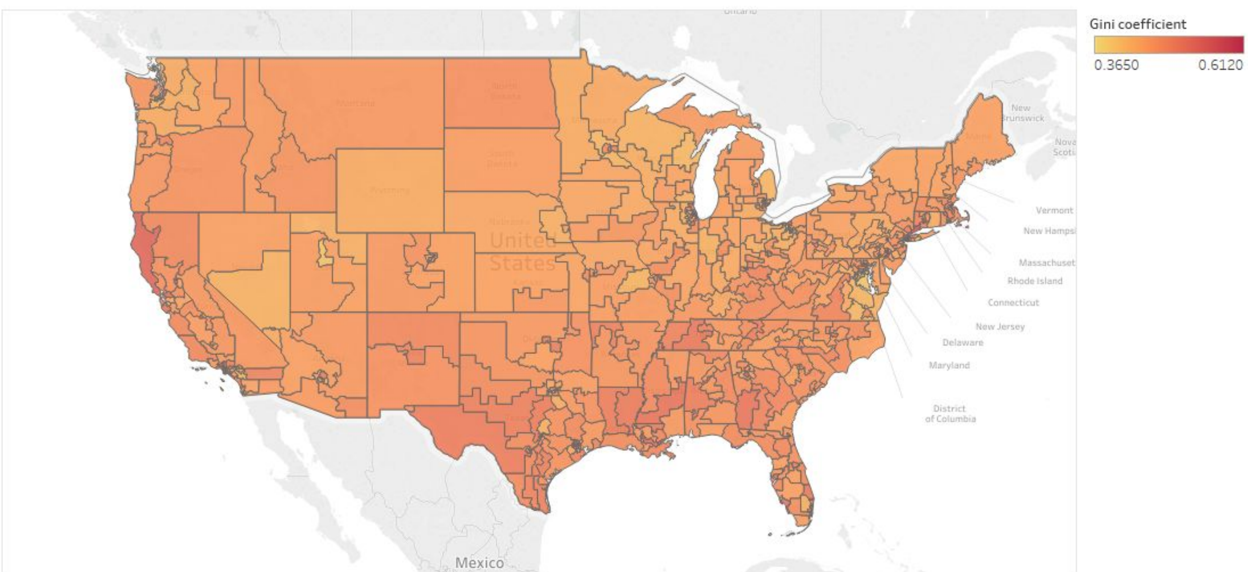
Appendix A:

Raw data:

Gini coefficient by congressional district - 2006

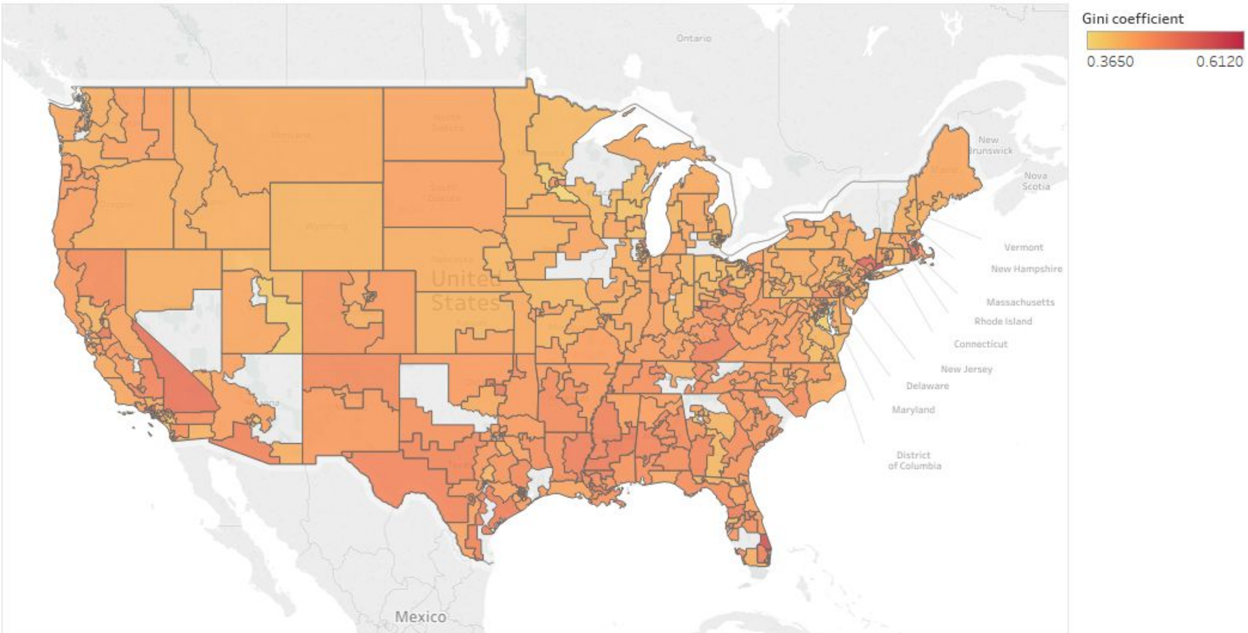


Gini coefficient by congressional district - 2014

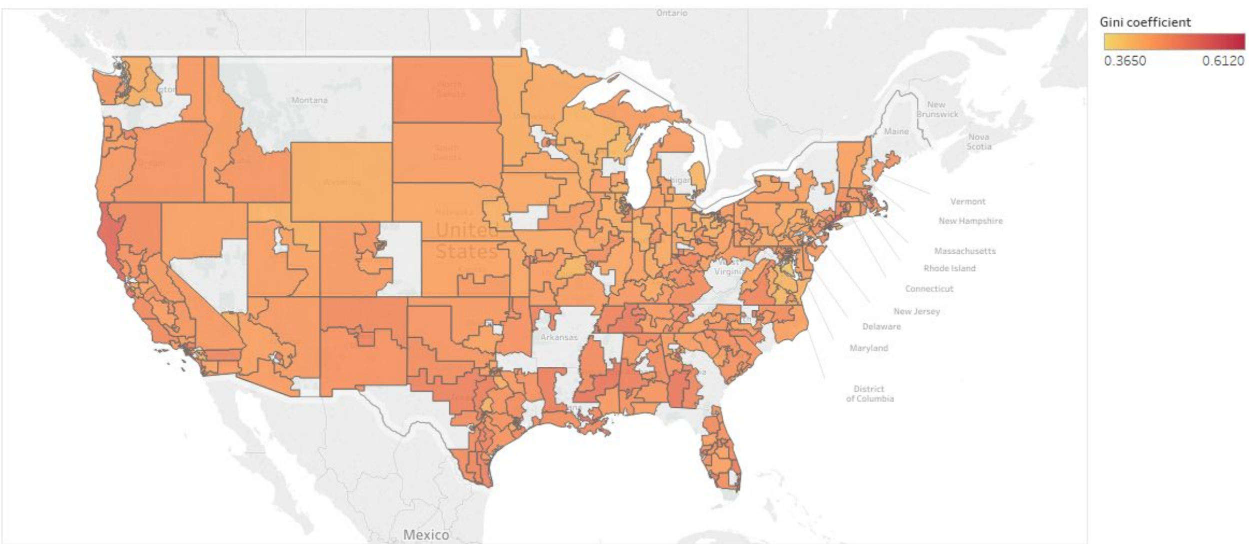


Merged data:

Gini coefficient by congressional district - 2006

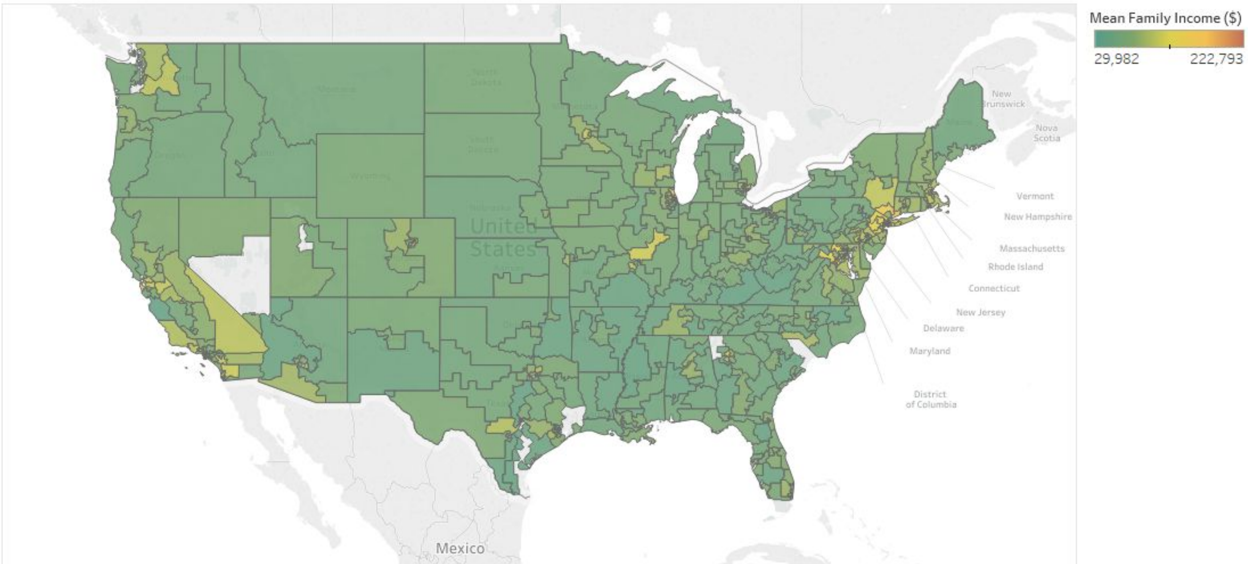


Gini coefficient by congressional district - 2014

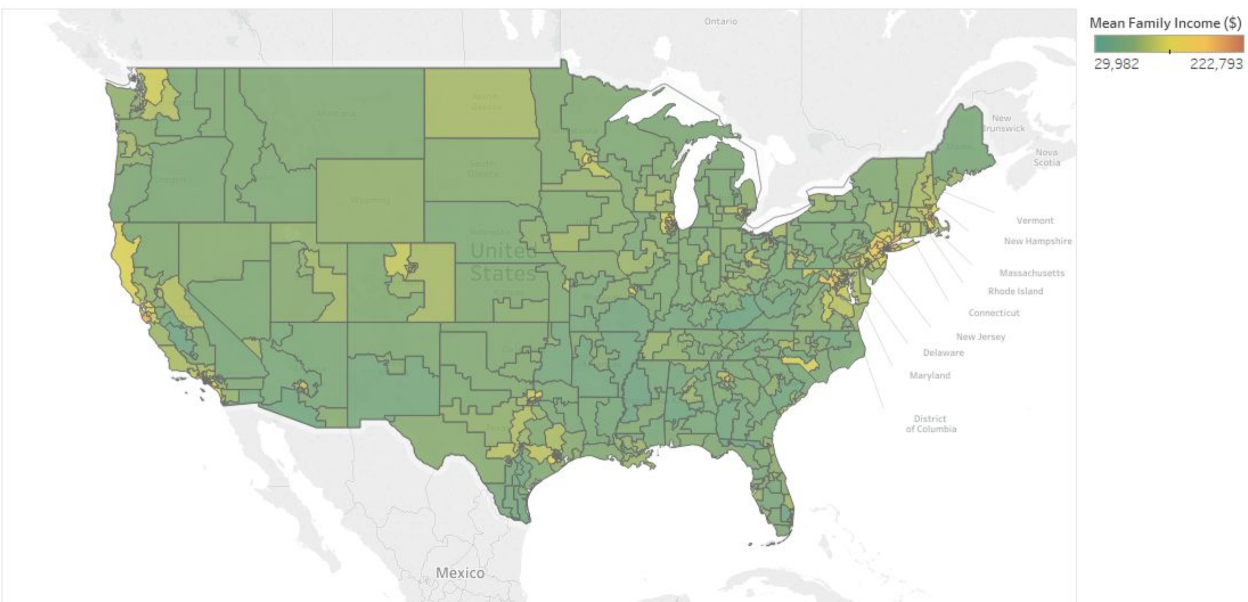


Raw data:

Mean family income by congressional district - 2006

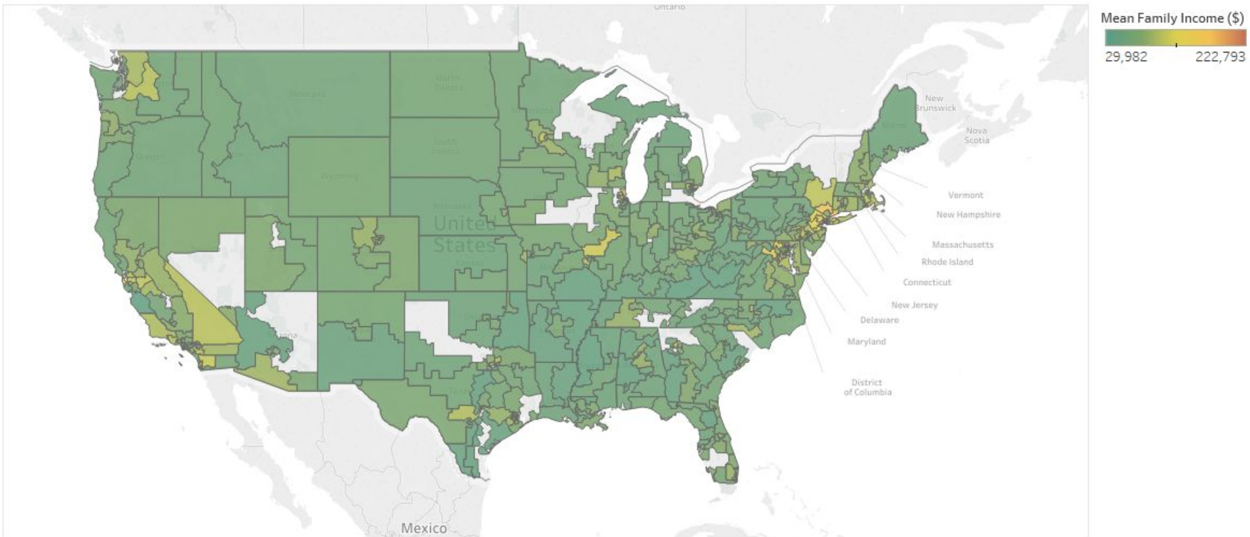


Mean family income by congressional district - 2014

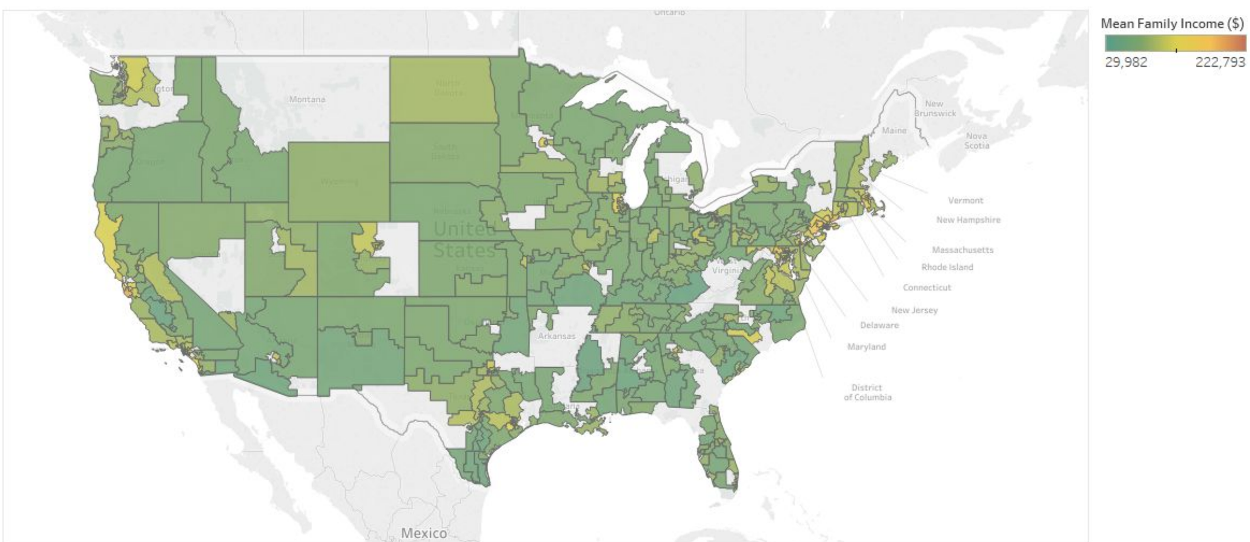


Merged data:

Mean family income by congressional district - 2006



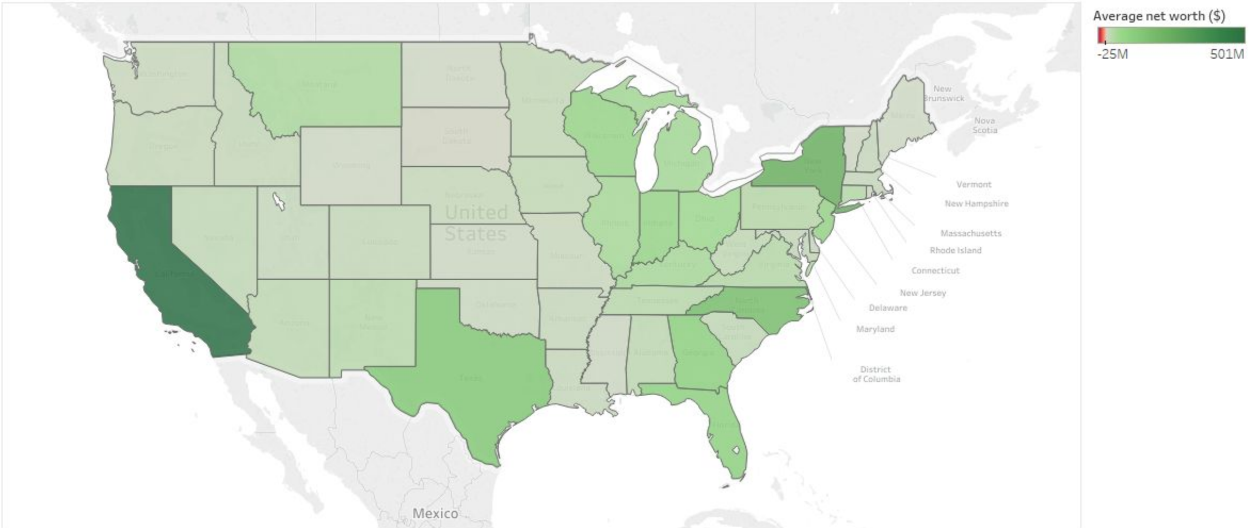
Mean family income by congressional district - 2014



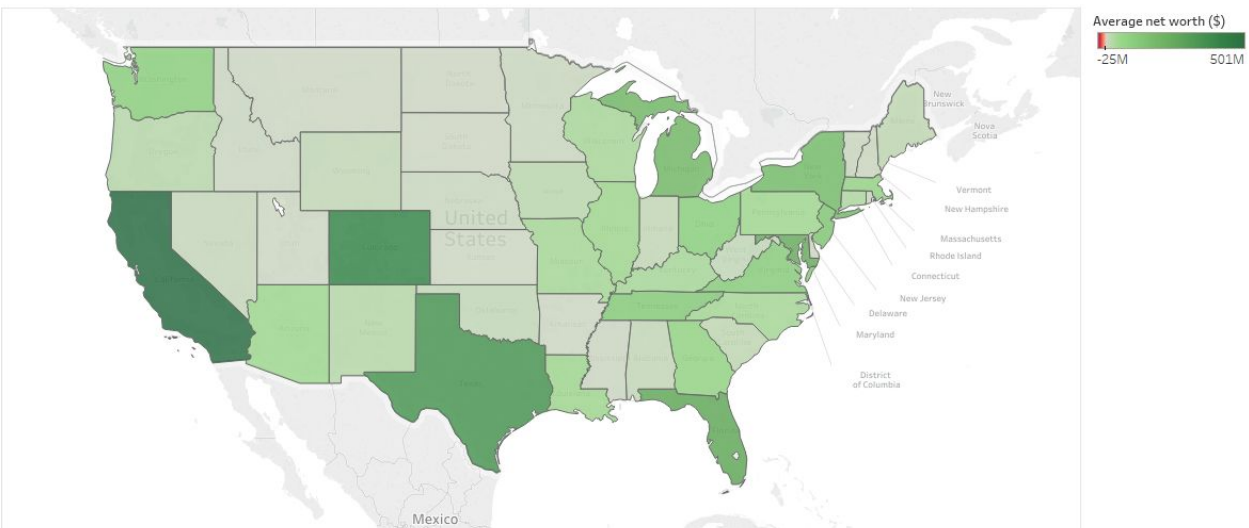
Raw data:

Note: The raw net worth dataset did not have a variable for congressional district. Therefore, the heat map for the raw net worth data is averaged over each state.

Net worth of U.S. Representative by state - 2004

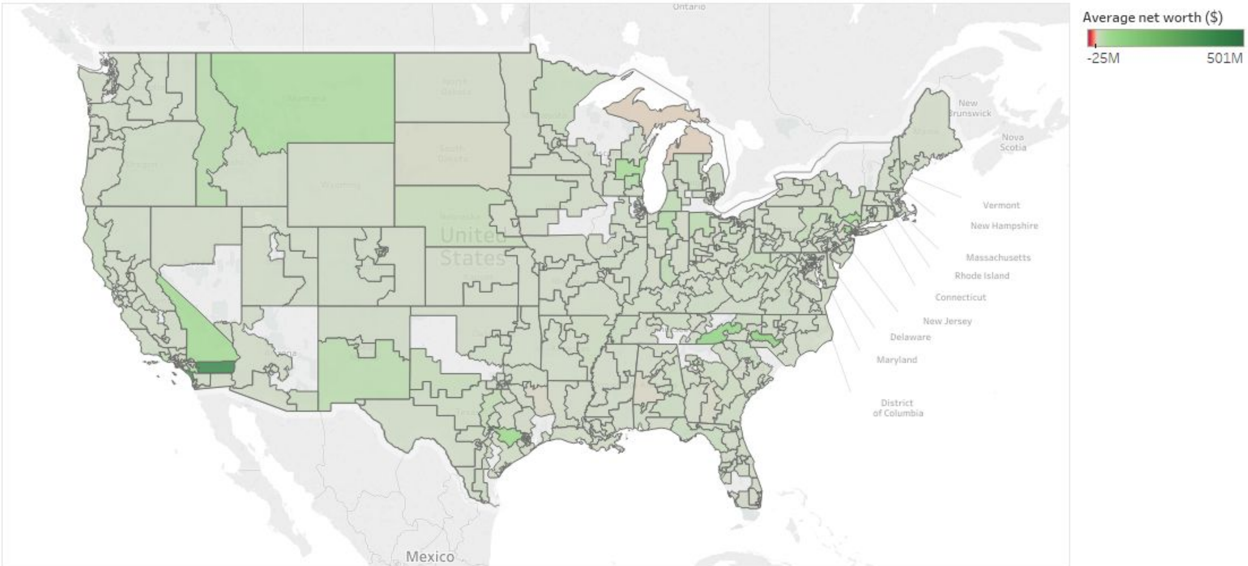


Net worth of U.S. Representative by state - 2014



Merged data:

Net worth of U.S. Representative by congressional district - 2006



Net worth of U.S. Representative by congressional district - 2014