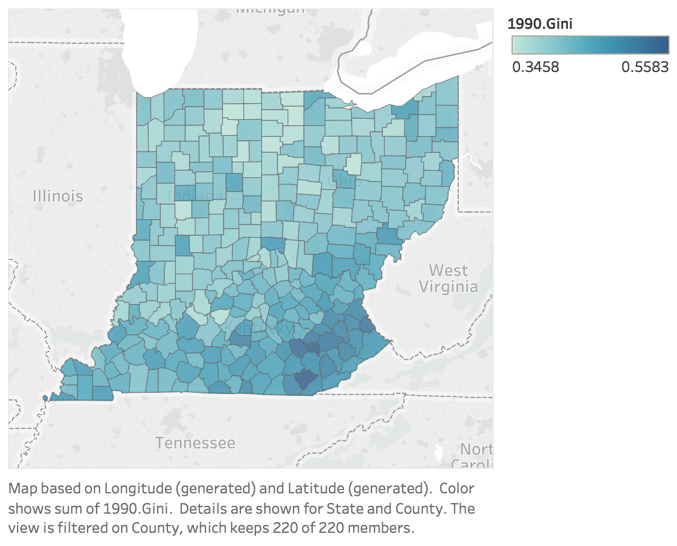
**Section I: Introduction and Literature Review**

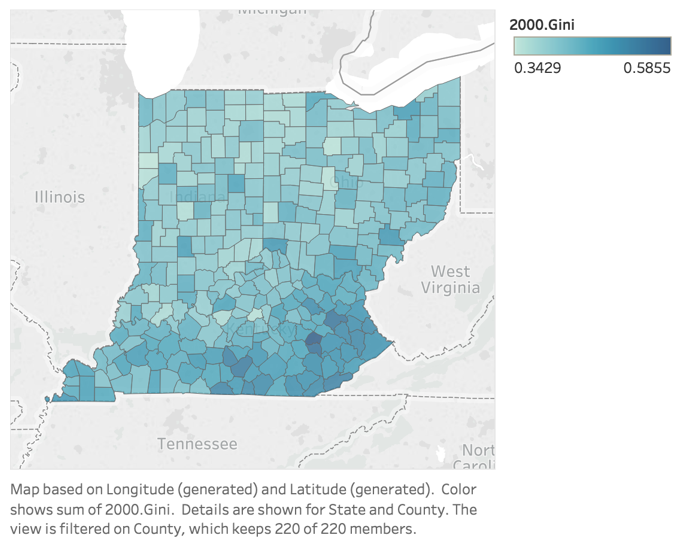
The topic of income inequality is one that has captured the attention and inflamed the passions of the general public throughout history, especially in recent years. The Occupy Wall Street movement of 2011 demonstrated the rage that many Americans felt about an economic system perceived as rigged to benefit the top 1% of wealthy Americans. Great and growing differences in earnings and wealth between members of society can incite resentment, fear, and premature death. Indeed, the case was made in the American Journal of Public Health by McDonough, et al, that low income and income instability can be linked to higher mortality rates (2011).

Public policymakers have every incentive to pay close attention to trends of and attitudes about extreme income disparities, as the consequences of rampant inequality can be a danger to society, whether locally, nationally, or globally. This paper will focus on causes of income inequality in the states of Ohio, Indiana, and Kentucky. We will examine the root causes of income inequality at the county level within these three states and model the impacts of various macroeconomic factors that have been postulated to drive variance in economic equality. Data will be derived from the Census Bureau’s American Community Survey, or ACS, for the periods 1990, 2000, and 2010. Figures 1, 2, and 3 show our dependent variable, the GINI coefficient, by county for each of these years.

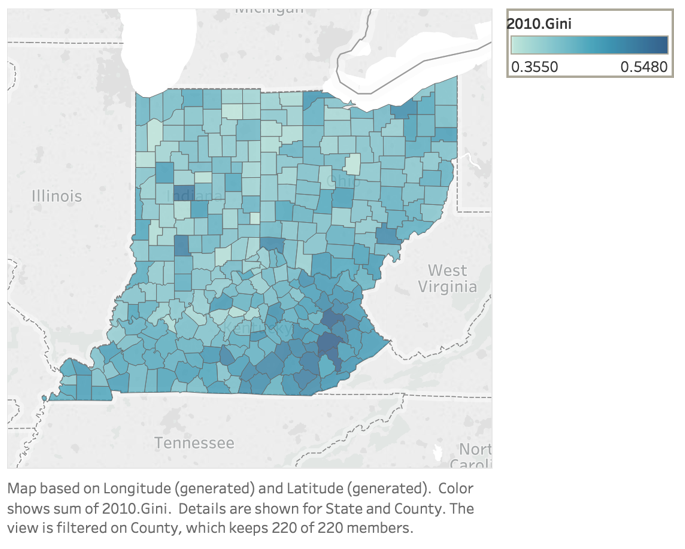
*Figure 1: GINI Coefficient by County, 1990*



*Figure 2: GINI Coefficient by County, 2000*



*Figure 3: GINI Coefficient by County, 2010*

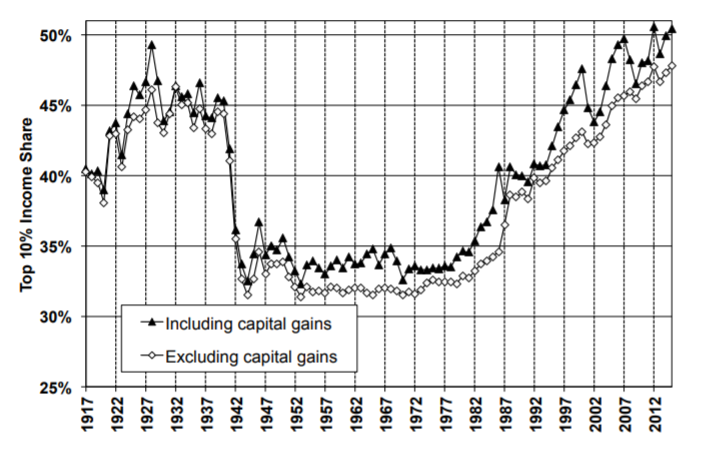


The GINI coefficient, developed by Italian statistician Corrado Gini in 1912, will function as our dependent variable. The coefficient, which ranges from zero (perfect equality in within a population) and one (all wealth or income in a population held by one individual). Using county-level data gleaned from the U.S. Census bureau and other sources, we will attempt to demonstrate the relationship between economic growth and income inequality.

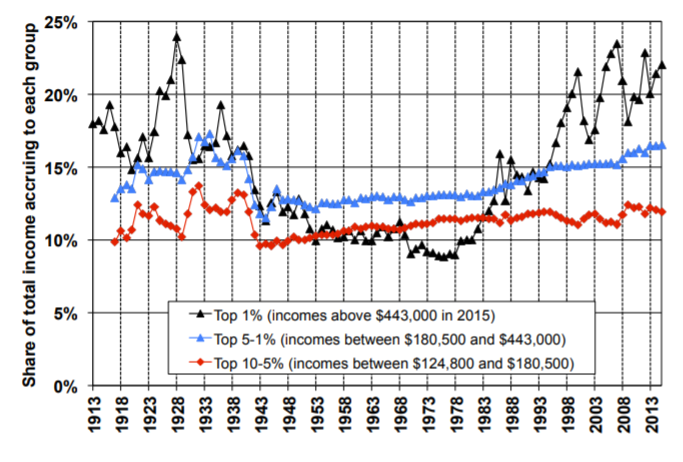
In the United States, the study of income inequality can be traced back to the early 20th century. Early works by Edwin Cannan’s “Division of Income” (1905) and H.J. Davenport’s “Value and Distribution” (1908) mark the earliest mentions of the issue of income distribution. For decades thereafter, most literature focused on methods of measuring inequality. Shortly after World War II, emphasis began to shift from simply measuring inequality to its causes. Simon Kuznets (1955) discussed the inherent difficulties of pinpointing the “character and causes of long term changes in the personal distribution of income” due to “looseness in definitions, unusual scarcity of data, and pressures of strongly held opinions” in “Economic Growth and Income Inequality.”

More recently, the focus has shifted to economic, political, and societal forces driving the mostly consistent growth in inequality over the last few decades. Nationally, a study out of UC Berkely (Saez, 2016) tracked the rise in income share of the top 10% and top 1% of incomes over nearly the last century. As can be seen in Figure 4 below, the top 10% of earners in the U.S. have never accounted for less than 30% of all income nationally. Further, the top earning decile recently accumulated as much as half of all income. Figure 5 splits out the top 5% and top 1% shares to further illustrate the growing increase in the rate of income dominance amongst the very wealthiest Americans.

*Figure 4: The Top Decile Income Share (1917-2015)*



*Figure 5: Decomposing the Top Decile US Income Share into 3 Groups (1913-2015)*

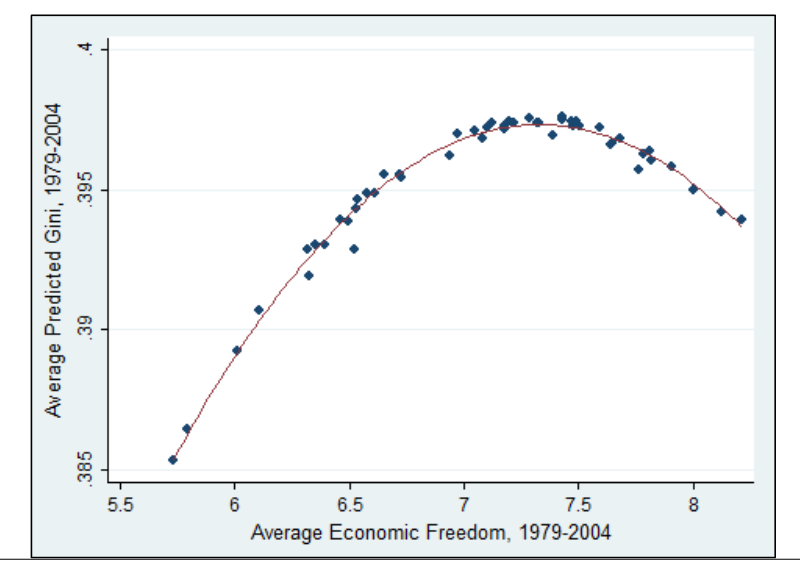


Various studies have localized the efforts to understand income inequality’s effects. The work of Brian Glassman of the U.S. Census Bureau (2017) used MSA data to analyze where along the income distribution inequality takes place. Within these metro areas, Glassman used several percentile ratios of income such as the 90-10 ratio, the 99-90 ratio, the 90-50 ratio, and the 50-10 ratio. It allowed him to discover that metro areas of varying sizes (small, medium, and large) tend to have larger ratios than others at differing income levels.

Levernier, Rickman, and Partridge (1995) examined economic data for the contiguous 48 states in the U.S. and studied inequality trends at a regional level. Key findings of their research were that some notable causes of inequality increase during the 1980’s were due to heightened international immigration and the rise of households headed by a single female. Additionally, their research supported arguments that factors such as greater high school attainment, labor force participation, goods-producing employment share, and transfer payments aided in reducing or mitigating the rise of income inequality.

The work of Bennett and Vedder (2013) also used state level data. Their research postulates that increasing degrees of economic freedom is associated with lower degrees of income inequality. Building on the work of Bergen (1999), Scully (2002), and Ashby and Sobel (2008), they argued that the relationship between economic freedom and inequality was represented by an inverse U-curve. At low levels of economic freedom, a marginal increase will yield higher inequality. However, as economic freedom continues to rise, marginal increases will begin to lower inequality. Figure 6 below illustrates this relationship.

*Figure 6: GINI coefficients as a function of Average Economic Freedom (1979-2004)*



**DATA:**

1. **Describe dependent variable – 2**
2. **Describe key x’s and explain why use these x’s (link to literature) – 2**
3. **Explain how you built your model (testing up, testing down, theory as a guide) – 2**
4. **Table with definition, source, mean, min, max, std dev. – 3**

The dependent variable, GINI*ij*, is a value calculated by the US Census that is specific to each county in the sample population. Where, *i*, represents the respective county in the analysis, and *j,* represents the respective census year for which the value was calculated. The analysis was composed of data collected from American Community Survey, collected by the US Census, which aggregates survey data from households in a certain geographic to confidently assign aggregate values for calculations regarding income inequality. The values are calculated using the varying income levels respective to their share of the population area (US Census).

Data was collected at a county-wide level for Indiana, Kentucky, and Ohio, which comprises some 300 independent county locales. All data was compiled by variable and by year, with the time and state variables being controlled by factored dummy time and dummy state variables. The observations were collected in 1990, 2000, and 2010, using the ACS 5-year estimate values, when available.

Some data was unobtainable from the US Census, which was supplemented from third parties. Data from places such as the Department of Agriculture, are indicated with an “^”, and is further detailed in the works cited information.

Descriptive Statistics Table **1**

                                                Will add in Word later

The explanatory variables, unemployment, per capita income, population, and percent of people over 25 with a high school degree (education), were common factors used in the OLS regression suggested by Levernier and Rickman in 2013 (amongst other variables). The percent of people below the federal poverty line, Poverty, was added to see how it would affect the regression results. The data was collected to a single datasheet that is provided at the end of this investigation with a data source page included.

In their analysis of a state by state income inequality, Levernier and Rickman discuss how multicollinearity affected their results. They note that single period regressions have fewer degrees of freedom by nature, and that further, measures that affect both measures that labor markets and demographic variables, like education, could affect each other. Controls for these interactions will be discussed in the empirical approach section.

The dataset was tested for several OLS assumptions to address issues such as multicollinearity, heteroskedasticity, and stationarity. It was hypothesized that the effects of time would have a correlation on results, and this was identified once graphing GINI vs. time, which is displayed in Figure **2**.

Literature addressed for issues arising between using single period regressions by adding dummy variables for time. Multicollinearity was addressed by adding dummy variables, and by adding  more defining demographic statistics regarding both population including ethnicity and gender. These data could possibly positively affect the validity of the results, and need to be investigated further. This regression took the rates of change for the entire, unspecified population.

**Empirical Approach**

1. **Why technique appropriate (fixed effects, spatial Durbin, etc.) – 2**
2. **Five problems test for and correct for – 2 each**

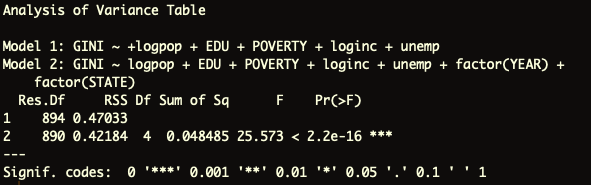
Fixed effects was supported by literature Levernier and Rickman

Fitted form was modified according to literature, which has called for taking the log of per capita income and population.

Saw that data had multicollinearity **1**  in several explanatory variables, and this problem is common across literature (), especially among education and unemployment values.

Tried to fix by adding fixed effects

        ANOVA Table **2**.



Fixed effects also tried to address the possibility of omitted variable bias **2**.

Marginal effects (dummies) Table **3.**

In order to address heteroskedasticity **3** and stationarity **4**, the variables were difference to an order of I(1), This was analyzed by the Breusch-Pagan test to address the degree to which this was effective at solving the problem

Differencing also addressed multicollinearity, which lowered VIF values for nearly all variables.

Endogeneity **5**?

The data looks to be clustered into years, so we will continue to investigate how this could be addressed by running different type of regressions. Spatial Dependent.

**Results**

**OLSTable 3.**

**Panel Table 4.**

The OLS regressions showed negative elasticities in the categories that were expected to decrease per unit GINI, namely, the percent of population over 25 with a high school degree, and unemployment. The negative effect on GINI from unemployment was supported by an investigation yielding similar results. (Bennet and Vedder, 2013)

1. **Elasticities or marginal effects of key x’s – 2**
2. **Compare importance of x’s – 2**
3. **Full and careful interpretation of an x (beta, marginal effect, or elasticity) – 3**

**Conclusions**

**Works Cited**  
McDonough, P., Duncan, G., Williams, D., & House, J. (1997). Income Dynamics and Adult Mortality in the United States, 1972 through 1989. Retrieved from <https://ajph.aphapublications.org/doi/pdf/10.2105/AJPH.87.9.1476>

Saez, E. (2016). Striking it Richer: The Evolution of Top Incomes in the United States. Retrieved from <https://eml.berkeley.edu/~saez/saez-UStopincomes-2015.pdf>

Kuznets, S. (1955). ECONOMIC GROWTH AND INCOME INEQUALITY. *American Economic Review*, *XLV*(1). Retrieved from <https://www.jstor.org/stable/pdf/1811581.pdf?refreqid=excelsior%3A5a1b7a97a79b79320f9062cb058a5a8e>

Glassman, B. (2017). Income Inequality among Regions and Metropolitan Statistical Areas: 2005 to 2015. Retrieved from <https://www.census.gov/content/dam/Census/library/working-papers/2017/demo/SEHSD-WP2017-41.pdf>

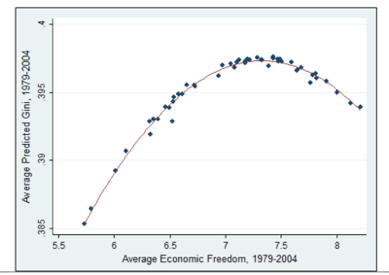
Levernier, W., Rickman, D., & Partridge, M. (1995). Variation in US. State Income Inequality: 1960-1990. *International Regional Science Review*, *18*(3), 355-378. Retrieved from <https://journals.sagepub.com/doi/pdf/10.1177/016001769501800305?casa_token=408gLj3fP-8AAAAA:j0EGJguJj_u2pd3HGQQhDDDRsYuqqf1JhnblR4eylm4_9Ifwqk-OtZARhD5jv-b-nti2wmYd5w>

Bennett, D., & Vedder, R. (2012). A Dynamic Analysis of Economic Freedom and Income Inequality in the 50 U.S. States: Empirical Evidence of a Parabolic Relationship. *SSRN Electronic Journal*, *43*(1), 42-55. doi: 10.2139/ssrn.2134650

[**https://www.census.gov/topics/income-poverty/income-inequality/about/metrics/gini-index.html**](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://www.census.gov/topics/income-poverty/income-inequality/about/metrics/gini-index.html%26amp;sa%3DD%26amp;ust%3D1550503094601000&sa=D&ust=1550503094638000&usg=AFQjCNH0Iom5zxhVSOOnj1oIxILeTjGFUw)

* [**https://journals.sagepub.com/doi/pdf/10.1177/016001769501800305**](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://journals.sagepub.com/doi/pdf/10.1177/016001769501800305%26amp;sa%3DD%26amp;ust%3D1550503094602000&sa=D&ust=1550503094638000&usg=AFQjCNGryfB_PaOXGdTQ2ZLCRWfnzoydOQ)
* [**https://ageconsearch.umn.edu/bitstream/243947/2/v43\_n1\_a5\_bennett\_vedder.pdf**](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://ageconsearch.umn.edu/bitstream/243947/2/v43_n1_a5_bennett_vedder.pdf%26amp;sa%3DD%26amp;ust%3D1550503094603000&sa=D&ust=1550503094638000&usg=AFQjCNE85QwqWx6n7W21_3RYbSPGVnseIQ)
* [**https://www.shsu.edu/eco\_mwf/ECIN%20-%20Frank%202009.pdf**](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://www.shsu.edu/eco_mwf/ECIN%252520-%252520Frank%2525202009.pdf%26amp;sa%3DD%26amp;ust%3D1550503094603000&sa=D&ust=1550503094639000&usg=AFQjCNH4GEOimQD3gqFYbUcII4IOaIrHQA)
* [**https://books.google.com/books?hl=en&lr=&id=fS4RDAAAQBAJ&oi=fnd&pg=PA280&dq=state+income+inequality+econometrics&ots=xiZ\_41k3LM&sig=B-8F0DYMe5UPEusTDe7PCxxKf\_M#v=onepage&q=state%20income%20inequality%20econometrics&f=false**](https://www.google.com/url?q=https://www.google.com/url?q%3Dhttps://books.google.com/books?hl%253Den%2526lr%253D%2526id%253DfS4RDAAAQBAJ%2526oi%253Dfnd%2526pg%253DPA280%2526dq%253Dstate%252Bincome%252Binequality%252Beconometrics%2526ots%253DxiZ_41k3LM%2526sig%253DB-8F0DYMe5UPEusTDe7PCxxKf_M%2523v%253Donepage%2526q%253Dstate%252520income%252520inequality%252520econometrics%2526f%253Dfalse%26amp;sa%3DD%26amp;ust%3D1550503094605000&sa=D&ust=1550503094639000&usg=AFQjCNGwA84vBzBkcSzyx8sYQrdzNTIKqQ)

**Data and Figures.**



*Figure 1: GINI coefficients as a function of Average Economic Freedom (1979-2004) SOURCE*

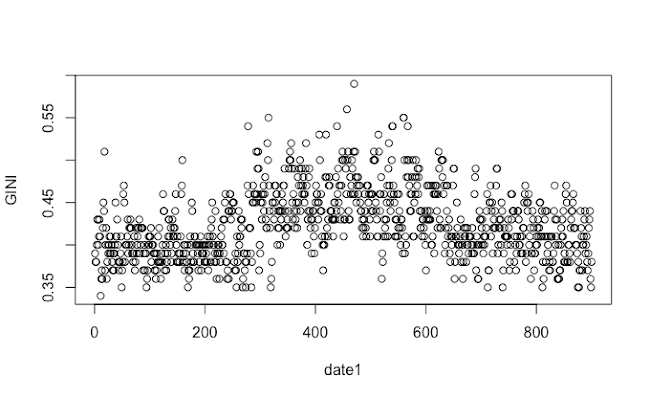
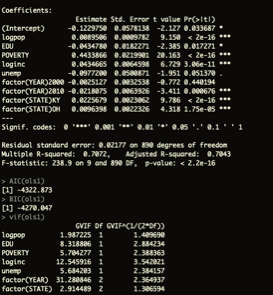


Figure **2**.

A plot of GINI versus time, where observations 1-300 represent values in 1990, 301-600 represent values in 2000, and 601-900 represent values in 2010. There are three visible cluster which represent nonstationarity and spatial dependence.

I(0) model



I(1) model

