

Median Income vs. Poverty:

New York, Texas and California

By: Luke Anderton, Mariana Gonzalez and Taylor Hill

Group ID: 172543

GEOG 312

I. Abstract

The objective of the research was to find a statistical relationship between median income and poverty levels. We chose New York, California and Texas as our three samples, comparing their annual median incomes and poverty levels from the years 1984 to 2017 and sourced our data from the US Census. Both the Pearson's correlation and the Student's Test were used to determine whether there was a statistical similarity between the two variables. In addition, we also found the spatial center for the median income in the US. Our results indicated that there was very little correlation between the two variables for each sample and our results were statistically insignificant. We were unable to reject the null hypothesis.

II. Introduction

Median income is a economic indicator surveyed by the US government on an annual basis. The data is collected by the American Community Survey (ACS) and the Puerto Rico Community Survey (PRCS). Both surveys are conducted annually by the US Census Bureau. A sample of 3.5 million housing addresses are interviewed over a twelve month period. The definition of household median income is income in the Past 12 Months, including "the income of the householder and all other individuals 15 years old and over in the household, whether they are related to the householder or not" (Median Household Income).

Poverty is an issue that the country is trying to solve. According to the U.S. Census Bureau, the U.S. median income increased by 1.8 percent and poverty rate decreased 0.4 percent between 2016 and 2017 (Income, Poverty, and Health Insurance Coverage in the United States:

2017). An alternative way they measure poverty is “Supplemental Poverty Measure (SPM)” ; it’s also an indicator of economic well-being. It was reported that in 2017, 39.7 million people, or 1 in 8 Americans, were in poverty (Income, Poverty, and Health Insurance Coverage in the United States: 2017).

According to the supplemental poverty measure, the most poverty stricken state in the U.S. is California. It’s more accurate than the traditional measure because it also considers living costs (Walters). Over 20 percent of Californians are living in poverty (Walters). Living costs for California are among one of the highest in the nation (Walters).

Texas’ poverty levels are lowering (Ura and Wang). Poverty peaked in Texas in 2011 at 18.5 percent and it dropped below the pre-recession levels in 2016 and kept dropping in 2017 (Ura and Wang).

About 40 percent of New Yorkers live in New York City (Demographics of New York (State)). A report that used data 2011-2015 stated that nearly 1.7 million people in New York lived in poverty (NYU Furman Center).

Our research questions seeks to find the relationship between median income an the poverty rate in a sample of states. The states of New York, California, and Texas are used as three samples, with data points being collected from the years 1984 to 2017. The objective of our research is to find correlation between median income and poverty, and use spatial analysis to find center of U.S. Median Income. The null hypothesis is that there is no correlation between median income and poverty.

III. Methodology

U.S. Census Data

The data was taken from the US Census website, "www.census.gov". On this page there is a "Browse by Topics" section in the menu, hovering over this a drop down menu appears with a list of census topics. When "Income and Poverty" is selected pulls up a page offering news, data, papers and publications on income and poverty in the United States. On the left hand side there is another menu with an option called "data", when hovering over this three options appear, the option "Income and Poverty Data Tables" is where the data for the research was found. Under the year 2016, the table "Historical Income Tables: Households" was selected, "Table H-8. Median Household Income by State" downloaded a CSV file with the median household income in each from the years 1984 to 2017. There were two median income values per state, the median income in past dollars and the median income adjusted for inflation. For this research, the current dollar value was used. The data points for California, New York and Texas were extracted for the correlation tests and the spatial tests.

The same process was conducted to find the poverty percentages with an adjustment to a few steps. Instead of selecting the year 2016, the year 2017 was selected, under this there is a table called "Historical Poverty Table: People and Families- 1959-2017". The data used from this group of tables was table 21 called "Number of Poor in Poverty, by State". This downloaded into a CSV file that was used to extract data from. The percentage of the state population in poverty was recorded for California, New York and Texas from the years 1984 to 2017.

Three separate CSV files were create, one for each sample. The CSV files contained the median income data, poverty percentage data and their corresponding year. This data was then used for the statistical analysis.

Statistical Analysis Methods

The goal for the analysis was to be able to confirm or rule out a positive relationship between median income and poverty in the U.S. by state, as well as locate the median income spacial center for the entire country. In accordance, we decided the best method to achieve that would be to use three different tests. The first being Pearson's correlation, which we used to attempt to find the value of the relationship between the two factors. The second, the Student's Test, would validate or otherwise our results, giving clairvoyance as to whether the correlation held any value. Finally for the last goal of finding the median income weighted center, we weighed each individual point with its location for the states for each year from 1984-2017, and calculated the center. All of the study was done in R, using data given by the US Census.

IV. Results

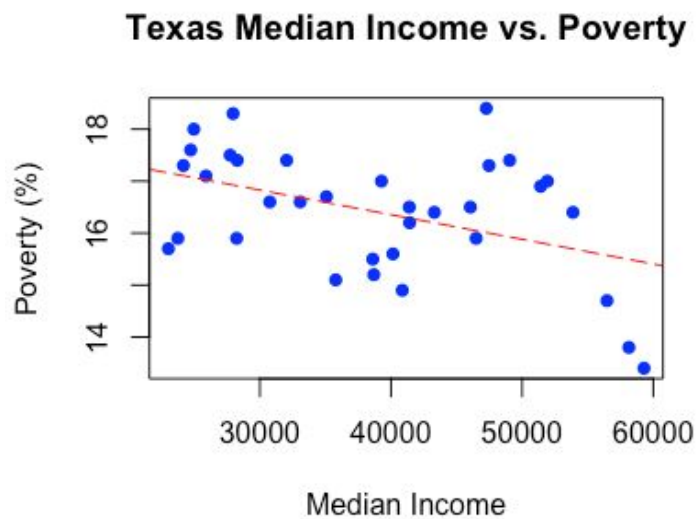
Median Income vs. Poverty

For this comparison, we used both the Pearson's correlation and the Student's Test to determine whether there was any significant statistical similarity between median income and poverty levels in states over time. Given the nature of the two's logical similarity, we assumed they would have a positive correlation at the beginning of this project. However, after completing both tests were were quite shocked to find that the two have very little in the way of

correlation. Firstly, Texas has the strongest correlation, with a negative Pearson's value given in chart (5). Also, it has the steepest slope of any of the three states (1). However, we can't successfully conclude there is a negative correlation between median income and poverty due to the t-score given by the student's test. The same is true for the other states as well. Despite their negative Pearson's correlation values (5), and negative slopes on the regression lines (2,3), they both have t-scores that are much too low to confirm the bizarre notion that median income has a negative effect on poverty rates.

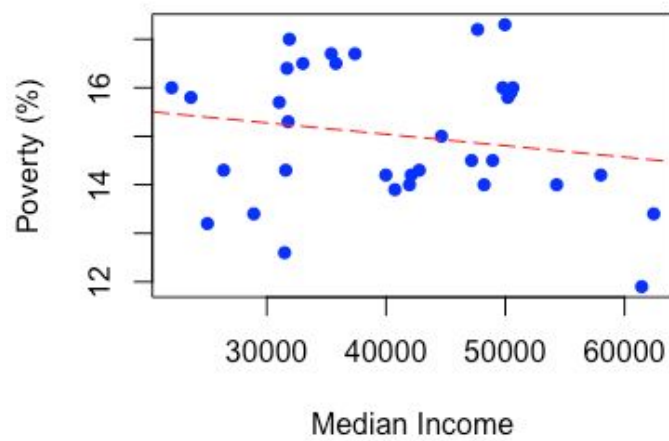
V. Figures and Tables

1



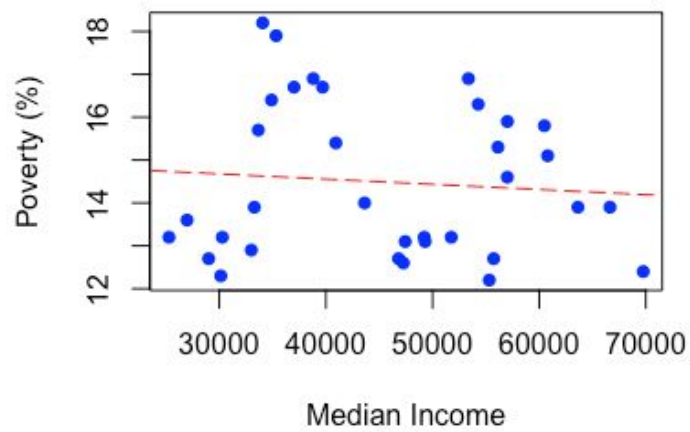
2

New York Median Income vs. Poverty

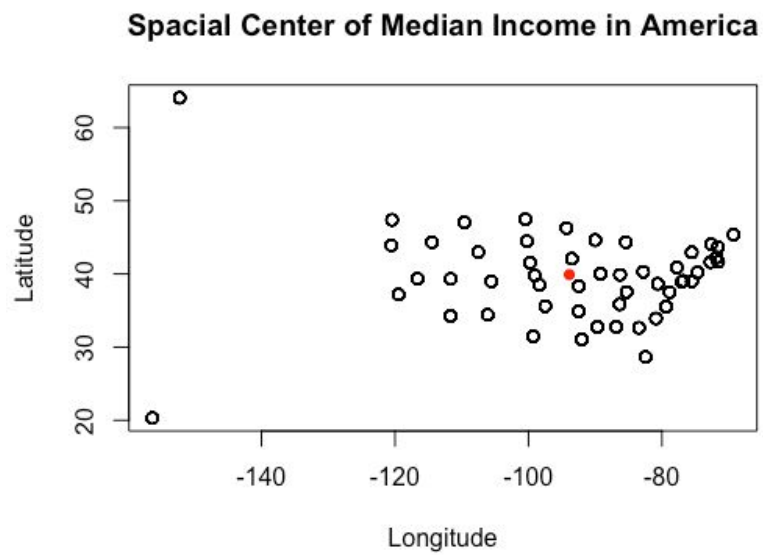


3

California Median Income vs. Poverty



4



5

State	Pearson's Correlation	Student's Test
California	$r = -0.002570496$	$t = -0.01454097$
New York	$r = -0.005604873$	$t = -0.03170645$
Texas	$r = -0.01332668$	$t = -0.07539377$

VI. Discussion & Conclusions

Discussion

Presumably, there are various factors that impact both poverty and median income. Although it would be natural to assume the two are much more closely linked, the analysis done here shows to prove otherwise. They are possibly in some way linked, but not closely correlated. This could be that the varying levels of poverty are naturally created by differing economic crisis and conditions unrelated to the overall health of the economy, but possibly more related to government action. On the other hand, Median income could be more directly fueled by the economy, while also signifying a rise in higher income wages, rather than an overall increase which would also reduce poverty.

Conclusion

In conclusion, there is not a strong relationship between median income and poverty rate. Thus, we cannot reject the null hypothesis to identify a relationship between the two economic measures. Although the hypothesis has been rejected, this information can still provide significant insight to the American economy. While many would believe that raising the median income would provide opportunities for lower class Americans to find employment, and escape poverty, this is not the case. In order to combat poverty levels, we will have to look at different factors impacting those with lower incomes or unemployed.

VII. References

“Demographics of New York (State).” *Wikipedia*, Wikimedia Foundation, 21 Feb. 2019, [en.wikipedia.org/wiki/Demographics_of_New_York_\(state\)](https://en.wikipedia.org/wiki/Demographics_of_New_York_(state)).

“Median Household Income .” *Census.gov*, www.census.gov/quickfacts/fact/note/US/INC110217.

Pressman, Steven. “The Range of Poverty in America.” *U.S. News & World Report*, U.S. News & World Report, 12 Sept. 2018, www.usnews.com/news/healthiest-communities/articles/2018-09-12/poverty-in-america-new-census-data-paint-an-unpleasant-picture.

Research & Policy. “Focus on Poverty in New York City.” *NYU Furman Center*, 7 June 2017, furmancenter.org/thestoop/entry/focus-on-poverty.

Ura, Alexa, and Elbert Wang. “Poverty in Texas Drops to Lowest Levels in More than a Decade.” *The Texas Tribune*, Texas Tribune, 13 Sept. 2018, www.texastribune.org/2018/09/13/texas-poverty-census-2017-lowest-levels-decade/.

US Census Bureau. “Income, Poverty and Health Insurance Coverage in the U.S.: 2017.” *The United States Census Bureau*, 12 Sept. 2018, www.census.gov/newsroom/press-releases/2018/income-poverty.html.

Walters, Dan. “Why Does California Have the Nation's Highest Poverty Level?” *Sacbee*, The Sacramento Bee, 14 Aug. 2017, www.sacbee.com/opinion/california-forum/article166553177.html.

VIII. R Code

```
dirPath = "/Users/Lukeanderton/Desktop/OOps/GEOG 312/Class/Project/"
```

```
fName = list.files(dirPath, ".csv")
```

```
fPath = paste0(dirPath, "/", fName)
```

```
tab = read.csv(fPath[1])
```

```
x = tab$Median.Income
```

```
y = tab$Poverty..
```

```
plot(x, y,
```

```
  xlab = "Median Income",
```

```
  ylab = "Poverty (%)",
```

```
  main = "California Median Income vs. Poverty",
```

```
  pch = 16,
```

```
  col = "blue")
```

```
n = length(tab$Year)
```

```
b = (n * (sum(x*y)) - (sum(x) * sum(y))) / (n*(sum(x^2)) - (sum(x))^2)
```

```
a = (sum(y) - b * (sum(x))) / n
```

```
abline(a,b, col = "red",
```

```
  lty = 5)
```

```
#Pearson's Correlation
```

```
c = cov(x,y)
```

```
sx = sd(x)
```

```
sy = sd(y)
```

```
r = (1/(n-1)) * (c/(sx * sy))
```

```
r
```

```
#Student's Test
```

```
t = r * sqrt(n-2) / (sqrt(1 - r^2))
```

```
t
```

```
tab = read.csv(fPath[2])
```

```
x = tab$Long
y = tab$Latitude
medianincome = vector()
medianincome = tab$Median.Income
site = as.vector(tab$State)

sumx = 0
sumy = 0
for (j in 1:length(site)){
  sumx = x[j] * medianincome[j] + sumx
  sumy = y[j] * medianincome[j] + sumy
}
sumx = sumx / sum(medianincome)
sumy = sumy / sum(medianincome)

plot(x, y,
      xlab = "Longitude",
      ylab = "Latitude",
      main = "Spacial Center of Median Income in America")

par(new=T)
points(sumx, sumy, pch = 16, col = "red")
```