American Journal of Advanced Research, 2022, 6-1
July. 2022, pages 01-10
doi: 10.5281/zenodo.6845951
http://www.ibii-us.org/Journals/AJAR/
ISBN 2572-8849 (Online), 2572-8830 (Print)

Time Series, Cross-Sectional and Spatial Analysis of US poverty and Income Inequality

Yonas Lemma

The University of North Texas at Dallas – School of Business- Department of Decision Sciences

*Email: Yonas.Lemma@untdallas.edu

Received on June 6, 2022; revised on July 11, 2022; published on July 16, 2022

Abstract

This paper describes how poverty rate is and has been associated with the inequality of income distribution in the United States of America and also examines if there is any spatial(spillover) effect. Poverty rate in US has been steadily increasing for most of the time since 1987 and the reason for this increase seems something which needs to be thoroughly studied. I know, there are studies conducted on how the inequality of income distribution in the united states of America affects poverty rate across the states and the effect of some social variables in increasing the poverty rate and also I know that there has been studies that examined the spillover effect of poverty and income inequalities focusing on certain regions of the U.S but to my knowledge there has been no study that uniquely studied the effect of those social variables analyzed in my paper (unemployment, percentage of people without health insurance, inequality of income distribution, per capital income, population growth and educational attainment) on the poverty rate across all the states of the united states of America including district of Columbia and also the spillover effect at the same time. This was the motivation for conducting this research on how inequality of income distribution with other social variable affects poverty rate in US and by how much the magnitude varies when the effect of inequality of income distribution is seen independently than with the other variable. Since this paper has three sections (time series, cross-sectional and spatial section), one would be able to see the impact from three different dimensions.

1 Introduction

Data Source and Analysis

This study aims at examining the effects of inequality of income distribution, unemployment, health insurance coverage, per capital income, educational attainment and population on the poverty rate in the United States of America and also examines spatial dependence and spillover effect between poverty and other social variable. The study has been conducted in three different analyses. The first part of this paper deals with time serious analysis and the second part cross-sectional analysis and the last part focuses on the spatial analysis.

Time Series Data source and Analysis

The time series part of this paper aims at analyzing the poverty rate and inequality distribution of income in the United States of America from 1987 to the year 2010 and shows the relationship between poverty and inequality of distributions prior the year 2010 and the 2010 poverty rate and income inequality of distribution and predicts the poverty rate and inequality of income distribution. The data on the time serious section of this study is taken from the U.S census bureau for the period stated above (for the poverty rate and Gini coefficient which is the measure of inequality of income distribution) and 2000-2010 for the population, so the data has 24 observations for the first two variables and 11 observations for the other variable (i.e., population) all measured annually. I chose the period 1987-2010 for the first two variables because it was simple to get data for this range of years and also because this period is pretty enough to analyze the trend of poverty rate and inequality of income distribution in the United States of America. I chose to take the data for the population only for the period 2000-2010 simply because I don't have a substantial back up to support the linear trend that I get from the data which states a constant change on the growth of population in the United States of America for the period prior to 2000.

The average poverty rate for the period 1987 to 2010 is 13.2 whereas the median poverty rate is 13.1 which tells US that more than half of the observations are below the mean and the maximum poverty rate for this period is 15.1 and it was observed two times, that is, in the years 1993 and 2010. On the other hand, the minimum poverty rate for the given range of periods is 11.3 and it was observed in the year 2000 (see table 1 at on page 9). The average value for the Gini coefficient (a measure of inequality of income distribution) is 0.425 whereas the median value is 0.43 which tell us that more than half of the observations are above the mean and the maximum inequality of income distribution was observed in 2006 with the value of 0.444 as measured in Gini index whereas the minimum inequality of income distribution was observed on the beginning year from when the data is taken (1987) with the Gini index value of 0.393. On the other hand, the population of the United States has steadily grown from 282,162,411 in 2000 to 308,745,538 in 2010 as shown on the table 1 at the back which is based on the data from the U.S census bureau.

Cross-sectional Data Source and Analysis

The cross-sectional part of this paper aims at examining the effect of inequality of income distribution, unemployment, health insurance coverage, educational attainment, population growth and per capital income of the states on the poverty rate across the state of the United States in 2010. The data for this section of the study was taken from the Bureau of labor statistics, U.S census bureau, Bureau of Economic Analysis under the U.S department of Commerce. The Data has 51 observations which is equal to the number of states of the United States of America including the District of Columbia. All data are the year 2010 data. I chose the year 2010 because the year is more current and is more accessible to get data than the year 2011.

At the state level and for 2010 the average poverty rate is 14.2% and it was observed in Oregon and the maximum poverty rate for the year 2010 is 22.7% and it was observed in Mississippi followed by Louisiana and District of Columbia (D.C) with the poverty rate of 21.6% and 19.9% respectively. On the other hand, the minimum poverty rate for 2010 was 6.6 and was observed in New-Hampshire followed by Connecticut and Wyoming with the poverty rate of 8.3 and 9.6 respectively. On the other hand the average Gini coefficient (the measure of inequality of income distribution) in 2010 is 0.454 (45.4%) and was observed in Oklahoma and the maximum Gini index value was 0.532 and was observed in District of Columbia followed by New York and Connecticut with inequality of distribution value of 0.499 and 0.486 respectively whereas the minimum value for the Gini index is 0.419 observed in Utah followed by Arkansas and Wyoming with the Gini value of 0.422 and 0.423 respectively. For more detail See table 5

The other social variables that are taken into account in this study are unemployment, per capital income, educational attainment, health insurance coverage and population. The average unemployment rate for the year 2010 is 8.6% and was observed in New Mexico and the maximum unemployment rate (i.e., 14.9%) was observed in Nevada followed by the big state California (12.5%) and Michigan (11.9%). On the other hand, the minimum unemployment rate was observed in North Dakota (i.e., 3.9%) followed by Nebraska (4.5%) and its neighboring state South Dakota (4.6%). Table 3 shows people without health insurance ranges over 19 percentage points from a minimum of 5.6% in Massachusetts to the maximum of 24.6% in Texas. The average per capital income for all the states in the United States of America in the year 2010 was $40114.55 and the maximum per capital income for that period was $71044 observed in District of Columbia and the minimum per capital income for the same year was $31186 and was observed in Mississippi. Of course, state with the highest population is obvious and is California (37,341,989 people as of 2010) and the state with lowest population is Wyoming (568,300 as of 2010). To measure the educational attainment, I used college graduation rate (percentage of people graduated from a college in 2010). And college graduation rate ranges over 32.6 percentage points from a minimum of 17.5 in West Virginia followed by Mississippi (19.5) and Arizona (19.5) to the maximum of 50.1 in District of Columbia followed by Massachusetts (39) and Colorado (36.4).

Spatial Data source and Analysis

The Spatial section of this study targets to specifically study the spillover effect (at the state-level) of poverty, inequality of income distribution and other social variable and spatial dependence between these variables. The state level Data was collected from the same sources mentioned in the cross-sectional part except the fact that on the estimation part of this spatial study, only 49(out of 51) observations are used dropping Alaska and Hawaii since they don`t have contiguous neighbors. The State-level shape file was taken from ArcGIS Data base.

For the spatial weighting object, I used contiguity method to evaluate states based on proximity which according to the summary shows on average a state has approximately four contiguous neighbors with a minimum number of neighbors a state has is observed to be 1 and the maximum number of neighbors a state has is observed to be 8. Since analyzing the extreme values are very important, we can also analyze it using the maps and the values of every attribute for a given state. Having the hypothesis on my mind that a state which has the highest value for an attribute (say, poverty rate) has neighbors with only approximately the same value or relatively the same value for the given attribute and vice versa, I would like to see the spatial relationship based on the map and compare it with empirical result.

The sate Mississippi has the highest poverty rate (22.7) followed by its neighbor Louisiana and this relationship is parallel to the hypothesis where as New-Hampshire with a poverty rate of 6.6 contains the minimum value followed by its second most close contiguous state Connecticut (8.3). See figure 3 for detail. When we look at the inequality of income distribution across the states, District of Columbia has the highest Gini index value (0.53) followed by New-York (0.49) and the two are not most close contiguous neighbors. On the other hand, the lowest inequality of income distribution is observed in Utah with Gini value of 0.419 followed by Arkansas (0.422) and these two states are located very far from each other and this shows my spatial relationship hypothesis doesn`t hold at all in this situation. See figure 4 for detail. On the same scenario, Nevada and California are closest contiguous neighbors with the highest value of unemployment rate whereas North Dakota and Nebraska are the second closest contiguous neighbors with the minimum value of unemployment rate at the state level. On the other hand, for District of Columbia (DC) and Connecticut with highest per capital income and for Mississippi (MS) and Idaho with the lowest per capital income the spatial relationship hypothesis doesn`t hold, the same is true for DC & Massachusetts and West Virginia & MS with the highest and lowest values of college graduation rate. The two closest contiguous neighbors (Texas and New-Mexico) have the highest percentage of people without health insurance coverage whereas Massachusetts and Hawaii which are not neighbors at all has the lowest. Figure 5 explicitly shows the spatial distribution of poverty and inequality of income distribution across 49 states of the United States including the district of Columbia but excluding Alaska and Hawaii since they don`t have contiguous neighbors. One can easily tell from the map that generally speaking areas of high poverty rate are also areas of high inequality of income distribution.

2 Methods & Results

Throughout the study I run a univariate and multivariate linear regression and spatial auto regression for different models. The model is presented as follows

Times series model 1

In this analysis, I used a multivariate method of data analysis with some further empirical works and the model in these times series analysis relates the dependent variable poverty (%) with to its lags (p_1, P_2, P_3). The multivariate regression (regression using Newey west standard) as shown in table 2 will be:

Poverty= 1.83 + 1.85P_1 – 1.36P_2 + 0.36P_3 + ε…………………………………… (1)

Where ε (epsilon) represents standard errors that are *heteroskedasticity* and auto-correlation consistent standard errors. P_1, P_2 and P_3 represents poverty rate of one, two and three years prior to a given years respectively.

From the model we can see that the coefficients of P_1, P_2 and P_3 are 1.85, -1.36 and 0.36 respectively which shows they have 1.85, -1.36 and 0.36 impacts on the predicted poverty rate. So, if poverty rate increases by 1 percent this year, then after three years the poverty rate will increase by 1.85 – 1.36 + 0.36= 0.85 percentage points. Testing the sum at 5% level of significance gave me a p value of 0 which shows the overall impact of the lags on the poverty rate is statistically significant. In addition to that I predicted the poverty rate for 2011 and based on the data I have, the predicted poverty rate for 2011 is 13.2 but the official poverty rate released on September showed 15.2% even though other alternative measure shows a higher poverty rate

Times series model 2

For this part of the time serious model, I used the same regression (multivariate regression using Newey west standard) and the model relates the dependent variable Gini (the measure of inequality of income distribution) to its lags (g_1, g_2, g_3) which is

Gini = 0.07 + 0.73g_1 _ 0.27g_2 + 0.36g_3 + ε…………………………………. (2) where ε is the standard error which is heteroskedasticity and auto correlation consistent and g_1, g_2 and g_3 are one, two and three years` prior Gini from the given period`s Gini.

The model explicitly shows the overall impact of inequality of income distribution in the lags on the given year`s Gini which is positive (0.73+ 0.27+ 0.36= 0.82> 0). So a percentage point increase in the inequality of income distribution (measured by Gini) this year will be responded by a 0.82 percentage points increase in the expected value of Gini (inequality of income distribution) after 3 years which is also statistically significant tested at 95% confidence interval with p value of 0. I also predicted the Gini for the year 2011 and the predicted Gini for 2011 is 0.439 based on the Data.

Time series model 3

Time series model 3 relates the dependent variable poverty rate (%) to Gini (measure for the inequality of income distribution and the population.

Poverty = -26.58 + 20.61gini + 0.0000103pop + ε………………………………………. (3) where ε represent heteroskedasticity and auto-correlation consistent standard error and Gini measure the inequality of income distribution and pop represents population.

Here we can see that poverty and inequality of income distribution are positively related. When gini increase by 1 percentage point, the expected value of poverty increases by 20.61 percentage points where as when population increase by 1 percentage point the expected value of poverty increases by 0.0000103 percentage points. Even though the coefficient of gini looks exaggerated, it a kind of matched my initial hypothesis which I said on the data analysis part, increase in the inequality of income distribution currently is observed in the way that rich are getting richer and the poor are getting poorer .This looks like a distribution towards a middle class in America is getting smaller which in my intuition, more of the middle class's income will be distributed towards the rich but more of the people in the middle class whose income is distributed towards the rich will be pushed down towards the poor and become under the income threshold set as poverty line and this will increase poverty rate in the united states of America but the coefficient is insignificant (with p value =0.66 > 0.05). On the other hand, the coefficient of population is significant with p value of 0.002.

Cross-sectional Model 1

In this analysis I used a univariate method of data analysis in which case the data is of all the states of the United States and the model relates the dependent variable poverty rate (%) to Gini (the measure of inequality of income distribution which is also given as a percentage).

Poverty rate = -27.8 + 0.92Gini + ε………………………………………………………….. (4)

Where ε represents the errors and Gini represents the inequality of distribution across the states of the United States.

The association seems obvious; the increase in the inequality of income distribution is positively and significantly (with the p value of 0) associated with the increase poverty rate across the states. So, a 1 percent increase in the value of gini is responded by a 0.92 percentage points increase in the poverty rate of a given state. In the time series part (times serious model 3) we have seen that a 1 percent point in the gin was responded by an increase in the poverty rate by more than one percentage points but here we can see that a 1 percent increase in the gini is responded by an increase in the poverty rate by less than 1percnetage points and the reason seem to be of three things. The first one is of course it is because one is times serious model and the other is cross-sectional and the second reason is because one of the models is for the United States and the other is across the states of the United States. Whereas the last reason is the inclusion of a variable (i.e. population) might have caused the magnitude of the relationship to change. See table 7

Cross-Sectional Model 2

In this section, a multivariate method of data analysis has been used and the model in this section relates the dependent variable poverty rate to social variable as follow:

Poverty rate= -36.88 + 1.25Gini percent – 0.00019percapital income +0.041unemployment – 0.89collegegrads + 0.29peopleuninsured -0.00000071population + ε……………………. (5)

Where ε is an error term, Gini percent is a measure of inequality of income distribution given as a percentage and per capital income, unemployment, people uninsured are the social variable in my model.

Table 7 shows that inequality of income distribution (measure by percentage Gini index) is positively (with a coefficient of 1.25) and significantly (with p value of 0) associated with poverty rate across the states ceteris paribus which implies a 1 percent increase in Gini is associated with 1.25 percentage points increase in expected value of poverty rate. A one percent increase in unemployment rate is responded by a 0.041 increase in the expected value of poverty but it is not found to be significant as shown in table 7, whereas per capital income is negatively (with a coefficient of -0.00019) and significantly (with p value of 0.009) associated with poverty rate, other things held constant. On the other hand, population and educational attainment (college grads) are not significantly (with p value of 0.08 and 0.29 respectively) associated with poverty rate. The table also shows, a 1 % increase in health insurance coverage across the states is positively and significantly (with p value of 0) responded by a 0.29 percentage points increase in the expected value of poverty rate across the states of the United States.

Cross-Sectional Model 3

I used a multivariate method of data analysis with further simple empirical works and the model in this analysis relates the dependent variable, p_std (poverty rate standardized) to the standardized forms of the other social variables. The Multivariate linear regression (regression using robust) model from the result shown in Table 8 will be:

P_std= 1.43+ 0.75g_std - 0.39I_std + 0.02u_std- 0.15coll_std+ 0.34ppl_un_std- 0.14pop_std + ε………………………………………… (6)

The Greek letter ε represents standard errors that are robust to heterogeneity in the error variance. g_std , I_std , u_std. coll_std, ppl_un_std, pop_std represent gini percent ,per capital income, unemployment, college graduates, people without health insurance and population standardized respectively.

From the model we can see that a marginal change in poverty rate for a change in gini (a measure of inequality of income distribution) of 1 standard deviation is 0.75 ,holding other variables in the model constant and the result shows the measure of inequality of income distribution(gini) is significantly (with the p vale of 0<0.05) and positively associated with the poverty rate across the states which implies the increase in the inequality of income distribution increases poverty rate across the states. The result of this analysis as shown in table 2 also tells us that when per capital income increases by 1 standard deviation, the poverty rate decreases by 0.39 standard deviation holding other variables in the model constant and this is relationship is significant with the p value of 0.022. The model specifies, this is true that when income of an individual in a given state increases the poverty rate for the state decrease since poverty is defined as a luck of a socially acceptable amount of money or a state of being under a certain income threshold. The model specifies that an increase in unemployment rate of 1 standard deviation is responded by an increase of 0.02 standard deviations in poverty rate but it is not significant (having the p value of 0.69 > 0.05) tested at 5% level of significance leading us to reject the alternative hypothesis that says a change of unemployment of 1 standard deviation is responded by a marginal change in poverty rate of a standard deviation which is different from 0 where as an increase in college graduation rate of 1 standard deviation is responded by a decrease of 0.14 standard deviations in poverty rate is not also significant (p value= 0.26 > 0.05) which would lead us to accept the null hypothesis that a marginal change in poverty rate for a change in college graduation rate of 1 standard deviation is 0. The percentage of people without health insurance in a given state is positively and significantly (p value = 0) associated with the poverty rate in the state implying when the percentage of people with health insurance increase by 1 standard deviation the poverty rate increase by 0.34 standard deviations. According to the model above, an increase in a population of 1 standard deviation in a given state decreases the poverty rate of the given state by 0.14 standard deviation, it is found to be non-significant (with the p value of 0.13 > 0.05).

Spatial Models

A Spatial Model is used in this part of the study. The dependent variable being the Standardized poverty rate, the included control variables are those related to poverty rate directly or indirectly. The spatial models used for comparison includes, Spatial Autoregressive model (SAR), Spatial Error model (SER), Spatial

Durbin model (SDM) and Spatial Autocorrelation model (SAC) and all the Models are given as follow based the result from MATLAB:

Spatial Autoregressive Model (SAR) ⇛ y=ρWy + Xβ+ ε ------------------ (7) where $\varepsilon \sim N(0_{nx1}, \sigma^2 I_n)$,WY is spatial lag of Y and ρ is the spatial lag parameter. Table 9 shows the value of ρ to be -0.11 which implies there is a negative spillover effect among the poverty rate of neighbors but the result is not statistically significant which would lead us to accept the null hypothesis that there is no spatial correlation between poverty rate of a given state and its neighbor. The Lagrangian Multiplier test (LM lag =1.101) implies there is no spatial dependence due to missing spatially lagged dependent variable. See table 9 for detail

Spatial Error Model (SEM) ⇛$y=X\beta+ (I- ʎW)^{-1}V$………………………… (8) Where ʎ indicates the extent to which the spatial component of the errors ε which are contained in V are correlated with one another for nearby observations. ʎ = 0.106 as shown in table 9 but is not significant implying the extent to which the spatial error components are correlated is not significantly different from 0 or simply, there is no spatial correlation between the errors for connected observations. LM error= 0.095 indicates the absence of spatial dependence in the errors

Spatial Durbin Model (SDM) ⇛ y=ρWy +X(β+γ)-WXρβ+V ----------- (9) where ρ is the spatial error parameter which is equal to -0.105 as shown in table 9 and indicates a negative but insignificant spatial dependence tested at 1% level of significance.

Spatial Autocorrelation Model (SAC) ⇛$y=\rho Wy+X\beta+(I-ʎW)^{-1}\mu$…..(10) where ρ is the spatial lag parameter and ʎ is the spatial parameter of the error in the model. The value of ρ in this model is as shown in table 9 is – 0.23, a negative but insignificant spatial dependence in the lags and the value of ʎ is 0.3 , a positive and also significant spatial dependence in the error part of this model but it could be wrong to conclude that because ʎ is significant only in this model, this model is the right model rather than comparing all the models. Now, I would like to go head and compare the models and choose the right model. Comparison of the above different model specifications using likelihood-based testing can easily be done as follow.

Comparing spatial autoregressive Model (SAR) and Spatial Error Model (SEM)

lr1=2*(sar1.lik-sem1.lik) gives you the likelihood ratio which in this case is equal to 0.8229 and

1-chi2cdf (sar1.lik-sem1.lik, 1) gives you the probability which in this case is equal to 0.5212 indicating to choose Spatial Error Model (SEM) in favor of Spatial Autoregressive Model (SAR).

Comparing Spatial Error Model (SEM) and Spatial Durbin Model (SDM)

lr2=2*(sdm1.lik-sem1.lik) =10.9486 is the value of the likelihood ratio where as

1-chi2cdf (sdm1.lik-sem1.lik, 1) = 0.0193 is the probability and it indicates that Spatial Durbin Model (SDM) is a better Model for compared to the Spatial Error Model (SEM)

Comparing Spatial Durbin Model (SDM) and Spatial Autocorrelation Model (SAC)

lr3=2*(sdm1.lik-sac1.lik) gives us the likelihood ratio which is equal to 8.1310 where as

1-chi2cdf(sdm1.lik-sac1.lik,1) gives us the probability and is equal to 0.0438 indicating to choose Spatial Durbin model (SDM) in favor of Spatial Autocorrelation Model (SAC).

From the law of transitivity (if A>B and IF B>C then A>C) we can conclude that if Spatial Durbin Model (SDM) is chosen over Spatial Error Model (SEM) and If Spatial Error Model (SEM) is chosen over Spatial Autoregressive Model (SAR) then Spatial Durbin Model (SDM) should be chosen over Spatial Autoregressive Model (SAR) and this leads us to the conclusion that Spatial Durbin Model (SDM) is the best model for this study. The Spatial Durbin Model is given as follow

y=ρWy +X (β +γ)-WXρβ+V --- (11) where y is the dependent variable, X is a vector of independent variables, W is the contiguity weight matrix, ρ is the spatial lag parameter and V is vectors of errors. Based on the result in MATLAB the Spatial Durbin Model will be

P_std= -.04 - .10W*P_std + .90g_std – .43I_std + .44u_std – .13coll_std + .26ppl_un_std – 0.11pop_std – (-0.25 W*g_std – 0.01 W*I_std - .08W*u_std – .13 W*coll_std – .13 W*ppl_un_std + .10W*pop_std) + V

……………………………………………….. (12)

Where P_std ,g_std, I_std, u_std, coll_std, ppl_un_std and pop_std are the standardized poverty rate ,gini, per capital income, unemployment, college grad rate, people uninsured and population respectively and W is the spatial contiguity weight matrix.

Since the spatial Durbin Model takes the neighboring states dependent and explanatory variables into account by adding their spatial lags, the model is expected to capture the direct and indirect effects easily. The negative value of the spatial lag parameter (-.10) seems unlikely as it means poverty rate of a given state is negatively associated on the poverty rate of its neighbor but it is statistically insignificant. Per capital income and inequality of income distribution (gini index) are found to significantly explain the poverty rate even though per capital income is negatively associated with poverty rate where as inequality of income distribution is positively associated with poverty. Generally, in almost all the models the spatial parameters implied the non-existence of spatial autocorrelation and Moran`s I did the same (0.0312). see table 9 & 10 for detail.

Conclusion

It is obvious that there are many social problems that are debatable and questioning across the states. But frankly speaking the results gave me a big picture on the percentage gap in the magnitude of having those social problems like unemployment, poverty, health insurance coverage across the states. One can just refer the 3 pages data section and know so many necessary things about what US used to look like since 1987-2010 than what he/she actually have in mind.

This study showed that inequality of income distribution has been significantly associated with poverty rates from 1987 to

2010. The study also showed that inequality of income distribution and per capital income significantly affects poverty rate across the states indicating the increase in the inequality of income distribution (as measure in gini index) increase the poverty rate whereas the increase in per capital income is responded by a decrease in poverty rate of a given state but there was little or no spatial spillover effect. In the time series and cross-sectional part, I found the data being consistent with my hypothesis that I had in mind that when inequality gap widens the number of people below the poverty line increases or some people would be pushed in to the poverty line since their income would be below the specified threshold but the result in spatial analysis part is not consistent with a hypothesis I had in mind (i.e. poverty rate in a given state is positively and significantly associated with poverty rate of its neighbor) . Even though my initial hypothesis on state unemployment rate and the educational attainment was found to be true from the regression result and from the interpretation of the coefficient, they suffer from being not significant as seen from p values they have which are greater than 0.05. It might be because other variables in the model are causing this relationship not to be significant.

As far as the limitations, I have no reference about whether such studies with all and same variables I have included in the model have been conducted or not. In addition to that I don't have clear evidence whether taking college graduation rate for my education attainment variable is more appropriate than other educational levels, say high school, except the fact I believe that on average people in the age of 20s are more productive than other ages and most people graduate from college in their 20s which would mean they would more likely be able to get a job and get income that is in most case above the poverty threshold income level. Finally, the linear tend I got from graphing the population in the United States of America made me say only a little bit about the relationship between population and poverty rate in the US.

Acknowledgements

I would like to thank Dr. James Murdoch and Dr. Nathan Berg for contributing ideas for my research work. I would also like to thank Dr. Karen Shumway for being a great supervisor of my academic career that helped me improve my research project.

Conflict of Interest: none declared.

References

1. Carmen DeNavas, Bernadette D. Proctor Jessica, Income, Poverty, and Health Insurance Coverage in the
2. United States: 2010. Current Population Reports, issued September 2011

3. Kathleen Short, The Research supplemental poverty measure: 2010. Current Population Reports, issued November 2011

4. Nathan Berg, Gerg Gigerenzer, As-If Behavioral Economics: Neo-classical Economics in Disguise? 2010

5. Sudiksha Joshi, Tesfa G.Gebremedhin, A spatial Analysis of poverty and Income Inequality in the Appalachian Region. Research paper 2010-1015

6. Tammy Leonard, James Murdoch, the Neighborhood Effects of Foreclosure. Journal of Geographical System, 11, Nov 2009, 319-332

7. Smeeding, Timothy, Poor People in Rich Nations: The United States in Comparative Perspective." *Journal of Economic Perspectives,* 2006, vol.20(1): 69–90

Table 1. Summary of poverty, inequality of income distribution (measured by gini) and population

Variable	Observation	Mean	Std. Dev.	Min	Max
poverty (%)	24	13.20833	1.064003	11.3	15.1
gini (0-1)	24	0.4245	0.01686	0.393	0.444
Population(million)	11	296	8.9	282	309

Table 2. Multivariate regression (using newey west standard errors) model of the association between poverty rate and its lags

poverty	Coefficients	Std. Err.	t	P>t	[95% Conf. Interval]
P_1	1.857804	0.164418	11.3	0	(1.510912, 2.204696)
P_2	-1.35885	0.299642	-4.53	0	(-1.99104, -0.72666)
P_3	0.364892	0.186526	1.96	0.067	(-0.02864, 0.758428)
_cons	1.830717	0.816574	2.24	0.039	(0.107897, 3.553537)

Table 3. Multivariate regression (using newey west) of inequality of income distribution (measured by gini) and its lags

gini	Coefficients	Std. Err.	t	P>t	[95% Conf. Interval]
g_1	0.733648	0.152871	4.8	0	(0.411117, 1.056178)
g_2	-0.26528	0.156137	-1.7	0.108	(-0.5947, 0.064139)
g_3	0.364344	0.037617	9.69	0	(0.284979,0.443709)
_cons	0.074241	0.055802	1.33	0.201	(-0.04349, 0.191973)

Table 4. Multivariate regression Model of the association between poverty, gini (measure of inequality of income distribution) and population

poverty	Coefficients	Std. Err.	t	P>t	[95% Conf. Interval]
gini	20.61706	44.83201	0.46	0.658	(-82.7657, 123.9998)
population	0.0000013	0.00000222	4.63	0.002	(-5.14000, 1.54000)
_cons	-26.5769	18.95486	-1.4	0.198	(-70.2869, 17.13307)

Table 5. Summary of poverty rate, gini index, per capital income, unemployment, educational attainment, health insurance coverage, and population across the states(n=51)

ariable	Mean	Std. Dev.	Min	Max
overty rate(%)	14.18627	3.493195	6.6	22.7
umber of people in overty	905509.8	1143414	51000	6077000
ni percent(%)	45.37647	2.089075	41.9	53.2
ni index	0.4537647	0.0208908	0.419	0.532
r capital income	40114.55	7051.909	31186	71044
nemployment rate)	8.635294	2.086511	3.9	14.9
ollege grads (%)	27.93333	5.723309	17.5	50.1
ople uninsured (%)	14.86471	4.043257	5.6	24.6
opulation	6074219	6845288	568300	37341989

** gini = the measure of inequality of income distribution= the ration of the area under the Lorenz curve and the total area under the 45º line.

Table 6. Bivariate regression model of the association between the inequality of income distribution and poverty rate across states

Poverty rate	Coefficients	Std. Err.	t	P>t	[95% Conf. Interval]
gini index	0.925447	0.198954	4.65	0	(0.525635, 1.32526)
_cons	-27.8073	9.037197	-3.08	0.003	(-45.9682, -9.64634)

Table 7. A multivariate regression model of the association between poverty rate and gini index, per capital income, unemployment, educational attainment, health insurance coverage, and population across the states

poverty rate	Coefficients	Std. Err.	t	P>t	[95% Conf. Interval]
gini percentage	1.254845	0.15728	7.98	0	(0.937868, 1.571821)
Per capital income	-0.00019	0.000705	-2.71	0.009	(-0.00033, -000049)
unemployment	0.041344	0.14123	0.29	0.771	(-0.24329, 0.325975)
College grads	-0.08987	0.084633	-1.06	0.294	(-0.26044, 0.080693)
People uninsured	0.294457	0.075404	3.91	0	(0.142491, 0.446424)
population	-0.00000071	0.00000039	-1.79	0.08	(-0.000015,0.0000008)
_cons	-36.8752	5.835607	-6.32	0	(-48.6361, -25.1143)

Table 8. Multivariate regression model of the association between standardize poverty rate and the standardize gini, per capital income, unemployment, educational attainment, people without health insurance and the population.

p_std	Coefficient	Std. Err.	t	P>t	[95% Conf. Interval]
g_std	0.750449	0.104106	7.21	0	(0.540636, 0.960261)
I_std	-0.38587	0.161908	-2.38	0.022	(-0.71217, -0.05956)
u_std	0.024695	0.063093	0.39	0.697	(-0.10246, 0.151851)
coll_std	-0.14725	0.129849	-1.13	0.263	(-0.40894, 0.114444)
ppl_un_std	0.340825	0.077687	4.39	0	(0.184257, 0.497392)
pop_std	-0.14035	0.090935	-1.54	0.13	(-0.32362, 0.042916)
_cons	1.43006	0.067343	0	1	(-0.13572, 0.135721)

Table 9. Spatial Auto regression, Spatial Error, Spatial Durbin and Spatial Autocorrelation Models

	SAR ⇒Y=ρWY + XB + e	SEM⇒Y=XB+(I-ʎW)$^{-1}$V	SDM⇒Y=ρWY +XB(I-ρW)+V	SAC⇒Y=ρWY+XB+(I-ʎW)$^{-1}$μ
	Coefficients	Coefficients	Coefficients	Coefficients
Constants	-0.035757	-0.024907	-0.040134	-0.020869
g_std	0.825998	0.824066	0.901788	0.896857
I_std	-0.463677	-0.443102	-0.430622	-0.475146
u_std	0.013809	0.023003	.442585	0.018601
coll_std	-0.130482	-0.123108	-0.133600	-0.141143
ppl_un_std	0.383449	0.330138	0.268737	0.352512
pop_std	-0.121716	-0.137299	-0.118452	-0.108921
W*Xi			**************	
Rho	-0.111970		-0.105995	-0.231999
Lamda		0.106000		0.393991
Log-like	-11.113589	-11.524771	-6.0500351	-10.133566
R^2	0.8196	0.8145	0.8497	0.8345
Moran`s I=0.0312				

W*Xi= spatial lags of the explanatory variable and are available in the SDM model

Table 10. Coefficient Estimates of the Spatial Durbin Model (SDM) including direct, indirect and total effect for states of the United States in 2010.

	SDM General		Direct effect		Indirect effect		Total effect	
	Coefficient	z-prob.	Coefficient	t-prob.	Coefficient	t-prob.	Coefficient	t-prob.
Constants	-0.040134	0.489299	-	-	-	-	-	-
g_std	0.901788	0.000000	0.904747	0.000000	-0.33004	0.030590	0.574699	0.000042
I_std	-0.430622	0.000886	-0.43146	0.001816	0.033493	0.898188	-0.39797	0.195885
u_std	.442585	0.658066	0.049405	0.633394	-0.084586	0.655546	-0.035181	0.802937
coll_std	-0.133600	0.276975	-0.127420	0.322531	-0.107412	0.645055	-0.234833	0.362013
ppl_un_std	0.268737	0.014553	0.269863	0.023676	0.094827	0.582455	0.364690	0.005441
pop_std	-0.118452	0.126897	-0.120483	0.132963	0.101074	0.597828	-0.019409	0.932089
W*g_std	-0.258277	0.220772	-	-	-	-	-	-
W*I_std	-0.004846	0.986817	-	-	-	-	-	-
W*u_std	-0.084733	0.647508	-	-	-	-	-	-
W*coll_std	-0.126247	0.614086	-	-	-	-	-	-
W*ppl_n_d	0.133705	0.457276	-	-	-	-	-	-
W*pop_std	0.100269	0.622840	-	-	-	-	-	-
Rho	-0.105995	0.595535	-	-	-	-	-	-

Appendix

Figure 1. US Poverty rate from 1987-2010

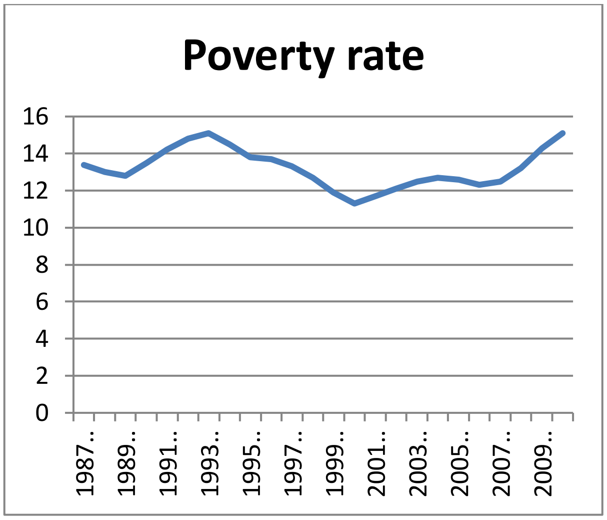


Figure 2. Inequality of income distribution (measured by Gini index) in USA from 1987-2010

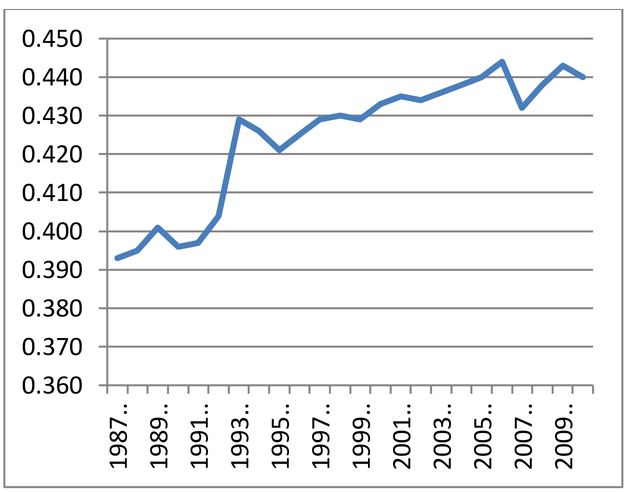
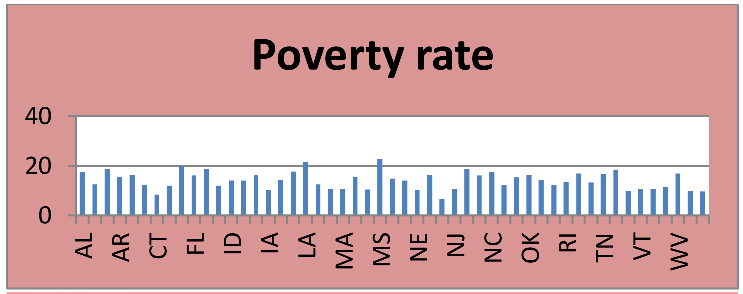


Figure 3. Poverty rate across states of the United States of America (2010)



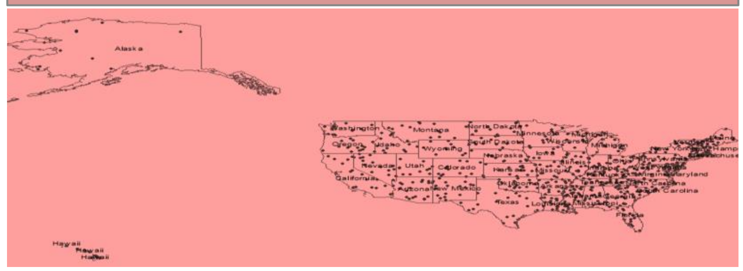


Figure 4. Inequality of income distribution (measure by Gini index) across US states

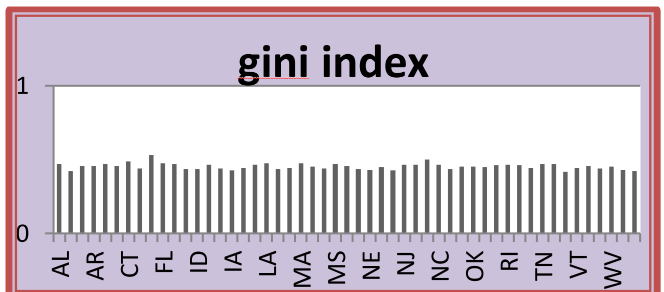
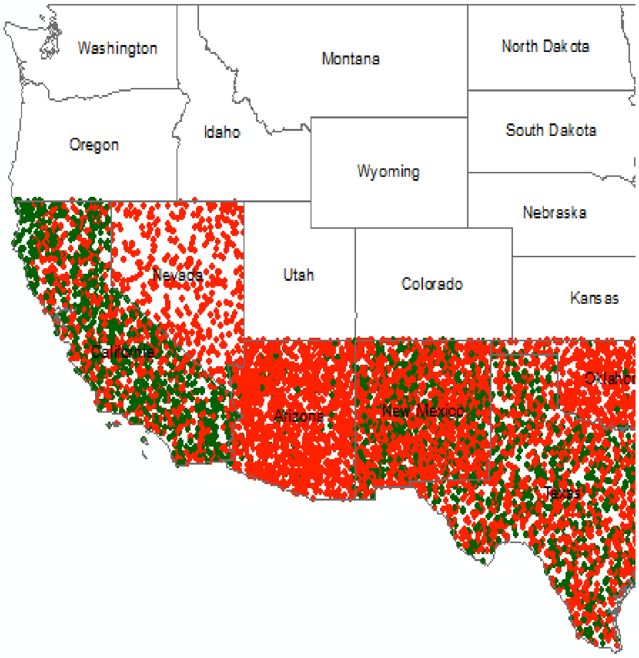


Figure 5. Spatial Distribution of Standardized poverty rate (p_std) and Standardize Gini index (g_std) at the state level



p_std=red
g_std=green
min=white
max=green/red