Factors Affecting the Poverty Rate in North Carolina Counties

Kimberley Mitchell, Fall 2012

Introduction

In this project, we look at demographic and economic variables affecting poverty rates by examining census data by county from a representative state, North Carolina.

In this paper we assume that poverty is affected by various demographic factors; among them age, education, and family composition (female headed households with young children being the most likely demographic to be living in poverty). The primary paper used as a reference (Levernier, et al, 2002) examined county data for all 50 states in the United States as reported for 1990, consisting of over 3000 observations. In this paper, we use a reduced sample, only looking at North Carolina data, and using more current data, from the years 2006 to 2010.

The intractability of poverty in the United States, which has one of the highest poverty rates in the developed world, while being one of the richest countries in per capita income, has long been a topic of interest to researchers both within the United States and worldwide (Smeeding et al, 2001). While much work has been done recently on the growing inequality of incomes in the United States (Piketty and Saez, 2003), research into factors affecting poverty itself is most prevalent when welfare reform is on the United States legislative agenda, which it was most recently in the 1990's and may well be again this year as the US government addresses spending excesses in an attempt to reduce the fiscal drag on the economy.

In this paper, we explore several economic factors expected to also affect the prevalence of poverty, including workers in relevant industries, labor force participation rate and population density. It is supposed that certain industries provide an economic benefit, not only through the people they employ, but also through "feeder processes" such as parts suppliers or services that are needed to supply their operations. The loss of jobs in manufacturing has often been pointed to an economic problem for the communities that lose the jobs and revenues, so it will be of interest whether regions with a high concentration of manufacturing jobs also tend to have lower rates of poverty.

A factor not included in the Levernier study was the rate of workforce participation. In the latter stages of our slow recovery from the financial crisis of 2008, the proportion of labor force participation may illustrate the effect of those who have dropped out of the active work force. A study by Mills and Hazarika (2003) found that increased workforce participation in the wake of welfare reform measures did improve economic conditions for families headed by young mothers.

The population density was included to investigate whether there was an "inner city" poverty issue, which might be indicated by higher poverty being associated with higher population density, or contrarily, that low population densities as seen in rural areas, may present difficulties in commuting.

In this study, our independent variables are the demographic and economic characteristics discussed above, and the dependent variable is the incidence of poverty as measured by the percent of the population living in poverty as defined by the census bureau's thresholds. (more detailed explanation at http://www.census.gov/acs/www/Downloads/data_documentation/SubjectDefinitions/2010_ACSSubjectDefinitions.pdf)

We were readily able to identify demographic and economic factors that had a significant effect on the poverty rates in this study. Many variables showed signs of autocorrelation by showing statistical significance when tested alone or with one set of factors, and registering as insignificant when tested

with another set of variables. In the end, our final model includes a trimmed-down set of only 4 factors,

that together explain almost 80% of the exhibited variability in poverty rates.

Theoretical Model

The question of what causes poverty has long been a moral and political, as well as an economic question, but economics certainly has much to say about this question. The two most common causes cited are 1) that poverty is an inevitable result of systemic structures and processes in our economic system and 2) that poverty is primarily caused by cultural influences that render the poor unable or unwilling to take the actions that would improve their lot (Andy, 2011). There is now a third thread or reasoning for the occurrence of poverty that uses the relatively new branch of Behavioral Economics to postulate that the poor are plagued with the same irrational tendencies as the rest of the population, but because of their reduced resources, face more severe economic consequences. (Bertrand et al, 2004)

Our paper focuses more on the systemic structure theory of poverty; we look at objective variables in the environment that these families inhabit to identify factors that affect the community poverty rates. Here we rely on some of the hundreds of variables collected by the US Census Bureau in the ongoing American Family Survey. This data has the advantage of being carefully and systematically collected in the same dimensions no matter what the political or trending focus may be. The data is credible, copious, consistent and available for every county. Its disadvantage is that it provides a summary only; there is no way to drill down and take a closer look at individual households when factors warrant a more detailed analysis.

North Carolina was chosen as the study locale because of its wide range of economic industries and conditions, incorporating growth in modern industries such as finance, information technology and biotechnology, mirroring economic trends seen elsewhere in the country, as well as maintaining some traditional industries such as agriculture (hog farming) and manufacturing (wooden furniture). (SSRI, 2012) With 13 percent poverty, North Carolina has a somewhat higher than average poverty rate with

respect to the majority of states, yet has some counties with as low as a 6% rate. North Carolina was also a good study subject because of the availability of 100 counties of census data.

Variables considered

Poverty is a condition that is affected by so many factors, and the American Family Survey data is so detailed that there were many avenues to explore. The factors we considered in this study are described below.

Percent of households with a single parent: Both from past data and from evaluation, we expect families with children and only a single adult to have a more difficult time, due to the need for child supervision and the cost of this child care. Both male-headed and female headed single parent families were included in this analysis.

Educational attainment: As education is an important investment in human capital, we included the percent of the population at various levels of educational attainment in our model. Categories included the percent of the population who had not completed high school, the percent with a high school diploma but not a bachelor's degree, the percent who had completed a bachelor's degree and those who had received a graduate degree.

Industry Sector Composition: As it is expected to be advantageous for any group to focus their efforts on endeavors in which they have a comparative advantage, it was anticipated that the proportion of the workforce in various industry sectors would have a significant effect on the poverty rate. All the standard industry sectors were included in the analysis, but it was expected that the manufacturing sector (a traditional strength in North Carolina) and the Finance and Technologies sector (especially strong in the Research Triangle area) would show significant improvements.

Labor Force Participation: labor force participation is traditionally considered a choice freely made, but in the midst of our current slow recovery from the financial meltdown of 2008, many people who would otherwise been employed, have been unable to find full-time employment. It has been noted that the drop in unemployment rate from close to 10% to just under 8% has been almost entirely due to a drop in labor force participation of the same magnitude, from 62.9% in January 2009 to 58.2% in December 2009, and has remained in the region of 58.5% since.(Rothstein, 2012) For this reason, we thought that the rate of labor force participation may have a significant effect and be inversely related to the poverty rate.

Median Income: Median family earnings and median earnings for female full-time year-round workers a (in 2010 inflation-adjusted dollars). Finally, this variable was not included in the final model, due to suspected endogeneity. It seems self-evident that the median family income may be affected by the poverty rate, as much as the poverty rate is affected by the median income (in that each additional poor household reduces the median household income by a small amount), and thus the median income of households is likely to be endogenous to our model. When median family income is included in the model, it is measured as being significant, but the variable associated with the percent in the labor force falls out of significance. It is the opinion of the author that this variable is distortionary due to endogeneity and so it was not included.

Fortunately, it was not necessary to identify specific instrumental variables, as exogenous variables were available that have a direct effect on poverty and were well-correlated with county median incomes.

Population density: Population density per square miles of land mass in the county; the first variable is population density, the second is population density by housing unit, and the third is the log of the population density per square miles of land mass. None of the population density variables was found to

be significant in the population studied. Examining the scatter plot of population density versus poverty

rate shows that the most densely populated counties have relatively low poverty rates (less than 10%)

and the lesser populated counties vary widely in rate of poverty. In addition, there are very few North

Carolina counties that can be said to be densely populated (3 counties with greater than 1000 people

per square mile), so there is probably not an adequate sample size to detect a significant effect if it

exists.

Racial composition: It was assumed for this study that being of one race, white, may have an effect on

the prevalence of poverty. In fact, the percent of white residents did have a significant effect of

improvement in the poverty rate, but the data were strongly correlated with some of the other

significant variables, especially the percent of households with a female householder and children under

the age of 18. Based on the expectation that the poverty rate may vary between white and non-white

populations, the percent of white residents was also considered, and in fact was highly correlated with

the poverty rate (-0.5978). However, it was not included in the final model due to its high correlation

with the rate of female headed households with children (-0.8214), which was deemed to be a more

direct causes of poverty, and which was not statistically significant when race was included as a factor.

Age distribution: The population of each county was broken down into several categories thought to
have a likely effect on the poverty rate as follows:

- Less than 20 years old
- 21 to 34 years old
- Over 65 years old

In particular, it was thought that the extremes of age, the very old and the very young, would be more

susceptible to a higher rate of poverty. None of these were found to have a significant effect with the

other factors in play.

Population Density: The density of dwellings per square mile of land mass of the county was also

considered as a predictor variable. Two mechanisms were under consideration: Frist, that the most

densely populated areas may have the "inner city" effect of entrenched poverty, and contrarily, that sparsely populated areas may exhibit increases in poverty due to a scarcity of local employment opportunities.

Estimation Method

The estimation method used in the main reference is ordinary least squares (OLS), and this is the method that is also used in this paper.

Empirical Model, Data and Estimation Technique

Poverty is a complex phenomenon that has given rise to a wide variety of explanations; and it is not possible in a paper of this size to examine them all. Factors presumed to influence poverty rates that were examined in this paper include educational attainment, median family income, median income of female workers, age distribution, industry sector composition, prevalence of single-parent families, population density, and the percent of the population in the labor force.

Data Details

Observational units: Each of the 100 counties in North Carolina were included, using data from the 2006-2010 American Community Survey 5-Year Estimates. This data is collected on an ongoing basis, by the United States Census Bureau. All data in this survey are estimates based on ongoing samples. The 5 Year estimates (2006 to 2010) are used, because the 1-Year and 3-Year estimates have population size thresholds, and so data are not available for all of the North Carolina counties.

Units: In order to normalize across highly populated and low populated counties, all variables are measured in terms of the percent of the population of the county, with the exception of population density, which is measured in terms of people per square mile of land (excluding water area) in the county, and median income (measured in inflation adjusted 2010 dollars).

Source: This data was obtained from the US Census website at www.census.gov, via the American FactFinder utility, which enables the extraction of data subsets of interest. The datasets used for this study are ACS_10_5YR_DP05 and ACS_10_SF4_DP02.

Variables Included in Final Model

The variables listed below are measured in terms of the county total population. In each case, there is one value per North Carolina county, for a total of 100 observations.

Poverty Rate: Percentage of families and people whose income in the past 12 months is below the poverty line, and log of this percent. (p_fam_pov, l_p_fam_pov).

Labor Force Participation: Percent of the population in the labor force (p_inlab) is expected to vary inversely with the poverty rate.

Industry Sector Composition: Percent of the population 15 years and over whose most recent job was in the industry sector being measured. Of the 15 industry sectors examined, only the percent involved in the manufacturing industry was significantly correlated with the poverty rate, and it varies directly with the poverty rate. (p_ind_manu)

Head of household: The percent of single parent households as follows; percent of male headed households, no wife present; female headed households, no husband present, and female headed households, no husband present with children under 18. Since raising a family as a single parent, makes the responsibility of child care more challenging to manage, it was thought that perhaps single parent homes headed by men would also face an increased likelihood of poverty. Such was not found to be the case; and the most significant factor affecting the countywide poverty rate was the rate of families which included minor children with only a single, female parent. The percentage of these households varied directly with the poverty rate. (p_femhh18)

Educational Attainment: Variables of educational attainment included the following:

- Percent not completing high school
- Percent with a high school diploma, no Bachelor's degree
- Percent with a bachelor's degree, no graduate degree
- Percent with a graduate degree

The category found to have a significant effect on the poverty rate was the percent of the population without a high school diploma. (p_nohighsch), which was found to track directly with the poverty rate.

The population's percentage of bachelor's and/or graduate degrees was not found to have a significant effect on the poverty rate once the rate of high school completion was taken into account.

Data description and summary follows:

Data Summary					
Variable	# Obs	Mean	Std.Dev.	Min	Max
% under poverty line	100	13.011	4.137801	6	26
% in labor force	100	59.91	5.995495	43.7	73.3
% in manufacturing	100	15.022	6.202664	2.9	31.2
% fem head w/children	100	7.341	2.367119	2.9	14.2
% no high school	100	19.718	5.365837	8.2	3.3
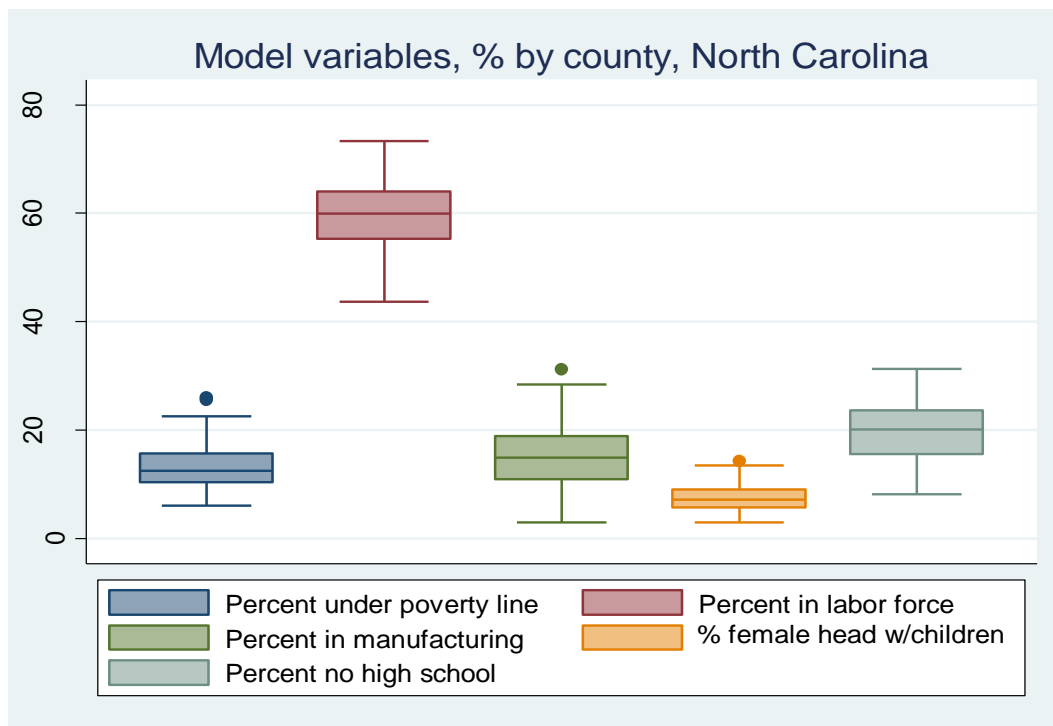
Correlation matrix:

Correlations	% under poverty line	% in labor force	% in manufacturing	% fem head w/children	% no high school
% under poverty line	1.0000				
% in labor force	-0.5690	1.0000			
% in manufacturing	0.1949	-0.0523	1.0000		
% fem head w/children	0.6143	0.0027	0.1459	1.0000	
% no high school	0.7004	-0.5403	0.5835	0.2981	1.0000

Scatter Plot Matrix

Empirical Results

The table of results, below, includes only those factors found to be significant.

Variable	Coefficient (Robust Std. Err.) No. Obs	Coefficient (Robust Std. Err.) No. Obs
Dependent Variable (name)	Percent of households under poverty line (p_fam_pov)	Log of percent of households under poverty line (l_p_fam_pov)
Percent of adults (age > 18) not completing high school (**p_nohighsch**)	0.3721523 (0.0711974) 100 p>t=0.000	0.0313543 (0.0059973) 100 p>t=0.000
Percent over 16 who did paid work during the reference week (**p_inlab**)	-0.2200396 (0.0470088) 100 p>t=0.000	-0.0145359 (0.0043584) 100 p>t=0.001
Percent of workers in the manufacturing industry (**p_ind_manu**)	-0.1173524 (0.0452484) 100 p>t=0.011	-0.0083465 (0.0038658) 100 p>t=0.033
Percent of female-headed households (with their own children under the age of 18) (**p_femhh18**)	0.8687601 (0.0948376) 100 p>t=0.000	0.0608785 (0.0068876) 100 p>t=0.000
Constant	14.24077 (3.445492) 100 p>t=0.000	2.447054 (0.3249152) 100 p>t=0.000
Model Statistics R-squared Adjusted R-squared SSE	Number of obs = 100 F(4, 95) = 89.58 Prob > F = 0.0000 R-squared = 0.7904 Adj R-squared = 0.7816 Root MSE = 1.9336	Number of obs = 100 F(4, 95) = 70.88 Prob > F = 0.0000 R-squared = 0.7490 Adj R-squared = 0.7384 Root MSE = .16373
Tests for heteroscedasticity		
Breusch-Pagan / Cook-Weisberg	chi2(1) = 4.55 Prob > chi2 = 0.0330	chi2(1) = 1.54 Prob > chi2 = 0.2153
White's test	chi2(14) = 18.13 Prob > chi2 = 0.2008	chi2(14) = 17.10 Prob > chi2 = 0.2509

The Final Model

The final model estimated is recorded below.

$$\beta_5 \text{l\_fam\_pov} = \beta_1 + \beta_2 \, \text{p\_inlab} + \beta_3 \, \text{p\_ind\_manu} + \beta_4 \, \text{p\_femhh18} + \beta_5 \, \text{p\_nohighsch}$$

Our final model is accounts for almost 80% of the variability of poverty rates by county with a very compact model, consisting of only 4 independent factors. The variables included in the final model are all significant at the $p=0.01$ level, except for the percent of workers in the manufacturing industry (p_ind_manu), which is significant at the $p=0.05$ level, and all of the signs of the coefficients are as expected.

All four of the significant factors are expressed in percent of the population, as is the dependent variable, making it easy to assess the results. The effect of a change of 1 in the coefficient of the factors, corresponds directly with a change of 1% in the poverty rate. So by looking at the factor coefficients, it is easy to see that the factor with the highest magnitude of influence is the percent of households with a female head of household and children under 18. This factor has an estimated coefficient of 0.867. This means that all else being equal, according to this model, every percent increase in this demographic leads to an increase in the poverty rate of 0.867%

The next most influential factor is the percent of the population that have not completed high school, where an increase in the high school completion rate by one percent (the inverse of the variable studied) would tend to decrease the poverty rate by 0.37%.

The other 2 factors, labor force participation, and percent of workers in the manufacturing sector have an inverse relationship with the poverty rate, in the magnitude of 0.22 to 1 percent and 0.117 to 1 % respectively.

Heteroscedasticity

Although there was an indication of an issue with heteroscedasticity when tested with the White test, the significance of every factor held up against the robust standard errors and also after the log of the poverty rate was regressed against the significant factors.

The pattern in residual variability is visually apparent in a graph of residuals against the predicted values (see figure 1). The residuals for fitted values under 10 seem to be significantly less variable than the residuals for higher fitted values. However, the significance of the estimates holds up when tested with robust variances..
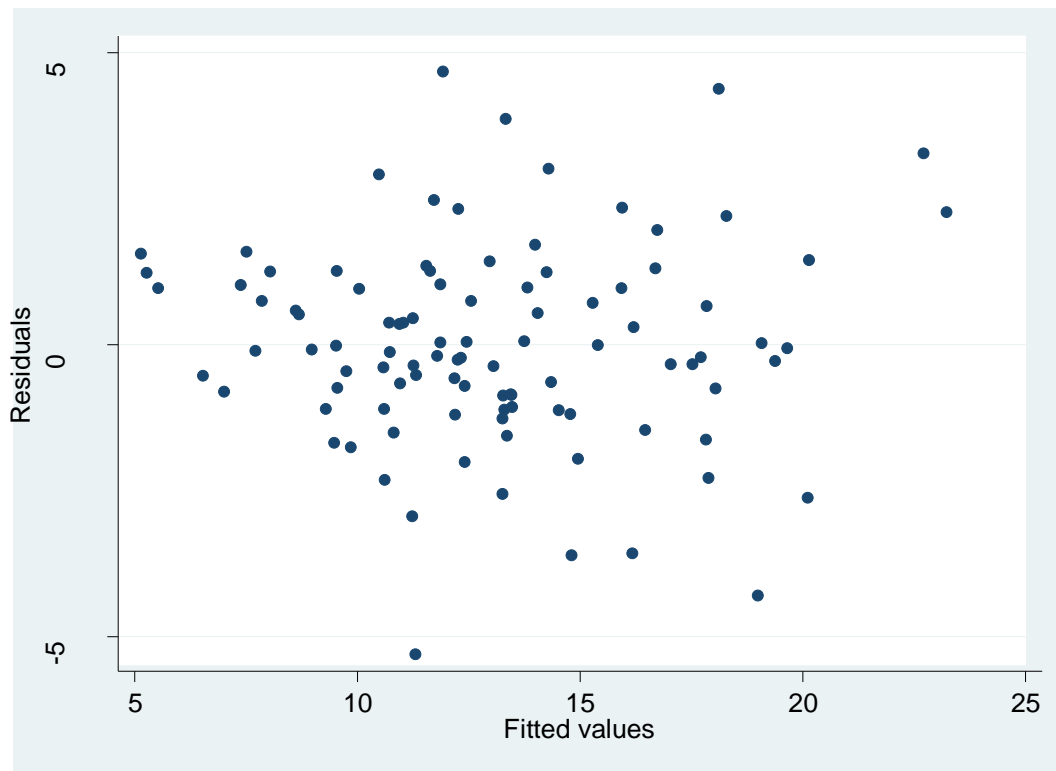
To mitigate this problem of heteroscedasticity, the dependent variable of the log of the poverty rate was run against the same regressors, with the same resulting significance and with the assumption of homoscedasticity no longer rejected at the alpha = 0.05 level. Significance levels were p =0.2153 and p=0.2509 for the Breusch Pagan and White tests respectively.

Since heteroscedasticity does not generally introduce bias into the estimates of the model coefficients, and all the final model variables were found to maintain their significance with robust variances and also when tested against the log of the poverty rate (which did not exhibit significant heteroscedasticity) we are reporting our final model in its original form for its ease of interpretation and communication.

Conclusions

In this study, we found that almost 80% of the variability in poverty rates between counties in North Carolina were accounted for by a set of four explanatory variables; the percent of the population in the labor force, percent of adults who had not completed high school, percent labor force participation, and the proportion of the population in the manufacturing industry.

The first level of further work would be to attempt to replicate more precisely the results of Levernier et al. by using the entire United States data set. Some of the variables that they found to be significant may also show up in the 2010 data by using the larger data set. They also looked at economic disruption as a causal factor of poverty, which we did not do in this study. It would be interesting to fully replicate their work to identify any change in the 20 years since their study was carried out.

In future work, I would also like to explore some alternative measures of poverty. In many developed countries, poverty is defined as an income level less than 50% of the median.(Smeeding et al, 2001) It would be interesting to find out how that metric would look in the United States. Since the poverty line in the US is known to be out-of-date and not truly representative of current conditions, other researchers have also been looking into issues such as the extent to which in-kind programs such as food stamps (SNAP) and the local cost of living affect whether a household can still be considered impoverished after in-kind programs are taken into account.

Other researchers have looked into the persistence of poverty. They have found that a large minority of impoverished households are in poverty for only a single year, and only 12% were impoverished for the several years of the study's range (Rodgers, 1993). It would also be interesting to look at the persistence of poverty over time for individual household data in a longitudinal study.

In short, there are many directions that the study of poverty can take, and much that can be learned to inform policy decisions.

References

Andy, Emmanuel; Critical Analysis of Poverty Theories;

http://www.academia.edu/1116707/Critical_Analysis_of_Poverty_Theories; (2011) AN71022C-

33213062-2011-12

Bertrand, Marianne, Sendhil Mullainathan, and Eldar Shafir. "A behavioral-economics view of

poverty." *American Economic Review* (2004): 419-423.

FILES: 2006-2010 American Community Survey [North Carolina]/prepared by the U.S. Census Bureau,

2011 http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml .

Levernier, William, Mark D. Partridge, and Dan S. Rickman. "The Causes of Regional Variations in US

Poverty: A Cross-County Analysis." Journal of Regional Science 40, no. 3 (2002): 473-497.

Mills, Bradford F., and Gautam Hazarika. "Do Single Mothers Face Greater Constraints to Workforce

Participation in Non-Metropolitan Areas?." *American Journal of Agricultural Economics* 85.1 (2003): 143-

161.

 Piketty, Thomas and Emmanuel Saez. "Income Inequality In The United States, 1913-1998," Quarterly

Journal of Economics, 2003, v118(1,Feb), 1-39.

Rodgers, Joan R. and Rodgers, John L.; "Chronic Poverty in the United States"; The Journal of Human

Resources; Vol. 28, No. 1 (Winter, 1993), pp. 25-54.

Rothstein, Jesse. *The Labor Market Four Years Into the Crisis: Assessing Structural Explanations*. No.

w17966. National Bureau of Economic Research, 2012.

Social Science Research Institute (SSRI), "North Carolina in the Global Economy", Duke University (2012)

http://www.soc.duke.edu/NC_GlobalEconomy/index.shtml

Smeeding, Timothy; Rainwater, Lee; and Burtles, Gary, "United States Poverty in a Cross-National Context" (2001). Center for Policy; Research. Paper 151. http://surface.syr.edu/cpr/151

Voss, Paul R., David D. Long, Roger B. Hammer, and Samantha Friedman. "County child poverty rates in the US: a spatial regression approach." Population Research and Policy Review 25, no. 4 (2006): 369-391.