Proceedings of the 2nd International Conference on Business and Policy Studies
DOI: 10.54254/2754-1169/17/20231121

The Effect of Basic Education, Health, and Racial Factors on the Spatial Distribution of Poverty Rate in Los Angeles County

Jianrui Liu[1,a,*]

[1]*University of California, Santa Barbara CA 93106, USA*
a. jianrui@ucsb.edu
**corresponding author*

Abstract: This work focuses on the spatial distribution of poverty rates in Los Angeles County and the impact of multiple socioeconomic factors on poverty rates and distribution. The scale of the study is based on census tract to ensure the fineness of the study. The three socioeconomic factors are basic education, health, and race. The primary research method was to perform regression analyses on socioeconomic factors and poverty rates. Initially, a conventional Ordinary Least Square regression analysis was conducted, and it was found that this model did not explain the distribution of poverty rates by region in Los Angeles County well. Therefore, a geographically weighted regression model was applied to explain the relationship between different socioeconomic factors and poverty rates. It was found that basic education has the effect of reducing local poverty in relatively poor areas; while for wealthier areas, merely receiving basic education is insufficient, which leads to higher poverty rates. Health insurance is more of a burden for relatively poor areas, thus worsening poverty, while for relatively affluent areas, health insurance can help alleviate local poverty by avoiding high medical bills. In areas where one ethnic group is concentrated, an increase in the proportion of one ethnic group has a positive effect on reducing the local poverty rate, but in areas where other ethnic groups are concentrated, it has a negative effect on reducing the local poverty rate.

Keywords: geographically weighted regression, poverty rate, Los Angeles County

1. Introduction

According to the U.S. Census Bureau's income and poverty report [1], the official poverty rate in 2020 was 11.4 percent, up 1.0 percentage point from 10.5 percent in 2019. This is the first increase in poverty after five consecutive annual declines. In 2020, there were 37.2 million people in poverty, approximately 3.3 million more than in 2019. Between 2019 and 2020, the poverty rate increased for non-Hispanic Whites and Hispanics. Among non-Hispanic Whites, 8.2 percent were in poverty in 2020, while Hispanics had a poverty rate of 17.0 percent. Among the major racial groups examined in this report, Blacks had the highest poverty rate (19.5 percent) but did not experience a significant change from 2019. The poverty rate for Asians (8.1 percent) in 2020 was not statistically different from 2019. From the above data, we can know poverty has been a complex social problem in the US, meanwhile, the appearance of poverty typically results from multiple factors, such as education, race,

DOI: 10.54254/2754-1169/17/20231121

health, food poverty, etc. Poverty has a significant impact on the children and teenagers, it usually results in juveniles dropping out of high school and are likely to have employment problems. Although only 1 percent of children who are never poor end up being poor as young adults, 32 percent of poor children become poor as young adults [2]. Their lack of education and other public resources in turn restricts them and their children to poverty, once again helping to ensure a vicious cycle of continuing poverty across generations. Also, low-income families tend to have low intakes of fruit and vegetables and high intakes of junk food. They also tend to suffer more health problems from cancer, diabetes, obesity, and heart disease, which result in higher mortality rates, higher prevalence of acute or chronic diseases, and more emotional and behavioral issues [3]. So, it is hard for people who are already restricted in poverty to get out of it. According to the report, Social Security kept 26.5 million people out of poverty in 2020, and economic impact or stimulus payments kept 11.7 million people out of poverty. Unemployment insurance payments and refundable tax credits kept 5.5 million and 5.3 million out of poverty, respectively [4]. We can see the government has a leading role in improving poverty. Therefore, the government and other poverty departments should actively respond to the poverty situation and make new policies and programs to reduce poverty growth. My research purpose is to help the government to have a thorough understanding of the urban poverty issue in LA so that they can respond to urban poverty issues accordingly, and help more people get out of poverty and get a better life in the future. My major study question is how the spatial distribution of the poverty rate is affected by basic education, health, and racial factors. More specifically, how do the number of secondary schools, education attainment, health insurance coverage, and percentage of specific ethnic groups influence the local poverty rate?

2. Data

To analyze and interpret the research questions, I obtained data from the US Census Bureau on poverty rates as well as basic education, health, and race. I attempted to regress them to find the relationship between the distribution of poverty rates and these factors. Since the study area is limited to Los Angeles County, I needed to keep the scale of the data as small as possible to provide a more refined analysis. The smallest administrative unit of the US Census Bureau data is the census tract, so all my data are obtained at the census tract scale. Also, all data were collected in 2019 and all have 2341 rows, corresponding to 2341 census tracts in Los Angeles County. Table 1 is a detailed list of the data.

Table 1: Data.

Dimension	Variables	Details
Poverty (Dependent Variable)	Poverty Rate	The percentage of the population below the poverty line among the total population.
Basic Education	Middle School Count	The number of middle schools within a 3km range of the census tract.
	High School Count	The number of high schools within a 3km range of the census tract.

DOI: 10.54254/2754-1169/17/20231121

Table 1: (continued).

	Middle School Enrollment	The percentage of the population who have enrolled in schools among those aged 10-14.
	High School Enrollment	The percentage of the population who have enrolled in schools among those aged 15-17.
	Middle School Degree 18-24	The percentage of the population who have less than a high school degree among those aged 18-24.
	High School Degree 18-24	The percentage of the population who only have a high school degree among those aged 18-24.
	Middle School Degree 25+	The percentage of the population who have less than a high school degree among those aged 25 or older.
	High School Degree 25+	The percentage of the population who only have a high school degree among those aged 25 or older.
Health	Uninsured Rate	The percentage of the population who do not have health insurance among the total population.
Race	White Ratio	The percentage of the population who are white among the total population.

DOI: 10.54254/2754-1169/17/20231121

Table 1: (continued).

African American Ratio	The percentage of the population who are African Americans among the total population.
Native American Ratio	The percentage of the population who are native Americans among the total population.
Asian Ratio	The percentage of the population who are Asian among the total population.
Pacific Islander Ratio	The percentage of the population who are Pacific Islanders among the total population.
Other Ratio	The percentage of the population who are of other races among the total population.

In addition, the number of high schools and the number of middle schools are the only ones not obtained from the US Census Bureau. I use the 2019 California school shapefile obtained from the California Geoportal and perform a spatial join to obtain the number of middle and high schools in the three kilometers around each census tract. Although other data have been considered in other similar studies, I will focus on these variables and examine their relationship with the distribution of poverty rates in Los Angeles County.

3. Method

3.1. Ordinary Least Square

Since the dependent variable of poverty rate is a percentage form of data, and the regression analysis model commonly used for percentages as the dependent variable is usually General Linear Regression, but according to the "Spatial Regression Analysis of Poverty in R" [5], their dependent variable is also the poverty rate and they use the standard regression model, which is the Ordinary Least Square model. Therefore, I will first use the OLS model to explore the relationship between the explanatory variables and the poverty rate. First, I selected all the explanatory variables and analyzed them using the OLS model in ArcGIS Pro, but not all the explanatory variables were significant. In this case, Middle School Degree 18-24, High School Degree 18-24, and Pacific Islander Ratio had p-values of 0.24, 0.99, and 0.94. This means that these three explanatory variables did not have a significant effect on the change in the poverty rate. I excluded these explanatory variables from the OLS analysis, and the obtained results are shown in Table 2.

DOI: 10.54254/2754-1169/17/20231121

Table 2: OLS coefficient.

Variable	Coefficient	Standard Error	t_Stat	P-value	Adjusted R-square	AICc
Middle School Count	-0.17807	0.075879	-2.34672	0.019007		
High School Count	0.249246	0.044683	5.578139	8.94E-08		
Middle School Enrollment	-0.02811	0.013874	-2.0258	0.042889		
High School Enrollment	-0.05159	0.01241	-4.15754	3.9E-05		
Middle School Degree 25+	0.531888	0.041685	12.75983	0		
High School Degree 25+	-0.15161	0.025918	-5.84982	2.7E-08	0.4582947	16307.5847
Uninsured Rate	0.462382	0.037355	12.3782	0		
White Ratio	0.159708	0.024243	6.587893	1.06E-09		
African American Ratio	0.258257	0.027841	9.275991	1.9E-14		
Native American Ratio	0.551268	0.056586	9.742173	3E-15		
Asian Ratio	0.182184	0.025852	7.047149	1.46E-10		
Other Ratio	0.15738	0.028437	5.53431	1.08E-07		

From the table we can see the p-value of all the explanatory variables is less than 0.05, which means that all these variables can explain the changes in the poverty rate. For example, Middle School Degree 25+ has a coefficient of 0.531888, a standard error of 0.041685, a t-test statistic of 12.75983, and a p-value of 0. This means that when all other explanatory variables are held constant, a 1% increase in the percentage of people with a middle school education among those aged 25 and older is associated with a 0.5312% increase in the poverty rate. Also, the p-value is very small; this then proves that the regression coefficient for the percentage of people with a middle school education among those aged 25 and older is not zero. In addition, the results include an adjusted R-squared of 0.4582947 for this model, indicating that about 45.8% of the change in the poverty rate can be explained by the relationship with the explanatory variables considered in this study.

DOI: 10.54254/2754-1169/17/20231121

3.2. Geographically Weighted Regression

In the previous section, I used an OLS regression model to analyze the relationship between poverty rates and explanatory variables. However, we did not consider the effect of spatial information on the relationship between these explanatory variables and poverty rates. To investigate whether the relationship between these explanatory variables and poverty rates varies with space, I used the Spatial Autocorrelation tool in ArcGIS Pro to analyze the residuals obtained from the OLS regression model. The results show that the residuals from the OLS analysis have a z-score of 27.626255, and there is a less than 1% likelihood that this clustered pattern has a spatial autocorrelation with each other, which would suggest that OLS is not the best model to study the relationship between poverty rates and explanatory variables. Thus, we need to consider spatial regression models for further analysis of the variables.

Before performing the spatial regression analysis, I need to perform exploratory regression analysis on the explanatory variables to select the desired explanatory variables again. As in "Spatial Regression Analysis of Poverty in R" [5], I took a similar approach. The explanatory variables used in the previous section were analyzed. The results showed that high school enrollment, the ratio of native Americans, and middle school enrollment had significant percentages of 8.76%, 1.64&, and 0.07%, indicating that these three explanatory variables did not have a significant effect on the change in the poverty rate, so I removed these three variables to obtain a more concise and efficient model. As mentioned by Kamenetsky [5] I need to perform model selection to find more parsimonious models that may have fewer explanatory variables. Also, I need to avoid multicollinearity between explanatory variables and thus reduce the redundancy of the data. In figure 2, it is shown that there is multicollinearity between ratio white and ratio other, so I eliminate ratio other for the following spatial regression analysis.

Summary of Variable Significance

Variable	% Significant	% Negative	% Positive
HIGH_COUNT	100.00	0.00	100.00
NOHIGH_25	100.00	0.00	100.00
UNINSURED	100.00	0.00	100.00
RATIO_AFRI	91.54	3.63	96.37
RATIO_OTHE	64.45	21.14	78.86
HIGH_25	56.66	67.31	32.69
MID_COUNT	55.56	32.61	67.39
RATIO_WHIT	49.19	67.29	32.71
RATIO_ASIA	41.87	53.27	46.73
HIGH_ENROL	8.76	98.73	1.27
RATIO_NATI	1.64	0.00	100.00
MID_ENROLL	0.07	57.43	42.57

Summary of Multicollinearity

Variable	VIF	Violations	Covariates
MID_COUNT	2.89	0	--------
HIGH_COUNT	2.75	0	--------
MID_ENROLL	1.73	0	--------
HIGH_ENROL	1.60	0	--------
NOHIGH_25	2.63	0	--------
HIGH_25	1.78	0	--------
UNINSURED	2.17	0	--------
RATIO_WHIT	9.34	276	RATIO_OTHE (1.16)
RATIO_AFRI	4.31	0	--------
RATIO_NATI	1.19	0	--------
RATIO_ASIA	6.32	0	--------
RATIO_OTHE	8.10	23	RATIO_WHIT (1.16)

Figure 1. & Figure 2.

Finally, I perform geographic weight regression analysis using the remaining explanatory variables with poverty rates to explore whether and how the relationship between explanatory variables and poverty rates changes with spatial relationships. In contrast to the spatial autoregressive regression used by Kamenetsky [5] and Liu [6] in their articles, I choose to use the geographically weighted regression tool in ArcGIS Pro to analyze the poverty rate with the explanatory variables and generate a series of visualizations to help better interpret results. Finally, the residuals generated by the GWR analysis were analyzed by spatial autocorrelation to explore whether their distribution was random and to see if GWR could better explain the relationship between the variables and the variation in poverty rates. The results show that the GWR-resulting residuals have a z-score of 0.263197, which means that they are randomly distributed and do not show significant patterns.

DOI: 10.54254/2754-1169/17/20231121

4. Results

4.1. Model Comparison

In the previous section I analyzed two regression models for the explanatory variables and poverty rates, an OLS regression model that does not consider the spatial component and a GWR model that considers the variation in the spatial component. To measure the performance of the two models, I have selected four important indicators to compare them. Moreover, the GWR model performs the best when there are 123 neighbors. These are adjusted R-square, AIC, MAE, RMSE, and p-value of the global Moran's I. The specific results obtained are in Table 3 below.

Table 3: Model comparison.

Model	AIC	Adjusted R-square	Moran's I
Ordinary Least Square	16307.3783048	0.458295	0
Geographically Weighted Regression	15495.5112	0.7480	0.792399

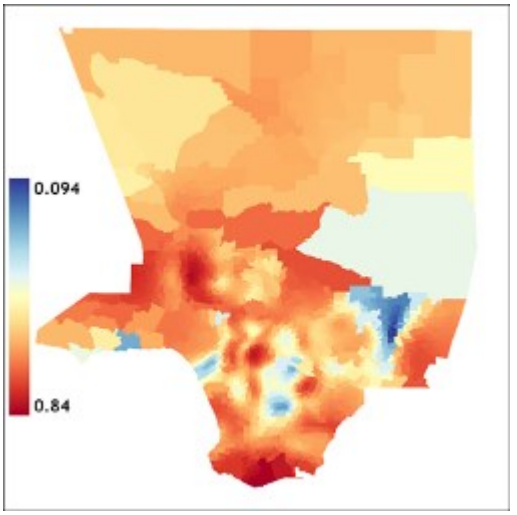

Figure 1: GWR local R^2 map.

From the table we can see that GWR outperforms OLS in any respect, where the AIC of the GWR model is much smaller compared to OLS, and the adjusted R-square and p-value of the global Moran's I of GWR have a significant improvement over OLS. From the local R-square map produced by GWR analysis, it shows high R-square values in most regions in Los Angeles County. These indicate that the data can better fit the GWR model rather than the OLS model, so I will next interpret the results obtained from the GWR model.

4.2. GWR Results

By analyzing the poverty rate with each explanatory variable, we ended up with eight coefficient maps that represent the impact of the variation of each explanatory variable poverty rate in different regions. The nine maps in the following sections include the coefficient maps of the eight explanatory variables with the poverty rate and the visualization maps of the distribution of the poverty rate in Los Angeles County.

DOI: 10.54254/2754-1169/17/20231121

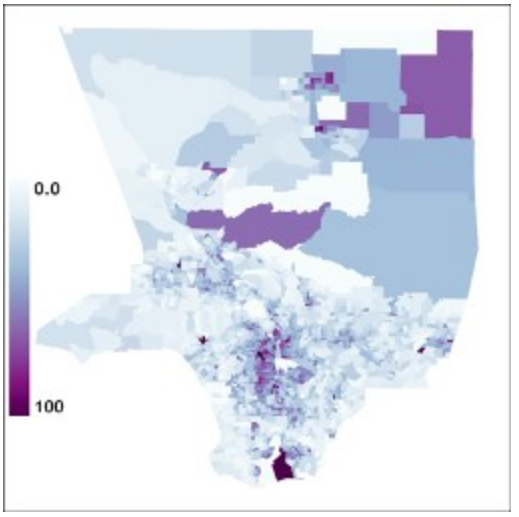

Figure 2: Poverty distribution map of Los Angeles County.

In the GWR coefficient maps, the redder parts indicate that an increase in this explanatory variable tends to have a decrease in the local poverty rate, and conversely the bluer areas indicate that an increase in this explanatory variable tends to have an increase in the local poverty rate. At the same time, the lighter the color means that the explanatory variable has a smaller effect on the poverty rate in this region. The results are then interpreted in terms of basic education, health, and race.

4.3. Basic Education

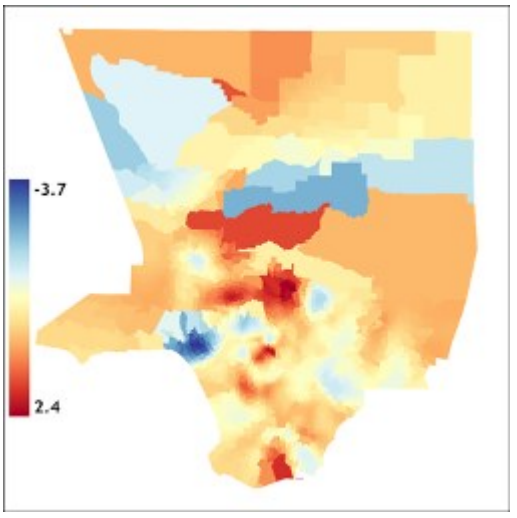

Figure 3: Coefficient map of middle school count.

DOI: 10.54254/2754-1169/17/20231121

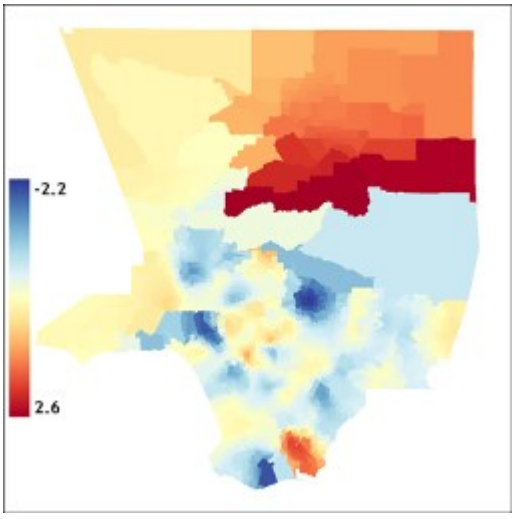

Figure 4: Coefficient map of high school count.

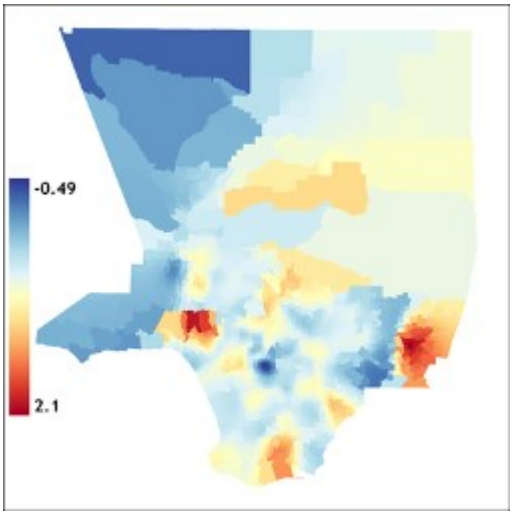

Figure 5: Coefficient map of middle school degree 25+.

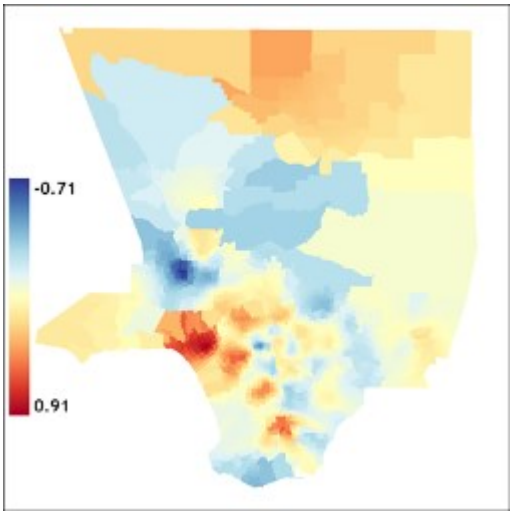

Figure 6: Coefficient map of high school degree 25+.

DOI: 10.54254/2754-1169/17/20231121

Looking at the level of education, we have four explanatory variables: the number of middle and high schools, and the percentage of people aged 25 and older who have middle and high school degrees. The number of middle schools shows a small effect on the poverty rate on most scales, but there are still some regions that show a large effect. In the Santa Monica and Hollywood areas, an increase in the number of middle schools tends to have a lower local poverty rate, while in Long Beach, Downtown, and areas north of Downtown, an increase in the number of middle schools tends to have a higher local poverty rate. Contrary to the number of middle schools, the increase in the number of high schools tends to have a lower local poverty rate in most areas. Comparing the map of poverty distribution, we can see that the increases in the number of middle schools correspond to the increases in the poverty rate in areas with high poverty rates, while in areas with low poverty rates, the increases in the number of middle schools correspond to the decreases in the local poverty rate. As for the number of high schools, an increase in the number of high schools tends to have a lower local poverty rate in both high and low poverty areas. This is more in line with our general perception of the improvement of poverty rates through basic education. This is because, in wealthier regions, people do not usually end up with lower secondary education, while in poorer regions, a certain percentage of the population stops studying after lower secondary education. Such differences create differences in the educational attainment of the population, which leads to differences in poverty rates. In contrast, high school education equips people to survive in society, so the number of high school schools has a positive effect on improving poverty in both poor and affluent regions.

In terms of educational attainment for those aged 25 and older, the higher percentage of people with a middle school education corresponds to a lower poverty rate in most regions and corresponds to the higher poverty rate only in the Beverly Hills region and in a few regions in the East. The effect of the percentage of people with a high school education on the poverty rate varies more spatially, as the higher percentage of people with high school education in areas with higher poverty rates corresponds to a lower local poverty rate, while the opposite is true in areas with lower poverty rates. In general, the proportion of people with a basic level of education can help improve poverty rates, but in a small number of wealthier regions, simply having a relatively low level of education can increase local poverty rates. This is consistent with our understanding that basic education does have a positive effect on poverty but having only a basic level of education in wealthier regions makes it difficult to find suitable jobs.

4.4. Health

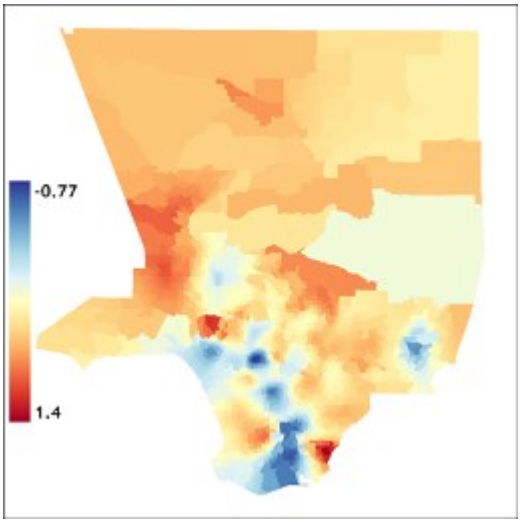

Figure 7: Coefficient map of uninsured population.

DOI: 10.54254/2754-1169/17/20231121

In terms of health, the increase in the percentage of people who do not have health insurance corresponds to lower local poverty rates in regions with high poverty rates; while in most other regions, the increase in the percentage of people who do not have health insurance tends to have higher local poverty rates. This may be because, in regions with high poverty rates, people cannot afford health insurance, or an additional expense would exacerbate their poverty. In other regions with lower poverty rates, having health insurance does not put much of a burden on people's lives, and even avoids high medical expenses at some levels. Without health insurance, these people may not be able to afford the high cost of health care and thus become poor.

4.5. Race

For the results on race, I combined the poverty rate with the percentage of each ethnicity in the population to interpret the results. The following six figures show the coefficient maps and the distribution maps of whites, African Americans, and Asians in Los Angeles County.

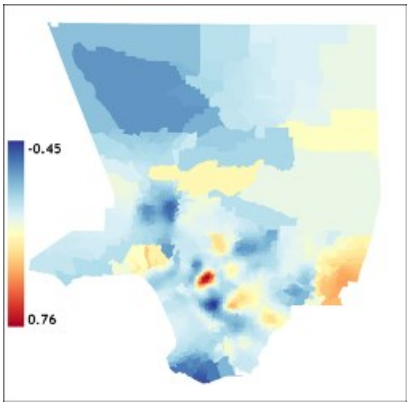

Figure 8: Coefficient of white ratio.

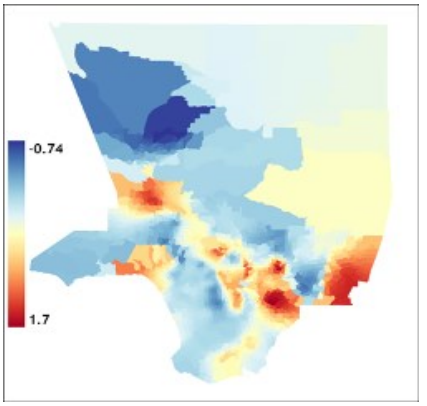

Figure 9: Coefficient of black ratio.

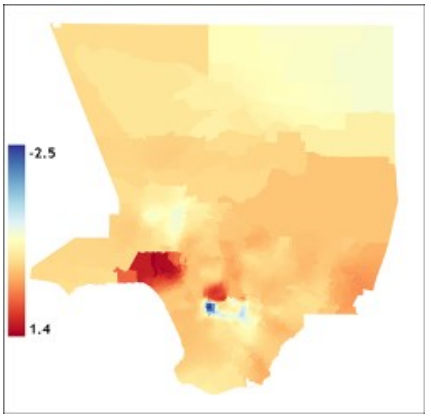

Figure 10: Coefficient of Asian ratio.

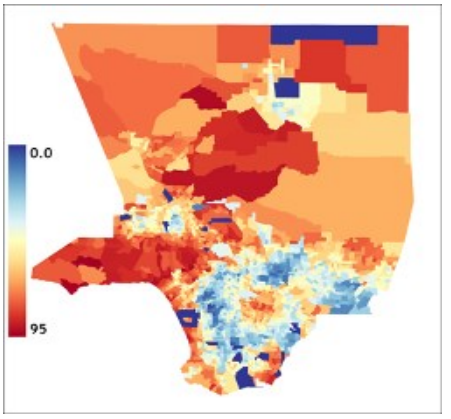

Figure 11: Distribution of whites.

DOI: 10.54254/2754-1169/17/20231121

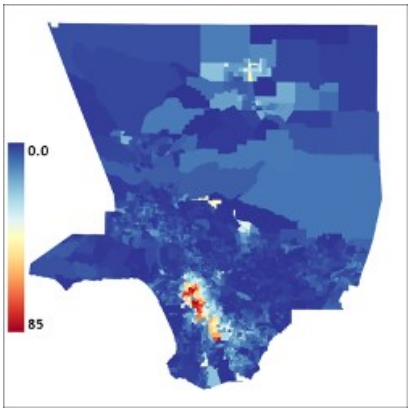

Figure 12: Distribution of black.

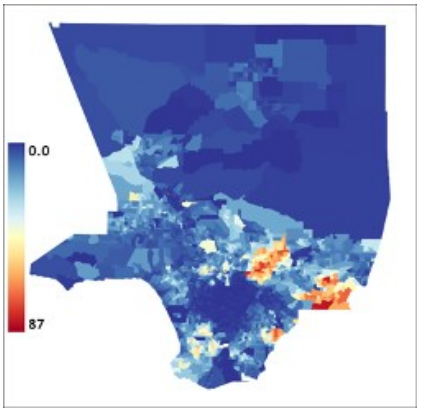

Figure 13: Distribution of Asians.

In terms of race, the increase in the proportion of Whites in the population tends to have a reduction in the local poverty rate in most regions; however, in the Downtown area of Los Angeles, the increase in the proportion of Whites in the population corresponds to a higher local poverty rate. This may be because people in Downtown, an area with high poverty rates, are also among the relatively poorer segment of the white population, and therefore do not have a positive effect on reducing poverty rates.

The effect of the African American proportion of the population on the poverty rate varies more significantly spatially. Combining the poverty rate with the distribution of African Americans, we can find that in areas with a high percentage of African Americans, an increase in the percentage of African Americans in the population tends to have a lower local poverty rate; while in areas with a high percentage of other minority populations, an increase in the percentage of African Americans in the population corresponds to a higher local poverty rate. This may be due to the inability of African Americans to adapt well to specific minority cultures in these regions, making it relatively difficult to find suitable jobs.

Finally, the percentage of Asians has a relatively small effect on changes in poverty rates in most regions and a positive or negative effect on improving poverty rates in only a few regions. In areas with a high white population, an increase in the Asian population corresponds to higher poverty rates, while in a small number of areas with a high African American population, an increase in the Asian population tends to have a lower poverty rate. This may be since Asians have lower income levels relative to whites, which does not help improve poverty rates locally, while for African Americans, Asians have relatively higher incomes, which have a relatively positive impact on improving poverty rates locally.

4.6. Conclusion

In general, basic education has the effect of reducing local poverty in relatively poor areas; while for wealthier areas, merely receiving basic education is insufficient, which leads to a higher poverty rate. For instance, the increases in the number of middle schools correspond to the increases in the poverty rate in areas with high poverty rates, while in areas with low poverty rates, the increases in the number of middle schools correspond to the decreases in the local poverty rate. Health insurance is more of a burden for relatively poor areas and thus worsens poverty, while for relatively affluent areas, health insurance can help alleviate local poverty more by avoiding high medical costs. The race is more complex, with the percentage of whites and African Americans having a significant impact on changes in poverty rates, but the percentage of Asians having a relatively small impact. In general, however, an increase in the percentage of a particular ethnic group has a positive effect on reducing

DOI: 10.54254/2754-1169/17/20231121

local poverty rates in areas where that ethnic group is concentrated, but in areas where other ethnic groups are concentrated, it has a negative effect on reducing local poverty rates.

5. Challenges/Gaps and Future Direction

In this analysis, I only considered the effects of basic education, health, and ethnicity on the change in poverty rate at the data level. However, more factors may have an impact on the change in poverty; therefore, in future studies, I may consider adding other factors in other dimensions to study the factors that have an impact on the poverty rate more comprehensively. Meanwhile, among the three dimensions, only considering the effect of basic education on the poverty rate may not be sufficient, and higher education and pre-school education can be added in the future. At the same time, only one factor was selected for the health dimension, which does not reflect the comprehensive effect of health factors on the poverty rate. At the methodological level, I used only two regression models, OLS and GWR, but it is possible that there are better models than these to analyze the effect of explanatory variables on the poverty rate. Therefore, in future studies, other spatial regression models, such as spatial lag regression models, could be considered. Overall, this study explains the effects of basic education, health, and ethnicity on the spatial distribution of poverty rates to some extent, but there are still many shortcomings that could be improved to provide a more thorough analysis of the factors that influence the spatial distribution of poverty rates.

References

[1] *Emily, A., Shrider, K., Frances, C., Jessica, S.: Income and Poverty in the United States: 2020, https://www.census.gov/library/publications/2021/demo/p60-273.html ,last accessed 2022/04.*

[2] *Childhood Poverty Persistence: Facts and Consequences, https://www.urban.org/sites/default/files/publication/32926/412126-childhood-poverty-persistence-facts-and-consequences.pdf , last accessed 2022/04.*

[3] *Bill, F.: Poverty in the United States, https://www.debt.org/faqs/americans-in-debt/poverty-united-states/ ,last accessed 2022/04.*

[4] *U.S. Census Bureau.: Income, Poverty and Health Insurance Coverage in the United States: 2020, https://www.census.gov/newsroom/press-releases/2021/income-poverty-health-insurance-coverage.html ,last accessed 2022/04.*

[5] *Kamenetsky, M., Chi, G., Wang, D., Zhu, J.: Spatial Regression Analysis of Poverty in R. Spatial demography, vol. 7(2-3), pp. 113–147 (2019).*

[6] *Liu, W., Li, J., Zhao, R.: The effects of Rural Education on Poverty in China: A spatial econometric perspective. Journal of the Asia Pacific Economy, 1–23 (2021). https://doi.org/10.1080/13547860.2021.1877240*