Spatial Analysis of Poverty: The case of Peru

Delgado Narro, Augusto Ricardo

Waseda University

23 May 2020

Spatial Analysis of Poverty: The case of Peru

Augusto Ricardo Delgado Narro †

May 2020

Abstract

The concept of Multidimensional Poverty traditionally was used for comparative analysis across regions or countries. Nevertheless, in this paper, we use the concept of Multidimensional Poverty calculated for each Peruvian region and analyze the spatial patterns and spatial autocorrelation observed across the country, and, later, identify the spatial spillovers in poverty across the country. We find evidence of spatial autocorrelation across the regions which are statistically significant across models; in other words, it means the poverty has spatial effects. Additionally, we find strong and significant evidence of spatial spillovers originated in the error terms rather than the endogenous variable, which has unstable effect. Finally, the set of covariates we use in our regressions are statistically significant and stable across the models.

Keywords: Poverty, Spatial econometrics, Peru

JEL classification: C21, O10

1 Introduction

Poverty is probably the most important concern of countries around the world, due to this importance, they established the reduction of extreme poverty as the first goal on the Millennium Development Goals (MDG) (Khawas, 2006). Following this goal, the government of Peru has been setting policies to reduce poverty and extreme poverty along the country. Since 2004 to 2015, the poverty was reduced from 58 to 22 per cent, while the extreme poverty fell from 16 to 4 per cent during the same period. In absolute numbers, nine million of Peruvians were able to escaped from poverty.

†Faculty of Political Science and Economics, Waseda University
Address:
Email: adelgado@aoni.waseda.jp

The concept of poverty is widely investigated in the literature, Sen (1976) argues that the concept of poverty is not only an income problem but it is composed by different dimensions that together compose a more encompassing concept of poverty. In other words, a person is not poor only because he does not have enough income to fulfill his needs, it is poor because his chances to escape from poverty are reduced due to insufficiency in education, health, and life quality. Nevertheless, the concept of poverty is not only a multidimensional concept, it is a spatial concept as well. Poverty is typically agglomerated in certain areas, regions with low quality infrastructure, regions where government policies do not reach, regions where people work in certain types of jobs. In the case of Peru, the poverty rates are higher in the highlands and the jungle regions, where indigenous population is concentrated and the basic public infrastructure (roads, schools, hospitals, security) is insufficient.

Taken these ideas into account, this paper tries to analyze the poverty in Peru by using Census data of 2017. Since the Census recollected information about population and their living conditions but not income information, this paper analyze poverty based in households living conditions rather than the classic monetary perspective by constructing the "Multidimensional Poverty Index" (MPI, hereinafter). Therefore, the main objectives of this research are:

1. Identify spatial patterns of poverty across the country.
2. Identify the presence of spatial autocorrelation and clusters of poverty among regions.
3. Find evidence of spatial spillovers across districts among Peru.

2 Literature Review

Wallender et al. (2011) use municipal level data to identify spatial patterns in rural areas in Brazil. By using Moran's I indicator, they are able to identify "hot spots" and "cold spots", it means, areas where the poverty is agglomerated or dispersed, respectively. They find that there is evidence of clusters among municipalities and the poverty reduction policies must be taken into account when those clusters are identified. However, those clusters are not limited by the state's boundaries.

Mussagy and Dambudzo (2015) analyze the poverty and road network situation in the Sofala province in Mozambique with the objectives of making a spatial analysis of poverty and relate it with the road networks. They use the National Household Survey of Living conditions and the Population Census and find that even when the policies reduced the poverty levels, it did not disappear completely. Additionally, they show a strong correlation between high poverty levels and weak road network.

Bigman and Fofack (2000) highlight the use of geographical data to analyze poverty and create poverty alleviation policies. They identify five advantages of the usage of geographical data for poverty alleviation. First, it provides clear criteria for identifying the target population; second, easy to monitor and administer; third, it has influence on a household's behavior; fourth, it is possible to improve targeting by combining with other criteria; and fifth, they can include not only direct income transfer but also other means to increase living standards.

Akinyemi and Bigirimana (2012) seek for emerging poverty patterns based on household living conditions in Kigali city in Rwanda; also, they look for the contribution of four indicators over the poverty, those are: expenditure, health, education, and services. For this purpose, they use data from the Integrated Living Condition survey between 2000-2001. In their findings, they show poverty patterns and the presence of urban-rural dichotomy.

Brunn and Wheeler (1971) by doing a geographical and factor analysis, they try to identify various faces of poverty among US counties based on data from the County and City Data Book of 1967. They use information for socioeconomic status, agricultural productivity, demographic composition, agricultural holdings and investment, and the degree of urbanization and manufacturing. They find that these factors have different importance once a county measures its poverty level. Therefore, some countries have similar poverty levels but the source of poverty differs among them.

Crandall and Weber (2004) focus on the analysis of the effect of job growth and social capital over the poverty rates, additionally, they identify the presence of spatial spillovers. By using two census tract-level data across US, the results suggest that job growth has poverty-reducing effect, social capital has strong effect for reducing poverty levels in high-poverty level areas. Finally, there is evidence of spatial spillover in poverty across areas.

Chen et al. (2015) combine spatial statistical analysis and GIS information to identify patterns and factors of spatial poverty distribution in Xianfeng in China. For this purpose, they use two key indicators, poverty headcount ratio and the per capita net income of poverty population. They find evidence of positive spatial autocorrelation and agglomeration of poverty levels across the county.

Rupasingha and Goetz (2007) investigate the determinants of poverty in US at county level data on 1999 by using spatial analysis techniques. They show that social capital, ethnic and income inequality, local political competition, federal grants, foreign-born population, and spatial effects are important determinants to explain poverty levels across the country.

Tanaka and Lee (2011) combines district level poverty rates, population census data, income data, and geospatial data in Ghana to investigate the impact over poverty of factors as human capital, structural change, infrastructure, and environmental degradation. They find that working age population, employment, and the service sector are key factors in order to reduce the poverty levels. Additionally, electricity, roads, and market conditions are important to improve the income level in non-agricultural sectors. Finally, all these findings are correlated with spatial patterns where is possible to identify the presence of "hotspots" or poverty agglomeration.

Grab (2009) analyzes the spatial income disparities among households in Burkina Faso by using three main nation-wide household surveys in 1994, 1998, and 2003. The author shows that spatial disparities are not only driven by spatial concentration of households with particular endowments but to a large disparity in those endowments. It means, communities are not only poor because households there are poor but also because the endowment of these communities are weak.

Holt (2008) uses socioeconomic and health-related data at county level in US from the Community Health Status Indicators database in order to describe spatial analysis of poverty across the country for 2000. The findings reveal significant and stark patterns of poverty that the author describes as "continental poverty divide". The poverty levels in US are concentrated in the south-east counties while low poverty levels are located on the north-west counties.

3 Methodology

3.1 Data

Three kind of data have been used for this research and build the "Multidimensional Poverty Index" (Odekon, 2015), all belonging to the National Census 2017 along Peru. These three types of data are the Housing Characteristics and Services, Households' characteristics, and Population Characteristics. Therefore, the MPI is built by using five dimensions: Education, Childhood and youth, Health, Employment, and Household; they are weighted on the following way:

$$MPI = 0.2(Education) + 0.2(Childhood) + 0.2(Health) + 0.2(Employment) + 0.2(Housing) \tag{1}$$

The "Education" dimension is considered a principal factor for households to adapt themselves for social changing conditions along the time. This dimension is composed by two factors:

- *Educational achievements (educ1).* Based on the Population Characteristics database, we built a variable for the schooling years for each member of the household since the first grade in the elementary school. Then we get the average schooling years for all members older than 15 years old, then, if the average is less than 9 years of schooling, the household is considered deprived.

- *Illiteracy (educ2).* We count the number of members in each household older than 15 years old that is not able to read or write. Those households with at least one member falling in this condition, it is considered deprived.

"Childhood and youth" dimension is considered important since it is stage where the crucial capabilities and skills are developed in order to have citizens that are able to self-sufficient. During this stage, people have higher probabilities to get infected with some diseases. On the other hand, for many developing countries, schooling and child labor are risk factors since many households do not have enough income and need young member leave school and start to work earlier. This dimension is composed by four factors:

- *Educational lag (child1).* We apply a filter to pick up the member between 7 and 17 years old, then we build the educational lag variables by considering the following rule: 7 years old and do not have at least 1 year of schooling; 8 years old and do not have at least 2 years of schooling, 9 years old and do not have at least 3 years of schooling; up to 17 years old and do not have at least 11 years of schooling. Finally, we count the number of members in the household that fall under this condition; therefore, if there is at least one member under this condition, then the household is deprived.

- *School absenteeism (child2).* We count the member between 6 and 16 years old that are currently attending a school. If there is at least one person among these ages that is not attending any kind of school, the household is considered deprived.

- *Childhood Care (child3).* We count the members younger than 5 years old that are not having any kind of insurance (public or private) and do not go to any kind of education institution where he or she can get care support. If there is at least one member under this condition, the household is considered deprived.

- *Child labor (child4).* We count the member younger than 14 years old that are currently working to collaborate with the household's income. If there is at least one member falling under this condition, the household is deprived.

"Health" dimension is crucial since the governments' duty is to supply a minimum level of healthcare system for their citizens in order to assure people's conditions to follow their own objectives. The dimension is composed by the following unique factor:

- *Healthcare insurance (health).* We count the number of members older than 5 years old that are not affiliate to any health insurance system (public or private). If there is at least one member falling under this condition, the household is deprived.

“Employment” dimension is crucial to go over the poverty stage by having a job and not belong to the informal sector. These two conditions help to assure a proper income and have a job with all its benefits. The dimension is composed by the following factors:

- *Employment (employment1).* We count the members older than 14 years old that currently do not have job and have been looking actively for a job. If there is at least one member falling under this condition, the household is deprived.

- *Informality (employment2).* We count the members that are working in a company with five or less employees. If there is at least one member falling under this condition, the household is deprived.

“Housing” dimension since it creates minimum conditions where families and their members develop their daily-life activities allow them to access for basic tools for develop their capabilities. This dimension is composed by seven factors:

- *Water access (house1).* We consider a household is deprived if they do not have access to water service inside the house, inside the building, or from a public sink. Additionally, we consider a household is deprived if they do not have access to water less than three days per week by any mean.

- *Sewage access (house2).* We consider a household is deprived if they do not have access to any sewage service inside the house or inside the building.

- *Floor (house3).* We consider a household is deprived if the house’s floor material is other than parquet, tiles, vinyl, or cement.

- *Walls (house4).* We consider a household is deprived if the house’s wall material is other than bricks, stones, mud bricks, or wood.

- *Roof (house5).* We consider a household is deprived if the house’s roof material is other than concrete, wood, or tiles.

- *Public lighting (house6).* We consider a household is deprived if the house does not have access to any public lighting.

- *Overcrowding (house7).* We consider a household is deprived if the house has more than three members per room.

Finally, The MPI is built as follows:

$$MPI = 0.2(Education) + 0.2(Childhood) + 0.2(Health) + 0.2(Employment) + 0.2(Housing) \tag{2}$$

where:

$$\begin{aligned}
Education &= \frac{1}{2}educ1 + \frac{1}{2}educ2 \\
Childhood &= \frac{1}{4}child1 + \frac{1}{4}child2 + \frac{1}{4}child3 + \frac{1}{4}child4 \\
Health &= health \\
Employment &= \frac{1}{2}employment1 + \frac{1}{2}employment2 \\
Housing &= \frac{1}{7}house1 + \frac{1}{7}house2 + \frac{1}{7}house3 + \frac{1}{7}house4 + \frac{1}{7}house5 + 1/7house6 + \frac{1}{7}house7
\end{aligned}$$

3.2 Spatial Autocorrelation Test

Global spatial autocorrelation analysis is used to identify a situation in which the dependent variable or the error term at the interested location is correlated with observations on the dependent variable or the error term at other locations. In other words, it measures how related are the observations in a specific area respect to its neighbors. One of the most common tests for this analysis is Global Moran's I, described as:

$$I = \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(x_i - \bar{x})}{\sum_{j=1}^{n} W_{ij}(x_i - \bar{x})^2/2}, \qquad \forall i \neq j \tag{3}$$

where n is the number of spatial units of analysis indexed by i and j. x_i are the values of the variable x in the unit of analysis, while the $\bar{x}$ is mean of the variable x. W_ij refers to the weighted matrix $n \times n$ squared matrix that define the relationships among spatial neighbors. For this research, we use contiguity row standardized weight matrix based on "Queen" method, which is defined as two regions are neighbors if they share any part of a common border, no matter how short is it.

$$\begin{aligned}
W &= \begin{bmatrix} 0 & w_{12} & w_{13} & \dots & w_{1n} \\ w_{21} & 0 & w_{23} & \dots & w_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \dots & w_{nn} \end{bmatrix} \\
w_{ij}^* &= \frac{w_{ij}}{\sum_{j=1}^{n} w_{ij}}, \\
W &= \begin{cases} 1, & i \text{ neighbor } j, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

The Moran's I is interpreted as a coefficient of correlation with a range of $[-1, 1]$. Then, a positive and significant value of the indicator represents positive autocorrelation among the spatial units analyzed and higher values tend to indicates the presence of clusters. Similarly, negative and significant values of the indicator show negative autocorrelation and tend to indicate the presence of "hotspots". Finally, values close to zero indicate a random distribution of the variable analyzed among the spatial units of analysis.

On the other hand, Local spatial autocorrelation analysis is used to determine the local spatial autocorrelation of the variable for each spatial unit of analysis. Regard to the relationship with the global autocorrelation indicator, when there is no global spatial autocorrelation, local spatial autocorrelation helps to focus more in sublevel; second, local autocorrelation helps to explore spatial patterns when the global autocorrelation is significant; third, local autocorrelation helps to identify any inconsistent pattern. Therefore, Local Moran's I is defined as:

$$I_i = \frac{n(x_i - \bar{x})\sum_{j=1}^{n} W_{ij}(x_i - \bar{x})}{\sum_{j=1}^{n}(x_i - \bar{x})^2}, \qquad \forall i \neq j \tag{4}$$

3.3 Spatial Autocorrelation Model

Based in the general form of spatial autoregressive model with spatial effects proposed by Anselin (1988), we can express the model as:

$$\begin{aligned} &y = \delta wy + X\beta + WX\gamma + \epsilon, \\ &\epsilon = \lambda W\epsilon + \mu, \\ &\mu \sim N(0, \sigma^2 I_n). \end{aligned} \tag{5}$$

Where y is the endogenous variable; δ is the coefficient to the spatial lag of the endogenous variable wy; W is the weighted matrix of spatial neighbors; X is the set of exogenous variables; β is the effect of the exogenous variables represented by $X\omega$ is the coefficient to the spatial lag of the exogenous variables WX; λ is the coefficient to the spatial effect into the error terms $W\epsilon$; finally, μ are the uncorrelated disturbances. By following the general model is the Spatial Autocorrelation Model (LeSage, 2008), our model is defined as:

$$\begin{aligned} IPM = W(IPM) + \beta_1 Migration + \beta_2 Language + \beta_3 Female + \epsilon,& \\ \epsilon = \lambda W\epsilon + \mu,& \\ \mu \sim N(0, \sigma^2 I_n),& \end{aligned} \tag{6}$$

where IPM is the "Multidimensional Poverty Index". Migration, is defined as the percentage of households in each district where there is at least one member that migrated within the last five years from another region (department for Peru). Language, is defined as the percentage of households in each district where there is at least one member that speaks another original language than Spanish. Female, is defined as the percentage households where the household's head is female. All the parameters are estimated using maxim likelihood methods in order to obtain robust estimated coefficients. Additionally, there are several diagnostic tests that can be applied in the spatial model contexts. Lagrange Multiplier tests contrast the presence of spatial effects. (Anselin, 2001, 2010)

Therefore, we have the Lagrange Multiplier Test for Spatial Error (LM-error):

$$\text{Hypothesis:} \qquad H_0 : \delta = 0 \qquad \text{vs} \qquad H_1 : \delta \neq 0$$

$$LM\_error = \frac{(\frac{e'We}{e'e/n})^2}{tr(W' + W'W)}$$

And the Lagrange Multiplier Test for Spatial Lag (LM-lag):

$$\text{Hypothesis:} \qquad H_0 : \lambda = 0 \qquad \text{vs} \qquad H_1 : \lambda \neq 0$$

$$LM\_error = \frac{(\frac{e'Wy}{e'e/n})^2}{D + tr(W' + W'W)}$$

$$D = \frac{(WX\beta)'(I - X(X'X)X')^{-1}(WX\beta)}{e'e/n}$$

Finally, the Akaike Criterion (AIC) remains as useful indicator in order to select the best model among all the estimated models.

4 Empirical results

Table 1 shows the descriptive statistics for the main data information we are using in our estimations. These statistics give us a clearer idea how varied are the 1874 districts among Peru. As we observe in the table, there are districts as small as 2 squared km and as big as 24 049 squared km, but with an average size of 690 squared km per district. Furthermore, respect to each dimension, we are able to identify some clear characteristics. Around 5.3% of the households did not achieve the average of nine years of schooling; on the contrary, 22.7% of households have at least one member without the skill of read or write properly. Regard to the childhood dimension, we observe that most of the households do not fall into the poverty category, less than 1% of households have at least one member that does not attend any school institution, has no health insurance (public or private), and has a job to contribute the family income with less than 14 years old; nevertheless, around 30% of households have at least one member suffering of schooling lag. The descriptive statistics shows that Peru is composed with districts where 18.7% of households have at least one member out of the healthcare system and 63% of those households have members working in the informal sector. In other words, in a 5 members family, 1 on 5 of them does not have any health insurance and 3 on 5 of them have a job in the informal sector. This description contrast with the 4.7% of households with members without a job.

Respect to the living conditions of those households, we can identify that 2.2% of the households have no proper access to water services; in contrast, 38% have no sewage access, 45% have no proper electricity access. Also, 45%, 22%, and 16% of the households have poor building conditions in their floor, wall, and roof, respectively. Finally, 15% of households are considered overcrowded since they have 3 or more members per room. These previous household conditions bring us a total 20.1% of households falling into poverty category under the concept of "Multidimensional Poverty Index".

Regard to the covariates we are using in our model, 6.4% of households have at least one member that migrated from another region, 41.6% of them have members that do not speak Spanish, and 31.1% of households have a female as household head.

Table 1: Descriptive Statistics

Statistic	**N**	**Mean**	**St. Dev.**	**Min**	**Pctl(25)**	**Pctl(75)**	**Max**
size	1,874	689.61	1922.27	1.99	92.91	500.09	24049.95
Dimensions							
educ1	1,874	0.0530	0.0370	0.0000	0.0300	0.0660	0.3210
educ2	1,874	0.2270	0.1310	0.0060	0.1220	0.3210	0.6890
childhood1	1,874	0.3010	0.0880	0.0430	0.2450	0.3560	0.5870
childhood2	1,874	0.0100	0.0080	0.0000	0.0050	0.0120	0.0910
childhood3	1,874	0.0010	0.0020	0.0000	0.0000	0.0020	0.0130
childhood4	1,874	0.0030	0.0090	0.0000	0.0000	0.0030	0.3150
health1	1,874	0.1870	0.1060	0.0100	0.1000	0.2600	0.6620
Job1	1,874	0.0470	0.0330	0.0000	0.0260	0.0620	0.4990
Job2	1,874	0.6300	0.1320	0.0950	0.5530	0.7210	0.9660
house1	1,874	0.0220	0.0420	0.0000	0.0030	0.0240	0.5480
house2	1,874	0.3810	0.2050	0.0000	0.2210	0.5290	0.9330
house3	1,874	0.4530	0.1920	0.0060	0.3280	0.5890	0.9550
house4	1,874	0.2180	0.2290	0.0000	0.0260	0.3680	0.9420
house5	1,874	0.1550	0.0950	0.0080	0.0900	0.1900	0.7020
house6	1,874	0.4520	0.2220	0.0020	0.2850	0.6210	0.9520
house7	1,874	0.1500	0.1210	0.0000	0.0710	0.1890	0.8650
Endogenous Variable							
IPM	1,874	0.2010	0.0380	0.0810	0.1760	0.2240	0.3600
Exogenous Variables							
migration	1,874	0.0640	0.0530	0.0000	0.0250	0.0920	0.4520
language	1,874	0.4160	0.3980	0.0000	0.0160	0.8860	0.9990
female	1,874	0.3110	0.0760	0.0920	0.2580	0.3680	0.5210

In Table 2 we observe the spatial autocorrelation for for our set of variables that compose the MPI. As it is shown in most of the cases, except *childhood3*, *childhood4*, *job2*, and *house1*, the Moran's I, which measures the spatial autocorrelation of a variable, are higher than 0.4 and in all the cases it is statistically significant at 1%. These results give a clue to the spatial relationships that multidimensional poverty can have. Most of the variables that characterize poverty on its different faces show positive relational behaviors across space. In other words, poverty expressed in low levels of education, poor levels of health, work, living conditions, and housing conditions show patterns of spatial behavior and positive spatial self-correlation. That is, areas with high levels of poverty influence other areas, making the latter more likely to be poor.

Table 2: Summary Statistics of Spatial Autocorrelation

Statistic	**Moran's I**	**p-value**	**Statistic**	**Moran's I**	**p-value**
educ1	0.6184	0.0000	health1	0.6246	0.0000
educ2	0.6888	0.0000	house1	0.2142	0.0000
childhood1	0.6709	0.0000	house2	0.5146	0.0000
childhood2	0.4348	0.0000	house3	0.5948	0.0000
childhood3	0.2426	0.0000	house4	0.7492	0.0000
childhood4	0.0614	0.0000	house5	0.6140	0.0000
job1	0.4849	0.0000	house6	0.6974	0.0000
job2	0.2725	0.0000	house7	0.4132	0.0000
Endogenous Variable					
MPI	0.5641	0.0000			

Figure 1 shows the distribution of the MPI throughout the Peruvian districts. The figure on the left shows the MPI values for each district and the histogram of the MPI across districts. In this figure, it can be seen that the majority of districts with high levels of multidimensional poverty are located in the central highlands and jungle areas of Peru. In contrast, the Peruvian coast has low levels of multidimensional poverty. The figure on the right shows the spatial distribution of Moran's local indicator that measures the spatial autocorrelation of the MPI for each district. The figure shows that the districts with high levels of spatial autocorrelation are located in the central highlands and jungle of Peru. These results are in accordance with the figure on the left.

In Table 3 we see the results of the estimations where the endogenous variable is the multidimensional poverty indicator, MPI, and the control variables were constructed to isolate possible household characteristics that may affect poverty levels within the household. The first control variable, *migrant*, captures the effect of migration between regions, taking the value of 1 when at least one member of the household has migrated from some other region and 0 in another case. The second control variable, *language*, measures the effect of the second language as a possible cause of poverty in the household. This indicator takes the value of 1 when at least one member of the household has another language other than Spanish as their mother tongue and 0 in another case. Finally, the third control variable, *female*, captures the effect of mothers who are heads of household, this indicator takes the value of 1 when the head of household is female and 0 in other cases. In this table, the columns represent the estimation methods used. The first column is the

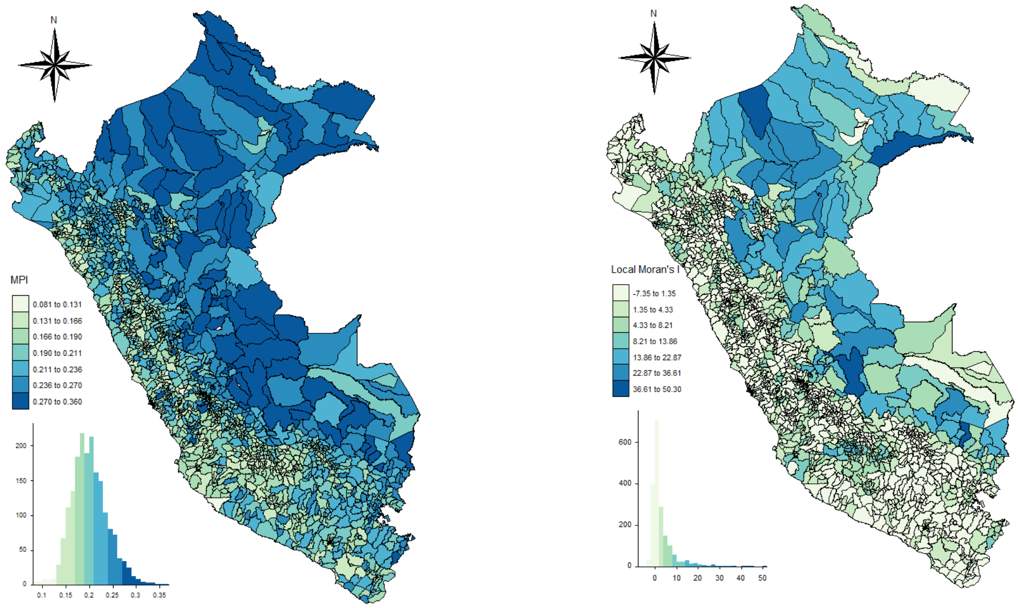

Figure 1: MPI and Local Moran's I per district

results of an OLS estimation, that is, when there are no spatial effects on multidimensional poverty. The second column shows the results of the spatial autoregression model, where externalities find their source in the endogenous variable; in other words, the MPI variable of a given district has spatial effects on the surrounding districts. The third column shows the results of the spatial error model, where the source of spatial autocorrelation is the errors. That is, the spatial autocorrelation is caused by variables not included in the model or by qualitative sources that could not be adequately captured in the model. Finally, column 4 shows the results of the autoregressive spatial model estimation with spatial errors. In this model, the source of spatial autocorrelation is not only the endogenous variable, but also errors can be possible causes of autocorrelation.

The first thing that can be observed from the results of these estimates is that the set of explanatory variables included in the model are statistically significant in all the estimated models. However, the signs appear to be somewhat counter-intuitive. In the case of the variable that captures migration, the results indicate that households with at least one migrant household member from another region are less likely to be in a multidimensional poverty household. This result may be due to the fact that much of the migratory flow between regions in Peru has been from the highlands and jungle regions to the coast and not vice versa. In this sense, it is very likely that migration has been from very poor households to coastal cities in search of better economic and living conditions. Therefore, families with migrant members are those who have possibly left poor households and have been able to escape in some degree from the poverty in which they

Table 3: Estimation Results

	Dependent variable: MPI			
	OLS	*Spatial Autoregressive*	*Spatial Error*	*Spatial Error Autoregressive*
migrant	-0.078***	-0.062***	-0.060***	-0.060***
	(0.015)	(0.015)	(0.015)	(0.015)
language	0.008***	0.008***	0.022***	0.023***
	(0.002)	(0.002)	(0.003)	(0.003)
female	-0.221***	-0.209***	-0.188***	-0.188***
	(0.011)	(0.011)	(0.011)	(0.011)
Constant	0.272***	0.249***	0.256***	0.256***
	(0.003)	(0.004)	(0.004)	(0.004)
δ		0.017***		-0.003***
		(0.002)		(0.002)
λ			0.127***	0.128***
			(0.003)	(0.003)
Observations	1,874	1,874	1,874	1,874
R2	0.196			
Adjusted R2	0.194			
Log Likelihood		3,701.63	4,111.52	4,112.32
sigma2		0.001	0.001	0.001
Akaike Inf. Crit.		-7,391	-8,211	-8,211
F-Statistic (df = 3; 1870)	151.596***			
Wald Test (df = 1)		85.818***	1,585.506***	
LR Test (df = 1)		87.192***	906.964***	908.570***
LM Test (df = 1)		1128.8***		
Note:	*p <0.1; **p <0.05***; p<0.01			

were in their regions of origin. Under this assumption, the resultant sign of estimation makes sense. For the variable that captures the effect of the second language, it is clear that households with members with a mother tongue other than Spanish are more likely to be in a poor household. This result is a consequence of the different social problems and cultural discrimination, which is associated with the mother tongue, that Peru has throughout its territory, the cultural and economic supremacy of those who speak Spanish over those whose mother tongue is Quechua, Aymara, Ashaninka, Aguaruna, among others. Finally, the variable that captures the effect of women as heads of household has a negative and statistically significant sign in all cases. These results are based on the assumption that families led by women are more likely to plan household spending and that family investment decisions are directed towards developing the capacities of household members, such as children. These assumptions are based on the results of other research done by researchers for different realities such as Duflo (2012) who holds that women tend to have better spending decisions within the household than their male counterparts.

Regarding the estimates of spatial effects, it is observed that both the spatial effects coming from the endogenous variable, as well as the errors, are statistically significant in the three estimated models. In the model where only spatial effects on the endogenous variable are incorporated, a positive and statistically significant coefficient is observed, suggesting that multidimensional poverty in a given district increases multidimensional poverty in a neighbouring district. In short, multidimensional poverty shows positive spatial effects. Similarly, when the model only includes spatial effects on errors, the spatial effect is positive and statistically significant, as shown in column 3. Finally, when spatial effects are incorporated in both the endogenous variable and the errors, the coefficients are statistically significant at 1% in both cases; however, the coefficient of the endogenous variable changes to values close to zero. This result could have two possible explanations. First, the spatial effects coming from the endogenous variable are unstable and when other explanatory variables or other sources of spatial effects are incorporated, the coefficient changes in a large proportion. Second, the interaction between the spatial effects of the endogenous variable and the errors produce changes in the coefficient of the spatial effects of the endogenous variable, making it even change its sign, although it is worth mentioning that the value is very close to zero. In short, this interaction makes the spatial effects coming from the endogenous variable lose strength until they are minuscule. What is important to note is that the spatial effects of the errors have remained stable and statistically significant, suggesting that there are factors not incorporated into the model that affect the spatial interaction of the multidimensional poverty indicator. It is observed that an increase of 1 in the MPI of a data district increases the multidimensional poverty of a neighbouring district by 0.128.

Furthermore, the Wald and LR test in all cases reject the null hypothesis of no spatial interactions; therefore, there is evidence of spatial effects. Also, the LM test of the column 2 rejects the null hypothesis

Table 4: Spatial Effects

	Spatial Autoregressive			**Spatial Error and Autoregressive**		
	Direct	**Indirect**	**Total**	**Direct**	**Indirect**	**Total**
migrant	-0.0626	-0.0062	-0.0688	-0.0593	0.0009	-0.0584
language	0.0083	0.0008	0.0092	0.0226	-0.0003	0.0222
female	-0.2092	-0.0208	-0.2300	-0.1875	0.0029	-0.1846

of no spatial interactions in the error terms; therefore, there is evidence that we must include the coefficient λ in our estimations. To support the LM test, the AIC criterion suggest that we must use the models where the spatial effect on the error terms is included.

Table 4 shows extra information where we include the direct and indirect effects originated in our set of covariates. The direct effects represent the effect of the exogenous variable over the endogenous variable without consider any spatial effects. Under no presence of spatial interactions, the direct effect corresponds to the OLS estimator. On the other hand, the indirect effects represent the effect of the exogenous variable over the endogenous variable in the neighbor districts and through this effects go back to the endogenous variable in the district of analysis. Direct and indirect effects can only be calculated when the spatial effects are in the endogenous variable or in the explanatory variables but not in the errors. Therefore, the table does not incorporate the externalities generated in the spatial error model.

Direct effects are the effects caused directly by the explanatory variables towards the endogenous variable, similar to the coefficients of an OLS model. However, in the case of indirect effects, these capture the externalities caused by the exogenous variables on the endogenous variable of a neighbouring district. The table shows that in some cases the effects change sign, more specifically, in the whole set of explanatory variables the externalities captured by the indirect effects change sign depending on the estimated model. This is due to the observed instability of the spatial effect coefficient of the endogenous variable incorporated in the model, as shown in Table 3.

5 Conclusion

This research studies the presence of multidimensional poverty in Peruvian households at the district level using data from the 2017 census. The main hypothesis of the study is that multidimensional poverty also has a spatial dimension that has been little studied in the literature.

Among the main results, we found evidence of spatial correlation of the variables that compose the distant dimensions of the multidimensional poverty indicator, MPI. A positive and statistically significant Moran indicator is observed in all cases. Likewise, it is observed that in most cases the value of spatial autocorrelation exceeds the value of 0.5. In other words, there are indications of the spatial dimension of poverty.

In the second part of the study, when spatial regressions are estimated for poverty using a set of control variables, two important results are observed. First, that the set of explanatory variables are all statistically significant across the estimated models. These results are explained by the social, cultural and economic changes that Peru has been undergoing in the last decades, among these changes we have the strong migration from the countryside to the city, which not only caused poor families to seek progress in the big cities, but in many cases people left their cultural background in search of better opportunities. Second, there is evidence of spatial effects on the endogenous variable and on errors. Moreover, there is evidence that the spatial effects originating in errors have a stable value among the models, while the spatial effects originating in the endogenous variable are unstable.

References

Akinyemi, F. and F. Bigirimana (2012): "A Spatial Analysis of Poverty in Kigali, Rwanda using indicators of household living standard," *Rwanda Journal*, 26, 3–22.

Anselin, L. (1988): *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer Academic Publishers.

——— (2001): "Spatial effects in econometric practice in environmental and resource economics," *American Journal of Agricultural Economics*, 83, 705–710.

——— (2010): "Thirty years of spatial econometrics," *Papers in Regional Science*, 89, 3–25.

Bigman, D. and H. Fofack (2000): "Geographical targeting for poverty alleviation: An introduction to the special issue," *World Bank Economic Review*, 14, 129–145.

Brunn, B. Y. S. D. and J. O. Wheeler (1971): "Spatial Dimensions of Poverty in the United States," *Geografiska Annaler*, 53, 6–15.

Chen, X., Z. Pei, A. L. Chen, F. Wang, K. Shen, Q. Zhou, and L. Sun (2015): "Spatial Distribution Patterns and Influencing Factors of Poverty - A Case Study on Key Country From National Contiguous Special Poverty-stricken Areas in China," *Procedia Environmental Sciences*, 26, 82–90.

Crandall, M. I. S. C. and B. R. A. W. Weber (2004): "Local Social and Economic Conditions, Spatial Concentrations of Poverty, and Poverty Dynamics," *American Journal of Agricultural Economics*, 86, 1276–1281.

Duflo, E. (2012): "Women Empowerment and Economic Development," *Journal of Economic Literature*, 50, 1051–1079.

Grab, J. (2009): *Econometric analysis in poverty research : with case studies from developing countries*, vol. 26.

Holt, J. B. (2008): "The topography of poverty in the United States: A spatial analysis using county-level data from the community health status indicators project," *Preventing Chronic Disease*, 5.

Khawas, V. (2006): *Human Development Report 2005-2006*, vol. 36.

LeSage, J. P. (2008): "An introduction to spatial econometrics," *Revue d'Economie Industrielle*, 123, 19–44.

Mussagy, I. H. and I. I. Dambudzo (2015): "A Spatial Analysis of Poverty and Roads Network in Sofala Province," *Journal of International Academic Research for Multidisciplinary*, 3, 333–341.

Odekon, M. (2015): "Multidimensional Poverty Index," *The SAGE Encyclopedia of World Poverty*, 1–8.

Rupasingha, A. and S. J. Goetz (2007): "Social and political forces as determinants of poverty: A spatial analysis," *Journal of Socio-Economics*, 36, 650–671.

Sen, A. (1976): "Poverty: An Ordinal Approach to Measurement," *Econometrica*, 44, 219.

Tanaka, T. and J. J. Lee (2011): "Spatial analysis of poverty : Combining geospatial data and survey data to study regional inequality in Ghana," *World Bank*, 32.

Wallender, W. W., L. N. Rodrigues, L. H. Bassoi, and J. A. Young (2011): "Spatial patterns of rural poverty: An exploratory analysis in the São Francisco River Basin, Brazil," *Nova Economia*, 21, 45–66.