CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A.C.

SPATIAL DEPENDENCE OF SLUM SEVERITY IN CDMX

TESINA

QUE PARA OBTENER EL GRADO DE

MAESTRA EN ECONOMÍA AMBIENTAL

PRESENTA

NADIA LIZBETH JIMÉNEZ RODRÍGUEZ

DIRECTOR DE LA TESINA: DR. ALFONSO MIRANDA

AGUACALIENTES, AGS. 2021

Para mi amada familia:
Mi madre,
mis hermanos Fefa, Evan y Dany.

Acknowledgments

Thanks to my advisor, Dr. Alfonso Miranda for his guidance and advice for this thesis. I admire him for his work, but more for the person that he is. I appreciate that he agreed to work with me. He gave me a lot of light on what I was doing.

Thanks to my reader Dr. Jaime Sainz for his valuable comments, his kindness, understanding and commitment.

A very special thanks to Dr. Elio Atenógenes Villaseñor for his support and his kindness.

Thanks to Dr. Pedro Hancevic and Dr. Héctor Núñez for their support, understanding and encouragement in the subjects assigned to work on the thesis.

Thanks to the CIDE for providing me with the conditions and support to study the master's degree, and in particular to conclude this thesis.

Finally, thanks to my mother who has given me her unconditional support throughout my life. And thanks to my dear brothers Evan and Dany for their support and love during this project.

Abstract

In this work a spatial econometric model is presented to analyze the vulnerability index in CDMX. Nutrition, population density, budget, economic units, gross value of production, number of hospitals and clinics, property deeds, percentage of income destined to pay rent, population with mental problems or conditions, maximum temperatures, illegal dumps, air pollution and green areas are used as covariates. I conclude that there is a phenomenon of segregation in the population of CDMX between neighboorhood areas.

Key words: spatial econometrics, slum severity, segregation

Contents

List of Figures

List of Tables

Chapter 1

Introduction

It is known that approximately 55% of the world's population lives in cities, according to UN-HABITAT data reported in (UN-HABITAT, 2020b). In that same report, projections are made for 2030 with 6 out of 10 people being urban inhabitants, and for 2050 with two-thirds of the population being urban inhabitants. Also in (UN-HABITAT, 2020b) it is pointed out that urbanization is more than a demographic or spatial phenomenon. This phenomenon could help the world population to overcome some of its main challenges, such as poverty, inequality, environmental problems, among other goals that also belong to the 2030 Agenda.

Sustainable cities and human settlements are essential for sustainable development, as UN-HABITAT points out in (UN-HABITAT, 2020b). In (UN-HABITAT, 2020b), it is emphasized that cities can act as a network that relates all the sustainable development objectives, because they relate the social and economic results of the population with the environment, energy and the economy, among others. However, in (UN-HABITAT, 2020b) states that achieving these goals are hampered by the inadequate functioning of the processes for sustainable urbanization, it may be due to several reasons, for example, problems with obtaining and quality of data from cities. And this causes, as noted in (UN-HABITAT, 2020b), that this network is not well connected because data is not well connected with knowledge, so that policies are proposed with this, and from there to adequate financing of integrated urban projects.

The 2016 World Cities Report reports that 75% of cities have higher levels of spatial inequality compared to levels two decades ago, as indicated in (UN-HABITAT, 2020a). Also in (UN-HABITAT, 2020a), it is explained that this comparison tells us that basic urban services, affordable housing, quality public spaces and livelihood opportunities in these cities are not evenly distributed. The division of urban spaces is made visible generally by phenomena such as segregation and the concentration of poverty ((UN-HABITAT, 2020a)). These zones of socioeconomic and spatial exclusion (where higher levels of poverty, crime and environmental dangers, unemployment, among others, can be observed, compared to the average) represent a considerable part of the world population ((UN-HABITAT, 2020a)). It is known that of the world population, one billion people live in informal settlements and, in addition, 600 million people live in inadequate housing UN ((UN-HABITAT, 2020a)).

According to the United Nations (UN), the estimated population for CDMX is 21 million 581 thousand inhabitants and, therefore, CDMX is the fifth most populated city in the world, according to data in (Venegas, 2018) . In the same study ((Venegas, 2018)) it is estimated that by 2035 it will have 24 million 490 thousand inhabitants, and by 2050, it will have 144.9 million inhabitants, which is equivalent to 88% of the total population.

As is stated in (Venegas, 2018), in order to face these social challenges, the Organization for Economic Cooperation and Development (OECD) recommends a housing financing model, to guarantee this right, giving more possibilities its inhabitants, and promote a model of smart and sustainable urbanization, and institutions with the capacity to put it into practice.

In (UN-HABITAT, 2020c) it is argued that the increase in population, especially in emerging countries, makes it necessary to pay attention to issues such as housing, transportation, energy, education, employment, among others, to meet the needs of their habitants; in such a way that they are inclusive regardless of residence.

In this paper I will study a phenomenon that frequently occurs in urban areas: slums. In particular, I will analyze the relationship they have with other variables such as nutritional risk, mental conditions and problems, economic units and illegal dumps, among others, within an

AGEB in CDMX. And also, the spatial relationship that slums have at the AGEB level. This analysis could help improve the urbanization process in many ways.

I will present the literature review first. Then the data I work with. Then I will make a brief description of the method I use, in this case a spatial econometric model. Then, applying this method to the data, I will show the results obtained. Finally I will discuss the principal results of this analysis and I will make my conclusions of this work.

Chapter 2

Literature review

According to the United Nations (UN), we can say that in 1900 one in 10 people lived in cities. Currently, almost 3 billion people, that is, almost half of humanity, reside in urban centers, and there are already 23 cities, 18 of them belonging to the developing world, with more than 10 million inhabitants. One of the consequences of this rapid increase in the urban population is that millions of poor people around the world live in crowded slums and illegal settlements, in living conditions below the minimum levels necessary to guarantee the health of families and communities. The issue has been worrisome around the world. For this reason, the United Nations Center for Human Settlements (Habitat) was created within the United Nations system in 1978, with the mission of coordinating the activities related to human settlements that the UN has generated in this regard (Velasco, 2019).

> The definition of the term "slum" includes the traditional meaning – that is, housing areas that were once respectable or even desirable, but which have since deteriorated as the original dwellers have moved to new and better areas of the cities. The condition of the old houses has then declined, and the units have been progressively subdivided and rented out to lower-income groups. Typical examples are the inner-city slums of many towns and cities in both the developed and the developing regions. Slums have, however, also come to include the vast infor-

mal settlements that are quickly becoming the most visible expression of urban poverty in developing regions cities, including squatter settlements and illegal subdivisions. The quality of dwellings in such settlements varies from the simplest shack to permanent structures, while access to water, electricity, sanitation and other basic services and infrastructure is usually limited. Such settlements are referred to by a wide range of names and include a variety of tenure arrangements. (UN-HABITAT, 2010, p. 10).

In the other hand, there are linkages of poverty and environment at the household level, for example in Philippine slums. Rapid urbanization and the inadequate infrastructure and basic services in large towns and cities have led to the proliferation of slums and informal settlements in the country. While poverty incidence of population in key metropolitan centers is on average 17% compared to the national average of 32%, slum population has been exponentially rising at an average rate of 3.4%. In Metro Manila, which is the prime city, an estimated 37% of population or over 4.0 million Filipinos live in slums in 2010 and slum population growth rate is at 8% annually. These slum dwellers and informal settlers confront on a daily basis another dimension of poverty which is environmental poverty. The underserviced and bad living conditions in slums impact on health, livelihood and the social fiber. The effects of urban environmental problems and threats of climate change are also most pronounced in slums due to their hazardous location, poor air pollution and solid waste management, weak disaster risk management and limited coping strategies of households. Bad living environment thus deepens poverty, increases the vulnerability of both the poor and non-poor living in slums and excludes the slum poor from growth. (Ballesteros, 2010, p. 1).

Using three simple frameworks it was created a meso level portrait of poverty and living conditions in the slums of Dakar, Senegal and Nairobi, Kenya. While slum residents in both cities share the challenge of monetary poverty, their experience diverges significantly relative to employment levels, education, and living conditions. Nairobi's relatively well-educated and employed residents suffer from poorer living conditions -as measured by access to infrastructure and urban services, housing quality and crime- than residents of Dakar, who report much lower levels of educational attainment and paid employment. The research findings challenge conventional development theory -particularly notions that education and jobs will translate into lower poverty and improved living conditions. (Gulyani, Bassett, y Talukdar, 2014, p. 98).

These findings suggest that reduction in income poverty and improvements in human development do not automatically translate into improved infrastructure access or living conditions. Since not all slum residents are poor, living conditions also vary within slums depending on poverty status. Compared to their non-poor neighbors, the poorest residents of Nairobi or Dakar are less likely to use water (although connection rates are similar) or have access to basic infrastructure (such as electricity or a mobile phone). Neighborhood location is also a powerful explanatory variable for electricity and water connections, even after controlling for household characteristics and poverty. Finally, tenants are less likely than homeowners to have water and electricity connections. (Gulyani, Talukdar, y Jack, 2010, p. 1).

Also, it was found that access to solid-waste removal services is reasonably high in Dakar but almost non-existent in Nairobi. About 73 percent of Dakar's slum households have some form of an organized garbage collection system, compared with only 12 percent in Nairobi. Of this 12 percent, private collection is a central part of that system, accounting for 11 of that 12 percent. For Dakar, 70 out of the total 76 percent of slum households with access to organized collection systems depend on city/municipal collection systems. For households without access to an organized garbage disposal system, the predominant method is "dumping in the neighborhood" in both cities. (Gulyani et al., 2014, p. 104).

The waste generated by cities is of enormous consequence, and solid waste is a pressing issue for urbanization, as it relates to public health, land use and climate mitigation. Solid waste generation is set to outpace population growth by more than double by 2050. Worldwide, approximately 2.0 billion tonnes of solid waste is generated annually; of this amount, around a third is not managed sustainably. Solid waste emits 1.6 billion tonnes of carbon dioxide, accounting for 5 per cent of emissions. Low-income developing countries struggle with the management and processing of waste; municipalities spend large shares of their budgets on waste management, approximately five times the share that highincome municipalities expend on average. Additionally, over 90 per cent of waste in lowincome countries is openly dumped or burned, rather than being collected and processed formally. Collection in low-income countries has significantly increased from 22 to 39 per cent. (UN-HABITAT, 2020c, p. 101).

Findings on environmental inequity has implications for the higher long-term health burden of air pollution suffered by the more marginalized communities across the Mexican territory. Environmental regulators can target more marginalized communities to achieve the largest reductions in air pollution levels. Related public policies such as health and those based on welfare can also target their limited resources at the most marginalized communities. (Chakraborti1 y Voorheis, 2021, p. 31).

Chapter 3

Data

3.1 Dependent variable

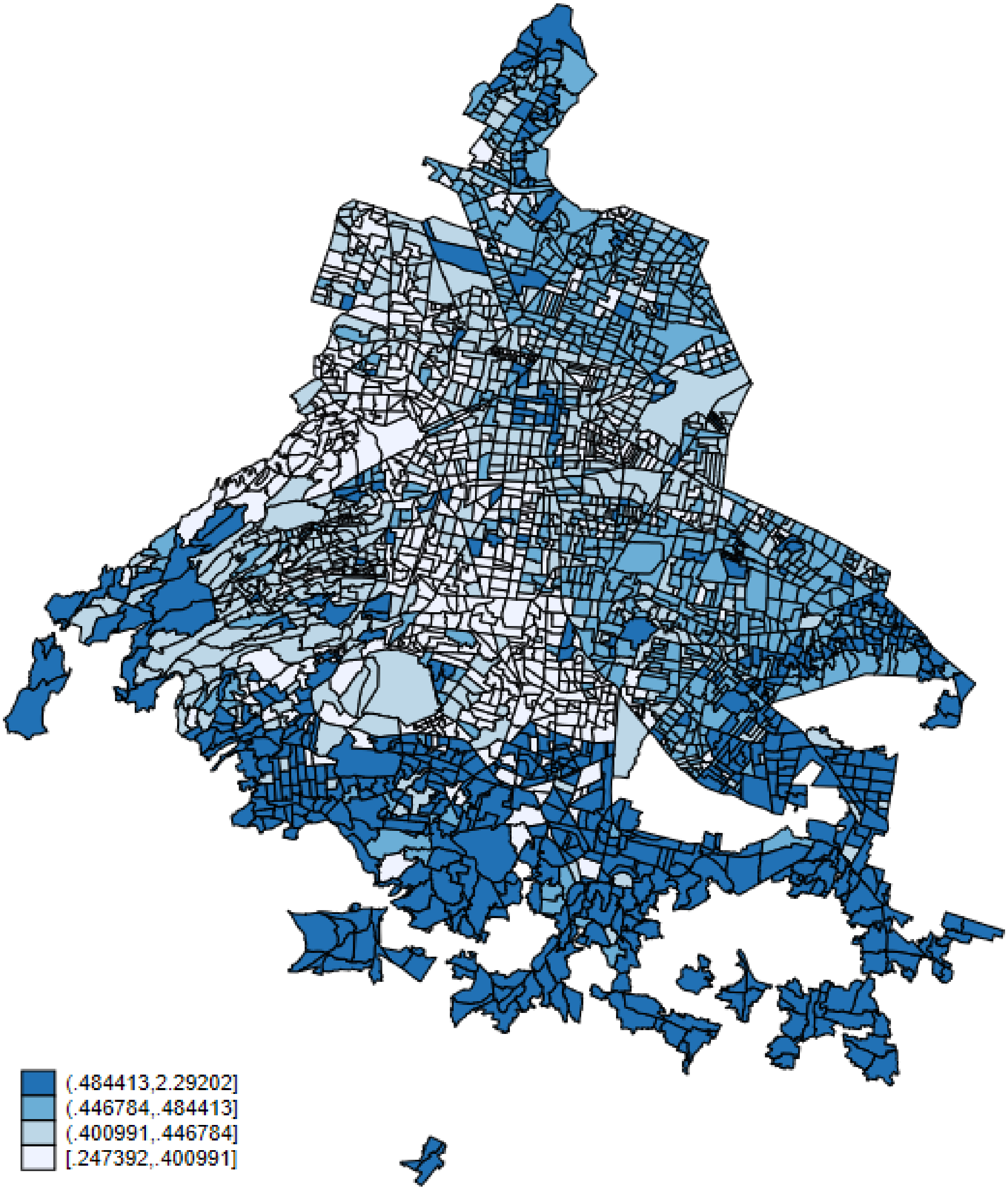
A vulnerable settlement is defined as an area with a roof shared by a certain number of people and used as housing in an urban area. It has one or more of the following characteristics: (1) the home does not protect them from natural conditions such as extreme climates; (2) it does not have enough space to be habitable, this is measured with more than 3 people in a single room; (3) it does not have access to potable water service in sufficient quantities and at an acceptable price; (4) there is not accessibility to a private or public bathroom shared by a reasonable number of people; and (5) it does not have deeds to ensure its belonging ((UN-HABITAT, 2006)).

As is stated in (UN-HABITAT, 2006), the level of slum severity measured by this index depends on how many of these characteristics of a vulnerable settlement there are and at what level. For this reason, marginal neighborhoods are not homogeneous, as they present different adverse conditions.

The National Institute of Statistics and Geography (INEGI) calculated a multidimensional slum severity index (SSI) in Mexico, through an exploratory factor analysis as in (Roy, Bernal, y Lees, 2019). INEGI uses the following proxy variables to develop it: (1) the average number of occupants per room per block; (2) the proportion of households per block with dirt floors; (3)

the proportion of households per block without drinking water; (4) the proportion of households per block without drainage; (5) the proportion of homes per block without electricity; and (6) the proportion of households per block without a bathroom. This SSI is the variable that I analyze in this work,but in the case of CDMX, it is the dependent variable. For this analysis I take the SSI by basic geostatistical area (AGEB). Figure 3.1 shows the distribution of SSI in CDMX. The data can be downloaded from https://www.inegi.org.mx/programas/ccpv/2020/.

Figure 3.1: Slum Severity Index (SSI).



Source: Own elaboration.

3.2 Independent variables

I have $2,431$ observations by AGEB. The independent variables are the following:

1. nutri risk:= Nutritional risk 2000. From the National Commission for the Knowledge and Use of Biodiversity (CONABIO). The data can be downloaded from http://geoportal.conabio.gob.mx/. The Figure 3.2 shows the distribution of "health" in Mexico City by basic geostatistical area (AGEB).

2. density:= Population density 2020. From the National Institute of Statistics and Geography (INEGI). The data can be downloaded from https://www.inegi.org.mx/programas/ccpv/2020/. The Figure 3.3 shows the distribution of "density" in Mexico City by basic geostatistical area (AGEB).

3. budget:= Budget that the city government assigns 2020. From the Open Data Portal. Government of Mexico City. The data can be downloaded from https://www.datos.gob.mx/. The Figure 3.4 shows the distribution of "budget" in Mexico City by basic geostatistical area (AGEB).

4. econ units:= Economic units 2019. From Timely Indicators of the City of Mexico of the Secretariat of Economic Development (SEDECO), Government of the City of Mexico. The data can be downloaded from https://www.sedeco.cdmx.gob.mx/. The Figure 3.5 shows the distribution of "employment" in Mexico City by basic geostatistical area (AGEB).

5. gvp:= Gross value of production 2019. From Timely Indicators of the City of Mexico of the Secretariat of Economic Development (SEDECO), Government of the City of Mexico. The data can be downloaded from https://www.sedeco.cdmx.gob.mx/. The Figure 3.6 shows the distribution of "gvp" in Mexico City by basic geostatistical area (AGEB).

6. hhc:= Number of hospitals and health centers 2020. Open Data Portal. Government of Mexico City. The data can be downloaded from https://www.datos.gob.mx/. The

Figure3.7 shows the distribution of "hhc" in Mexico City by basic geostatistical area (AGEB).

7. housing deed:= Number of households having housing deed 2008. ENIGH 2018. From the National Institute of Statistics and Geography (INEGI). The data can be downloaded from https://www.inegi.org.mx/programas/enigh/nc/2018/. The Figure 3.8 shows the distribution of "housing deed" in Mexico City by basic geostatistical area (AGEB).

8. rent:= Percentage of the household income assigned to pay the rent of the household 2018. ENIGH 2018. From the National Institute of Statistics and Geography (INEGI). The data can be downloaded from https://www.inegi.org.mx/programas/enigh/nc/2018/. The Figure 3.9 shows the distribution of "rent" in Mexico City by basic geostatistical area (AGEB).

9. pmpc:= Population with a mental problem or condition 2020. From the National Institute of Statistics and Geography (INEGI). The data can be downloaded from https://www.inegi.org.mx/progra The Figure 3.10 shows the distribution of "pmpc" in Mexico City by basic geostatistical area (AGEB).

10. max temp:= Maximum temperatures 2020. Open Data Portal. Government of Mexico City. The data can be downloaded from https://www.datos.gob.mx/. The Figure 3.11 shows the distribution of "max temp" in Mexico City by basic geostatistical area (AGEB).

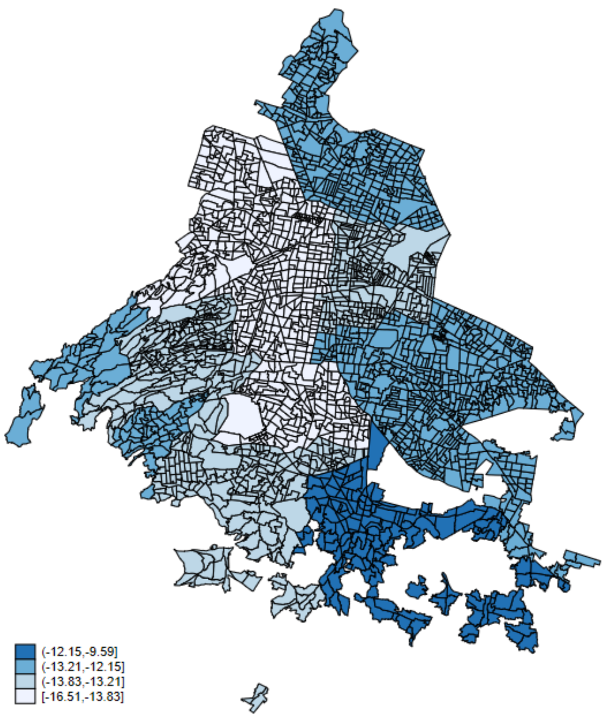
11. illegal dumps:= Number of illegal dumps 2017. Open Data Portal. Government of Mexico City. The data can be downloaded from https://www.datos.gob.mx/. The Figure 3.12 shows the distribution of "illegal dumps" in Mexico City by basic geostatistical area (AGEB).

12. air pollution:= Concentrations of pollutants in the air 2020. Open Data Portal. Government of Mexico City. The data can be downloaded from https://www.datos.gob.mx/. The

Figure 3.13 shows the distribution of "air quality" in Mexico City by basic geostatistical area (AGEB).

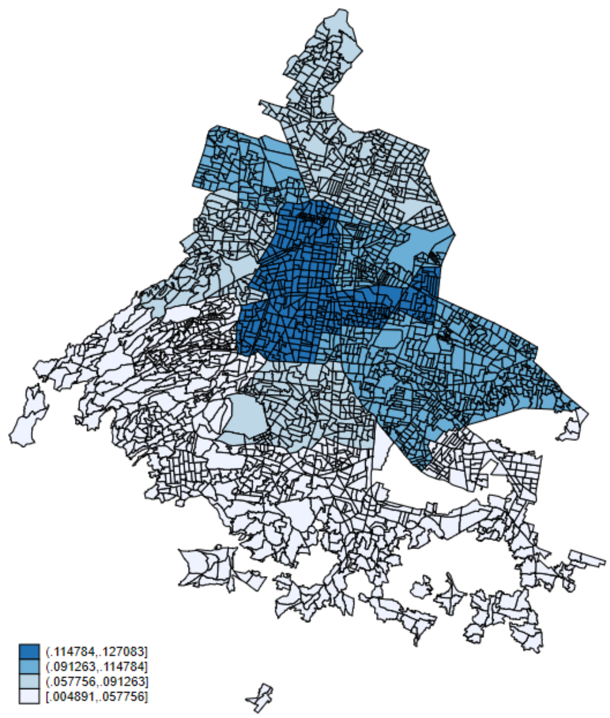
13. green areas:= Proportion of the green areas of Mexico City 2017. Open Data Portal. Government of Mexico City. The data can be downloaded from https://www.datos.gob.mx/. The Figure 3.14 shows the distribution of "green areas" in Mexico City by basic geostatistical area (AGEB).

Figure 3.2: health



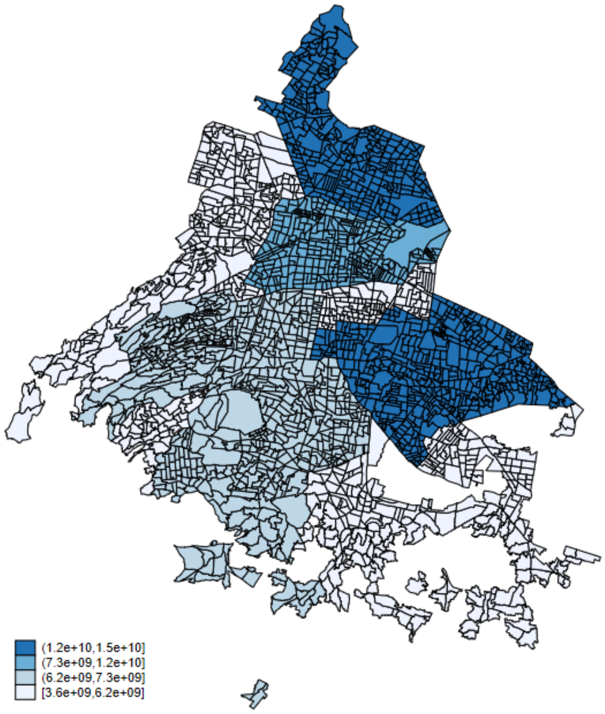
Source: Own elaboration.

Figure 3.3: density



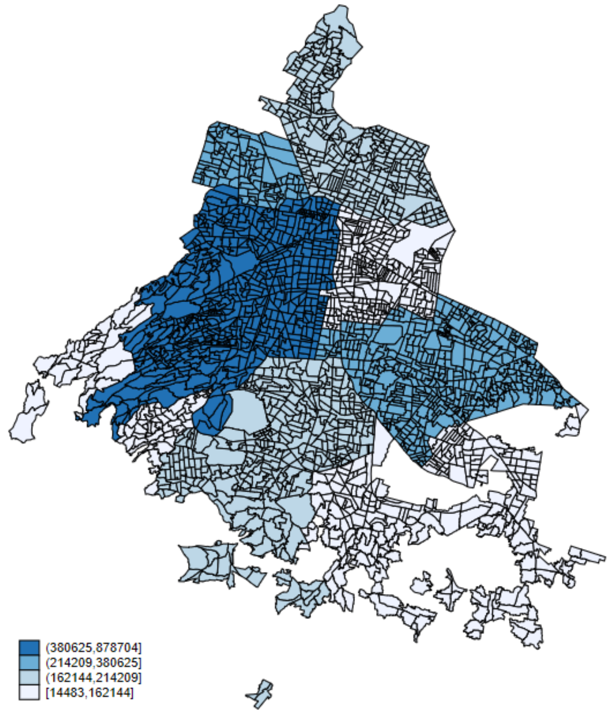
Source: Own elaboration.

Figure 3.4: budget



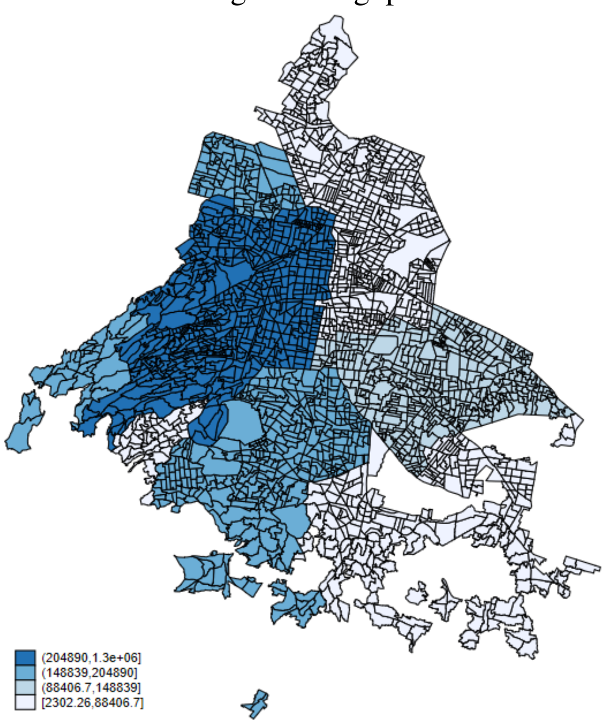
Source: Own elaboration.

Figure 3.5: employment



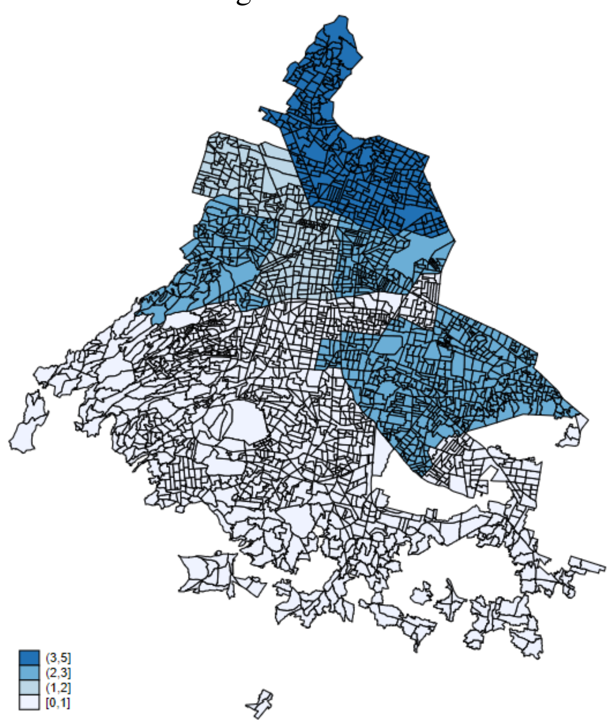
Source: Own elaboration.

Figure 3.6: gvp



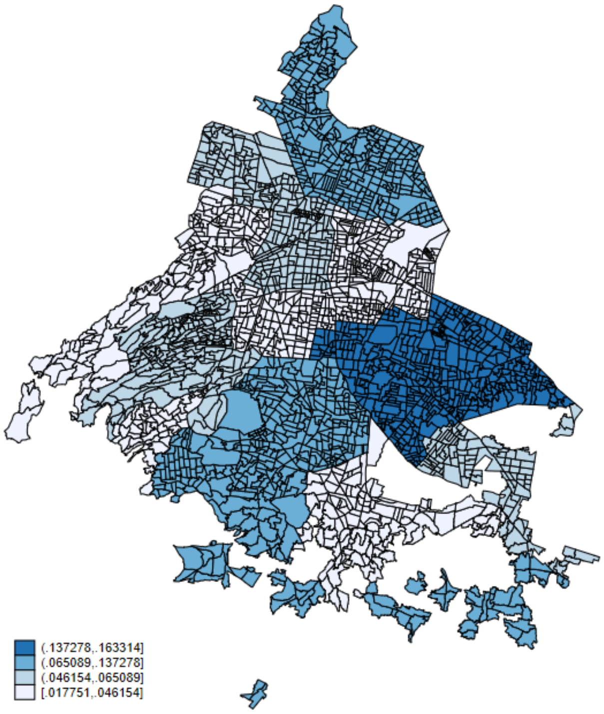
Source: Own elaboration.

Figure 3.7: hhc



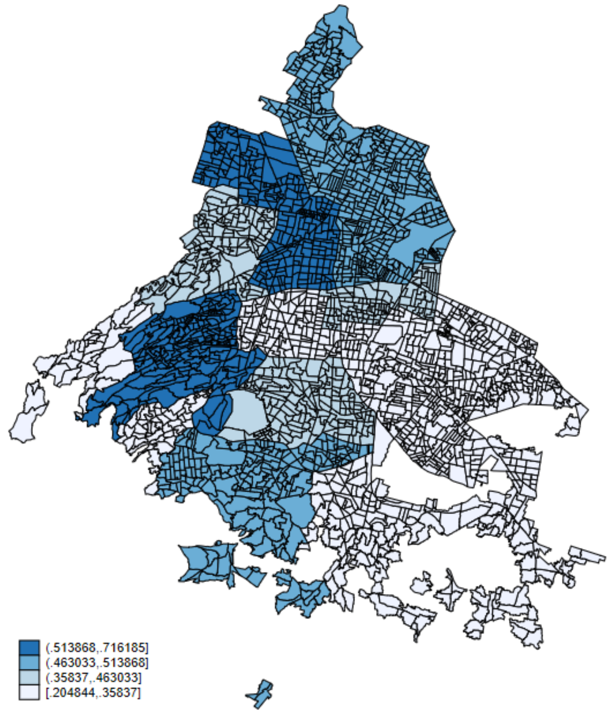
Source: Own elaboration.

Figure 3.8: housing deed



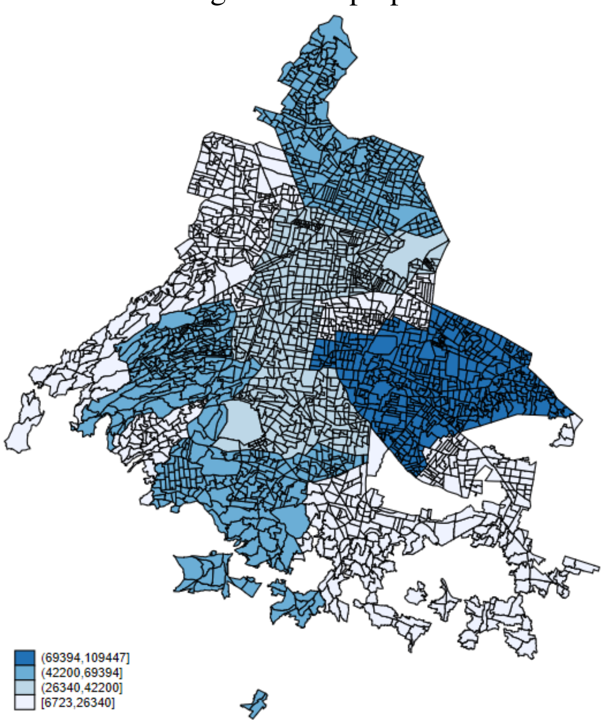
Source: Own elaboration.

Figure 3.9: rent



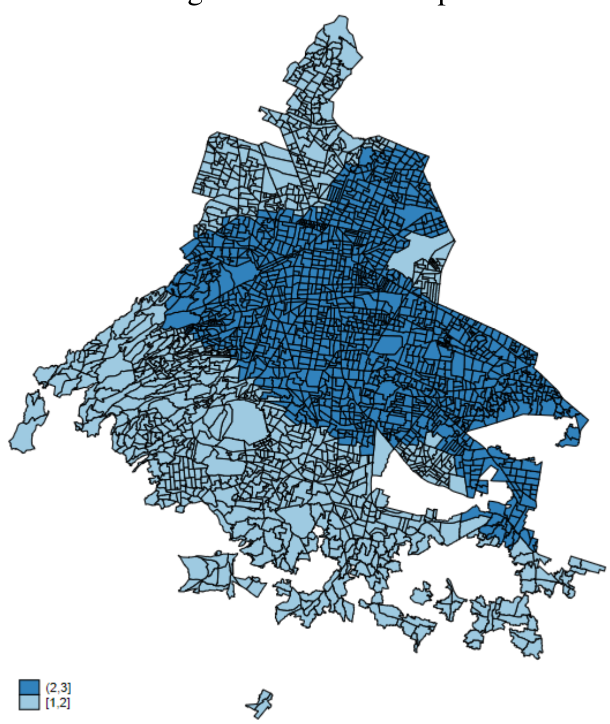
Source: Own elaboration.

Figure 3.10: pmpc



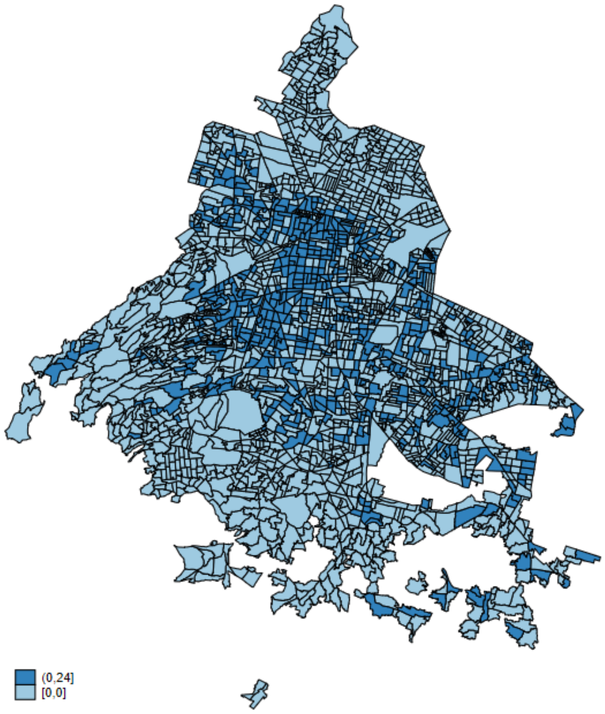
Source: Own elaboration.

Figure 3.11: max temp



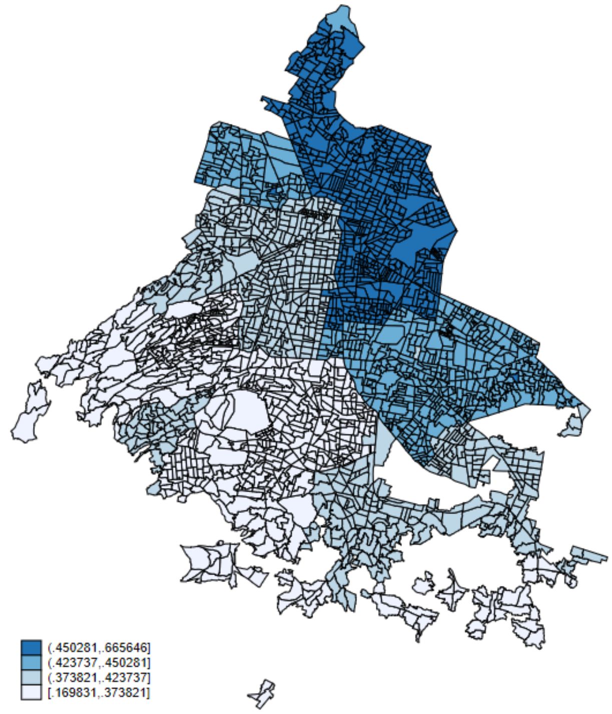
Source: Own elaboration.

Figure 3.12: illegal dumps



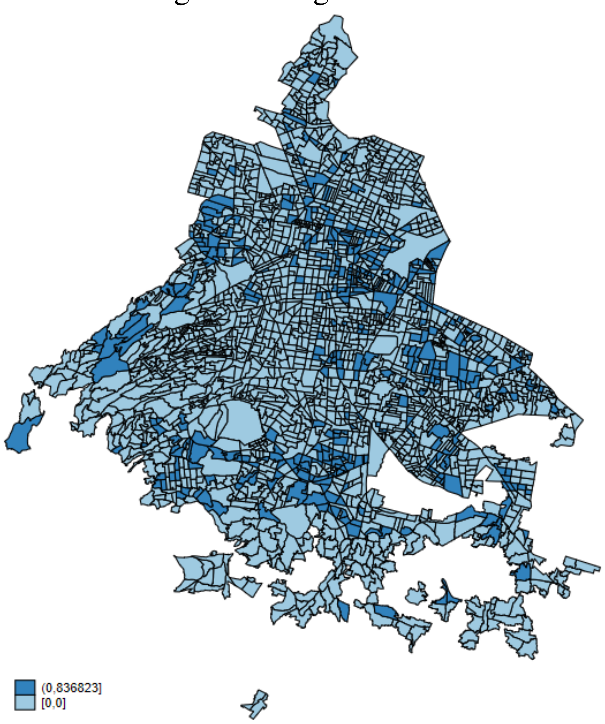
Source: Own elaboration.

Figure 3.13: air quality



Source: Own elaboration.

Figure 3.14: green areas



Source: Own elaboration.

Chapter 4

Methods

4.1 Research area

Spatial econometrics is a subfield of econometrics dealing with spatial interaction effects among geographical units. Whereas the time-series literature focuses on the dependence among observations over time and uses the symbol "$t - 1$" to denote variables lagged in time, the spatial econometrics literature is interested in the dependence among observations across space and uses the so-called spatial weights matrix W to describe the spatial arrangement of the geographical units in the sample. It should be stressed here that spatial econometrics is not a straightforward extension of time series econometrics to two dimensions. One obvious difference is that two geographical units can affect each other mutually, whereas two observations in time cannot (Elhorst, 2014, p. 1).

In this work we will take advantage of the data structure using a spatial econometric model. Therefore, there will be interesting interpretations of the estimated parameters for this model, in particular, interaction effects between geographic units.

4.2 SAR models

> SAR models are fit using datasets that contain observations on spatial units such as countries, districts, or even nongeographical units such as social network nodes. For simplicity, we refer to these spatial units as areas. Datasets contain at a minimum a continuous outcome variable, such as incidence of disease, output of farms, or crime rates, along with the other variables assumed to predict the chosen outcome. (StataCorp, 2017, p. 5).

The dataset could be used to fit a linear regression of the form

$$y_i = \beta_0 + x_{i,1}\beta_1 + x_{i,2}\beta_2 + \cdots + x_{i,k}\beta_k + \epsilon_i \tag{4.1}$$

> This linear regression is provided as a starting point; it is not a SAR model. To give this starting point a spatial feel, we will call the observations areas. The variables contain characteristics of the areas. (StataCorp, 2017, p. 6).

The notation we will use is

i area (observation), numbered 1 to N

y_i:= dependent (outcome) variable in area i

$x_{i,1}$:= 1st independent variable in area i

$\vdots$

$x_{i,j}$:= jth independent variable in area i

$\vdots$

$x_{i,k}$:= last independent variable in area i

ϵ_i:= error (residual) in area i

The linear regression model can be written in column-vector notation:

$$\mathbf{y} = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \cdots + \beta_k\mathbf{x}_k + \boldsymbol{\epsilon} \tag{4.2}$$

The boldfaced variables are each $N \times 1$ vectors.

> SAR models extend linear regression by allowing outcomes in one area to be affected by outcomes in nearby areas. Said in the spatial jargon, models can contain spatial lags of the outcome variable. These terms are borrowed from the time-series literature. (StataCorp, 2017, p. 6).

In time series, an autoregressive $AR(1)$ process is

$$y_t = \gamma_0 + \gamma_1 y_{t-1} + \epsilon_t \tag{4.3}$$

where y_{t-1} is called the lag of y. In vector notation, $L.$ is the lag operator, and the above equation could be written as

$$\mathbf{y} = \gamma_0 + \gamma_1 \mathbf{L}.\mathbf{y} + \boldsymbol{\epsilon} \tag{4.4}$$

> The time-series notation and jargon can be translated to the spatial domain. The lag operator becomes an $N \times N$ matrix $\mathbf{W}$. What was $\mathbf{L}.\mathbf{y}$ becomes $\mathbf{W}\mathbf{y}$, which means matrix $\mathbf{W}$ multiplied by vector $\mathbf{y}$. (StataCorp, 2017, p. 7).

The SAR model corresponding to the above time-series equation is

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{W}\mathbf{y} + \boldsymbol{\epsilon} \tag{4.5}$$

> $\mathbf{W}$ is called the spatial weighting matrix. The values in the matrix characterize the spatial relationships between areas. $\mathbf{W}$ is the spatial analog of $\mathbf{L}.\mathbf{y}$. Whereas $\mathbf{L}.\mathbf{y}$ measures the potential spillover from time $t-1$ to t, elements W_{i_1,i_2} specify how much potential spillover there is from area i_2 to i_1. W_{i_1,i_2} is zero if area i_2 can have no effect on i_1. The more potential spillover there is, the larger W_{i_1,i_2} is. The elements of $\mathbf{W}$ are specified before the model is fit. In the mathematics of SAR models: $\mathbf{W}\mathbf{y}$ is the spatial equivalent of $\mathbf{L}.\mathbf{y}$. Either way, it is the lag of the dependent variable. (StataCorp, 2017, p. 7).

Recall that the linear regression model we started with was

$$\mathbf{y} = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \cdots + \beta_k\mathbf{x}_k + \boldsymbol{\epsilon} \tag{4.6}$$

We could add $\mathbf{Wy}$ to the model:

$$\mathbf{y} = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \cdots + \beta_k\mathbf{x}_k + \lambda\mathbf{Wy} + \boldsymbol{\epsilon} \tag{4.7}$$

> The result would be that $\lambda\mathbf{W}$ would measure the amount that outcomes are affected by nearby outcomes. You can think of $\mathbf{W}$ as specifying the potential spillover as long as you realize that the actual spillover is the effect that y_i of area i has on nearby y's from the term $\lambda\mathbf{Wy}$. The weighting matrix $\mathbf{W}$ is effectively a constraint placed on the individual spillovers formulated as part of the model specification. (StataCorp, 2017, p. 7).

4.3 Direct and indirect effects

Also, I will present the exposition of Direct and indirect effects in (StataCorp, 2017). The solution to the SAR model is

$$\mathbf{y} = (\mathbf{I} - \lambda\mathbf{W})^{-1}(\mathbf{X}\boldsymbol{\beta} + \epsilon) \tag{4.8}$$

implies that the mean of $\mathbf{y}$ given the independent variables and the spatial weighting matrix is

$$E(\mathbf{y}|\mathbf{X}, \mathbf{W}) = (\mathbf{I} - \lambda\mathbf{W})^{-1}(\mathbf{X}\boldsymbol{\beta}) \tag{4.9}$$

> This is known as the reduced-form mean because the solution in equation 4.8 is known as the reduced form of the model. The predicted reduced-form mean substitutes estimates of λ and β into equation 4.9. (StataCorp, 2017, p. 97).

To define the direct mean and the indirect mean, let

$$\mathbf{S} = (\mathbf{I} - \lambda \mathbf{W})^{-1} \tag{4.10}$$

and let $\mathbf{S}_d$ be a matrix with diagonal elements of $\mathbf{S}$ on its diagonal and with off-diagonal elements set to 0. (StataCorp, 2017, p. 97).

The direct means are

$$\mathbf{S}_d \mathbf{X} \boldsymbol{\beta} \tag{4.11}$$

which capture the contributions of each unit's independent variables on its own reduced-form mean. Substituting estimates of λ and β produces the predictions. (StataCorp, 2017, p. 97).

The indirect means capture the contributions of the other units' independent variables on a unit's reduced-form prediction. (StataCorp, 2017, p. 97)

And they are

$$\{(\mathbf{I} - \lambda \mathbf{W})^{-1} - \mathbf{S}_d\} \mathbf{X} \boldsymbol{\beta} \tag{4.12}$$

The partial derivatives of $E(\mathbf{Y})$ with respect to the jth explanatory variable have the property that, if a particular explanatory variable in a particular unit changes, not only will the dependent variable in that unit itself change but also the dependent variables in other units. The first is called a direct effect and the second an indirect effect. (Elhorst, 2014, p. 21)).

The total impact of an independent variable x is the average of the marginal effects it has on the reduced-form mean. (StataCorp, 2017, p. 98),

$$\frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial E(\mathbf{y}_i | \mathbf{X}, \mathbf{W})}{\partial x_j} \tag{4.13}$$

where $E(\mathbf{y}_i | \mathbf{X}, \mathbf{W})$ is the ith element of the vector $E(\mathbf{y} | \mathbf{X}, \mathbf{W})$, whose formula is given in equation 4.9, and x_j is the jth unit's value for x. (StataCorp, 2017, p. 98).

The direct impact of an independent variable x is the average of the direct, or own, marginal effects. (StataCorp, 2017, p. 98):

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial E(\mathbf{y}_i|\mathbf{X},\mathbf{W})}{\partial x_i} \tag{4.14}$$

The indirect impact of an independent variable x is the average of the indirect, or spillover, marginal effects. (StataCorp, 2017, p. 98):

$$\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\frac{\partial E(\mathbf{y}_i|\mathbf{X},\mathbf{W})}{\partial x_j} \tag{4.15}$$

Chapter 5

Results

5.1 SAR model

The theory tells us that it makes sense to think of a spatial model for the SSI. It will be analyzed if the data indicates that this is the case. I will use the reference manual for SAR models (StataCorp, 2017) to obtain the results.

The descriptive statistics of the SSI are the following,

Table 5.1: SSI description.

Variable	Obs	Mean	Std.Dev.	Min	Max
SSI	2,431	.471174	.1282501	.2473925	2.292022

Source: Own elaboration.

The SSI varies between 0.25 and 2.29 in AGEBs. In Figure **??**, it is shown that the SSI is segregated. This leads us to hypothesize that there are negative spillover effects among AGEBs. The phenomenon of segregation in the population is observed as the separation of a part of its population that migrates to other places.

Spatial data lends itself to spatial and non-spatial analysis. For this reason, I will first do a linear regression of the SSI on the independent variables. Then, with the STATA commands, diagnose if the residuals are spatially correlated in this model. The results can be seen in table 5.2 and in table 5.3.

Table 5.2: Linear regression.

Source	SS	df	MS			
				Number of obs	=	2,431
				F(13, 2417)	=	91.24
Model	13.157312	13	1.01210092	Prob > F	=	0.0000
Residual	26.811518	2,417	.011092891	R-squared	=	0.3292
				Adj R-squared	=	0.3256
Total	39.96883	2,430	.016448078	Root MSE	=	.10532

Source: Own elaboration.

Table 5.3: Linear regression coefficients.

SSI	Coef.	Std. Err.	t	$P > \|t\|$	[95% Conf.	Interval]
nutri risk	.0124241	.0030414	4.09	0.000	.0064601	.018388
density	-2.096916	.2091745	-10.02	0.000	-2.507096	-1.686736
budget	-8.442532	7.210123	-1.17	0.242	-22.58119	5.696129
econ units	1.204431	.5861562	2.05	0.040	.0550109	2.353852
gvp	-1.172911	.378052	-3.10	0.002	-1.914251	-.4315715
hhc	-.002565	.0085965	-0.30	0.765	-.0194224	.0142923
housing deed	-.7294054	.2425072	-3.01	0.003	-1.204949	-.2538618
rent	.1729636	.0278321	6.21	0.000	.1183863	.2275409
pmpc	2.827374	.81995	3.45	0.001	1.219496	4.435252
max temp	.0061545	.0057847	1.06	0.287	-.005189	.017498
illegal dumps	.0070778	.0018249	3.88	0.000	.0034993	.0106563
air pollution	.0184041	.0855825	0.22	0.830	-.1494184	.1862267
green areas	-1.21416	.7917922	-1.53	0.125	-2.766822	.3385014
cons	.6931723	.0693497	10.00	0.000	.5571813	.8291633

Source: Own elaboration.

Now, a statistical test is performed to find if there is spatial dependency, Sp provides Moran's test for this. First, I have to build the spatial weighting matrix, which states what and how much the slums are "close". This matrix is created by the command texttt spmatrix. In this analysis

I use a contiguity matrix. This contiguity matrix defines "close" as "shares an edge". The result of Moran's test is in table **??**.

Table 5.4: Moran test.

Moran test for spatial dependence
Ho: error is i.i.d.
Errorlags: W
chi2(1) = 617.03
Prob $>$ chi2 = 0.0000

Source: Own elaboration.

The result of this test tells us that we can reject that the residuals are independent and identically distributed (i.i.d.). In particular, the test took into account the alternative hypothesis that the residuals of this model are correlated with the residuals of nearby AGEBs as defined by $\mathbf{W}$. So, I propose to model SSI by a SAR model:

$$\begin{aligned}\mathbf{SSI} =& \beta_0 + \beta_1 \mathbf{nutri\,risk} + \beta_2 \mathbf{density} + \beta_3 \mathbf{budget} + \beta_4 \mathbf{econ\,units} \\ &+ \beta_5 \mathbf{gvp} + \beta_6 \mathbf{hhc} + \beta_7 \mathbf{housing\,deed} + \beta_8 \mathbf{rent} + \beta_9 \mathbf{pmpc} \\ &+ \beta_{10} \mathbf{max\,temp} + \beta_{11} \mathbf{illegal\,dumps} + \beta_{12} \mathbf{air\,pollution} \\ &+ \beta_{13} \mathbf{green\,areas} + \lambda \mathbf{WSSI} + \boldsymbol{\epsilon}\end{aligned} \tag{5.1}$$

The model I fit will include the term $\lambda\mathbf{WSSI}$, because I assume the SSI spills over from nearby AGEBs. Behind this assumption is the idea that one might think that the population of slums in much worse conditions (SSI high) look for housing with better conditions to live when they have more income, and they are likely to look for it in nearby places for reasons such as work or school. Also because their income is not enough high, so that what they could find would be a house in nearby places, since the costs of the house will be similar, not so high, but with better conditions to live. The result of the spatial regression without independent variables, only with the lag in the dependent variable is,

Table 5.5: Spillovers.

SSI	Coef.	Std. Err.	z	$P > \|z\|$	[95% Conf.	Interval]
SSI						
cons	.5196539	.0086649	59.97	0.000	.502671	.5366368
W						
SSI	-.1298157	.022054	-5.89	0.000	-.1730407	-.0865906
Wald test of spatial terms:			chi2(1) = 34.65		Prob > chi2 = 0.0000	

Source: Own elaboration.

This spatial regression confirms the segregation in the population because the coefficient of **W** is negative. So, the population of a worse slum looks for a better slum nearby.

Now, I fit the SAR model proposed in equation 5.1. The results are in table 5.6.

Table 5.6: Spatial regression coefficients.

SSI	Coef.	Std. Err.	z	$P > \|z\|$	[95% Conf.	Interval]
SSI						
nutri risk	.0119669	.0030564	3.92	0.000	.0059764	.0179574
density	-2.171724	.2117771	-10.25	0.000	-2.5868	-1.756649
budget	-8.487893	7.234218	-1.17	0.241	-22.6667	5.690914
econ units	1.193507	.5881279	2.03	0.042	.0407977	2.346217
gvp	-1.179738	.3793231	-3.11	0.002	-1.923198	-.4362786
hhc	-.0032835	.0086295	-0.38	0.704	-.0201971	.01363
housing deed	-.7649288	.2436887	-3.14	0.002	-1.24255	-.2873077
rent	.1835107	.0282094	6.51	0.000	.1282214	.2388001
pmpc	2.95279	.8240582	3.58	0.000	1.337665	4.567914
max temp	.0072382	.0058185	1.24	0.214	-.0041659	.0186422
illegal dumps	.0073361	.0018336	4.00	0.000	.0037423	.0109299
air pollution	.0293068	.0859675	0.34	0.733	-.1391864	.1977999
green areas	-1.117835	.7952734	-1.41	0.160	-2.676542	.440872
cons	.6969324	.0695958	10.01	0.000	.5605271	.8333377
W						
SSI	-.044735	.0169444	-2.64	0.008	-.0779454	-.0115246
Wald test of spatial terms:			chi2(1) = 6.97		Prob > chi2 = 0.0083	

Source: Own elaboration.

Results for the spatial regression are similar to those reported by linear regression. However, when spillover effects are significant, frequently the other parameters change. In this case, we find that λ (which multiplies $\mathbf{W}SSI$) is significant, that is, the estimated coefficient on the spatial lag of SSI is -0.045, indicating negative correlation between the SSI in one AGEB and the SSI in a neighboring AGEB.

If, for example, $nutri\,risk$ increases, that increases SSI by β_1, and that increment in SSI spills over to produce a further increment in SSI of $\lambda\mathbf{W}$, and that increment spills over to produce yet another increment in SSI, and so on. the command in STATA `estat impact` reports the average effects from this recursive process.

5.2 Direct and indirect effects

In Table 5.7 it is show the effects. In the table it is reporting derivatives, but in this analysis I will interpret the results as if they were for a change in one unit. The interpretation of the coefficients from the table is the following. For example, the table reports average changes for a 1 unit increase in the dependent variable $nutri\,risk$. The direct effect is the effect of the change within the AGEB, ignoring spillover effects. The own-AGEB direct effect is to increase the SSI by 0.012 percentage points. The indirect effect is the spillover effect. A 1 unit increase in the dependent variable $nutri\,risk$ increase SSI, and that increment spills over to further reduce SSI. The result is a 0.0004 reduction in SSI. The sum of both effects is the total effect, which is $0.012 + -0.0004 = 0.11$. In the next section I discuss these results.

Table 5.7: Direct and indirect effects.

Average impacts Number of obs $= 2,431$

	Delta-Method					
	dy/dx	Std. Err.	z	$P > \|z\|$	[95% Conf.	Interval]
direct						
nutri risk	.0119695	.003057	3.92	0.000	.0059779	.017961
density	-2.172194	.2118706	-10.25	0.000	-2.587453	-1.756935
budget	-8.48973	7.235787	-1.17	0.241	-22.67161	5.692152
econ units	1.193766	.5882539	2.03	0.042	.0408091	2.346722
gvp	-1.179994	.3794066	-3.11	0.002	-1.923617	-.4363704
hhc	-.0032843	.0086314	-0.38	0.704	-.0202015	.013633
housing deed	-.7650944	.2437484	-3.14	0.002	-1.242832	-.2873564
rent	.1835504	.0282197	6.50	0.000	.1282408	.2388601
pmpc	2.953429	.8242644	3.58	0.000	1.3379	4.568957
max temp	.0072397	.0058199	1.24	0.214	-.004167	.0186464
illegal dumps	.0073377	.0018341	4.00	0.000	.003743	.0109324
air pollution	.0293131	.0859863	0.34	0.733	-.1392169	.1978431
green areas	-1.118077	.7954371	-1.41	0.160	-2.677105	.440951
indirect						
nutri risk	-.0004172	.0001813	-2.30	0.021	-.0007726	-.0000618
density	.075712	.0296479	2.55	0.011	.0176031	.1338209
budget	.2959102	.2747612	1.08	0.281	-.2426119	.8344323
econ units	-.0416088	.0254649	-1.63	0.102	-.0915191	.0083016
gvp	.0411288	.0201178	2.04	0.041	.0016986	.0805589
hhc	.0001145	.0003051	0.38	0.707	-.0004834	.0007124
housing deed	.0266674	.0132977	2.01	0.045	.0006043	.0527305
rent	-.0063977	.0026676	-2.40	0.016	-.011626	-.0011693
pmpc	-.102942	.048715	-2.11	0.035	-.1984217	-.0074623
max temp	-.0002523	.0002288	-1.10	0.270	-.0007008	.0001961
illegal dumps	-.0002558	.0001162	-2.20	0.028	-.0004836	-.000028
air pollution	-.0010217	.0030381	-0.34	0.737	-.0069764	.0049329
green areas	.0389707	.0305987	1.27	0.203	-.0210018	.0989431

Source: Own elaboration.

Chapter 6

Discussion

Under the assumption that the population wants to improve the conditions of their housing, which reduces the SSI, through the income they receive, the following interpretations of the direct and indirect effects in table 5.7 will be made. First, notice that the independent variables: $budget$, hhc, $max\, temp$, $air\, pollution$ and $green\, areas$ are not significant. So, there is not direct or indirect effects in the SSI. Then, it only remains to interpret the direct and indirect effects of the other independent variables.

The direct effect in the increment by one unit of the variable $nutri\, risk$ is an increase of $.012$ in the SSI within the AGEB, ignoring spillover effects. This may be due to the fact that their income is reduced by being used in health. So the lower-income population will look for cheaper homes that tend to have a higher SSI. This increase in the SSI have an indirect effect in the nearby AGEBS of a $.0004$ decrease in their SSI. This could occur because the population with a higher income prefers not to live in a place where the variable $nutri\, risk$ increases, and has the possibility of moving to a nearby place where they can improve the conditions of that new home, thus reducing the SSI of nearby places.

The direct effect in the increment by one unit of the variable $density$ is a decrease of 2.17 in the SSI within the AGEB, ignoring spillover effects. This may be due to a process of urbanization, since as the population grows, housing projects are created for the population. This decrease in the SSI have an indirect effect in the nearby AGEBS of a $.076$ increase in their SSI. This may be due to the fact that the population that does not have sufficient income to acquire the homes of the new projects migrate to nearby places with housing conditions similar to those they had before these urbanization projects, and this increases the SSI of these nearby places.

The direct effect in the increment by one unit of the variable $econ\, units$ is an increase of 1.19 in the SSI within the AGEB, ignoring spillover effects. This could be because increasing the economic units of an AGEB increases the supply of jobs and with this the population that wants to reside near these economic units that are not planned for housing. So unplanned settlements will be created near the economic units which for the same reasons will have a high SSI. This increase have an indirect effect of $-.042$, but it is not significant.

The direct effect in the increment by one unit of the variable gvp is a decrease of 1.18 in the SSI within the AGEB, ignoring spillover effects. This may be due to the fact that the population with more income can benefit from this increase in production by improving the conditions of their housing, and thus decrease their SSI. This decrease in the SSI have an indirect effect in the nearby AGEBS of a $.041$ increase in their SSI.This is because possibly the increase in production makes lower-income populations want to live in that AGEB, but as there is a high demand for housing, they will only be able to approach that AGEB and live in nearby AGEBS, thus increasing the SSI from the nearby AGEBS.

The direct effect in the increment by one unit of the variable $housing\, deed$ is a decrease of $.765$ in the SSI within the AGEB, ignoring spillover effects. This could happen because the population, knowing that the home is their property, is incentivized to improve its conditions, thus reducing the SSI. This decrease in the SSI have an indirect effect in the nearby AGEBS of a $.027$ increase in their SSI. This may be because having the house deeds, this increases the value of the property. Then, the population with less income probably cannot acquire it and decides to migrate to nearby places where it is cheaper and for this reason where there tends to be a higher SSI.

The direct effect in the increment by one unit of the variable $rent$ is a increase of $.184$

in the SSI within the AGEB, ignoring spillover effects. This is probably due to the fact that by increasing the percentage of their income that they allocate for rent, the populations will have fewer financial resources to improve their housing and even worsen the initial housing conditions. This increase in the SSI have an indirect effect in the nearby AGEBS of a $.006$ decrease in their SSI. This is probably because the population with less income decides to migrate to nearby places and not allocate that higher proportion of income to rent. And with that money that they no longer require for rent, improve the conditions of their new home, thus reducing the SSI in nearby places.

The direct effect in the increment by one unit of the variable $pmpc$ is a increase of 2.95 in the SSI within the AGEB, ignoring spillover effects. This may be because mental conditions and problems can affect the perception of well-being and quality of life. So it could happen that the population does not seek to improve the conditions of their housing and even worsens or chooses a housing in worse conditions, and this will increase the SSI for that AGEB. This increase in the SSI have an indirect effect in the nearby AGEBS of a $.103$ decrease in their SSI. This occurs because probably the population with more income migrate to nearby places where the incidence of mental conditions or problems is lower. Thus, as this population has a higher income, it can improve the conditions of its housing, thus reducing the SSI of nearby places.

The direct effect in the increment by one unit of the variable $illegal\ dumps$ is a increase of $.007$ in the SSI within the AGEB, ignoring spillover effects. This may be because it is more difficult to improve housing conditions if there are illegal dumps nearby. They are probably marginalized places where improving housing is not prioritized, and for this reason the SSI increases in that AGEB. This increase in the SSI have an indirect effect in the nearby AGEBS of a $.0002$ decrease in their SSI. This is probably due to the fact that garbage from nearby AGEBS is deposited where there is a greater presence of illegal dumps. And because of this, the homes of the nearby AGEBS are in better conditions, thus reducing their SSI.

The development of nations must be understood as a harmonious whole that involves strategies aimed at protecting the environment, economic growth and improving the living conditions of the entire population, particularly those most in need. (Velasco, 2019).

> Neglected historical sites in inner cities, dilapidated public housing in monofunctional residential zones, declining industrial areas and unplanned neighbourhoods in peri-urban areas are shared experiences in cities irrespective of their income classes. Such derelict and dysfunctional locations typically host a disproportionately high share of populations experiencing cumulative disadvantages due to their exclusion from prosperity and development opportunities generated by urbanisation. (UN-HABITAT, 2020a, p. 1).

> The UN states that we envisage cities and human settlements that fulfill their social function, including the social and ecological function of land, with a view to progressively achieving the full realization of the right to adequate housing as a component of the right to an adequate standard of living, without discrimination, universal access to safe and affordable drinking water and sanitation, as well as equal access for all to public goods and quality services in areas such as food security and nutrition, health, education, infrastructure, mobility and transportation, energy, air quality and livelihoods. (UN-HABITAT, 2020c, p. 61).

> Also, it is argued that financial tools involve direct exchange of funds between the public and private sectors for a regeneration project. These could include a variety of value capture methods such as impact fees, levies and special assessments. In this category, there are tools that are more sophisticated and require a high capacity within the government to execute and implement. More importantly these tools require that the city is creditworthy and can borrow in the financial markets. The second group of such tools do not require a linkage to capital markets and could be implemented by cities without such access. (UN-HABITAT, 2020c, p. 65).

> Finally, it is argued that people reside in slums because there are no other housing alternatives and the demand–supply gap for the low income sector continues

to grow. A large number of the population in slums are the low income workers that provide labor in the service sectors, industrial production and construction. They contribute substantially to productivity and growth in urban areas yet they are deprived of basic services in cities. The rising population in slums shows that inequality is rising and growth has not been inclusive. Improving slums would not only impact on poverty reduction but also bring about growth due to higher productivity of labor. Less slums will also attract tourist and investment in cities. Slum poverty cannot be addressed through traditional poverty programs such as cash transfer because bad housing significantly lowers health status of households especially children. It is noteworthy that among the housing components that tend to matter most in terms of health index and households' assessment of risk reduction are public good types- drainage, sewer facilities, asphalt roads, solid waste management, pollution enforcement etc- which the individual household cannot provide or enforce by itself. These "goods" require government investments and regulatory actions. It implies investments in basic infrastructure and flood mitigation measures and effective town planning and pollution controls. It also implies strong national government presence since public good investments and environmental concerns cut across administrative boundaries. (Ballesteros, 2010, p. 26-27).

Chapter 7

Conclusion

The results of this work show us the direct and indirect effects of the independent variables in the SSI. They are the following:

- The direct effect in the increment by one unit of the variable $nutri\,risk$ is an increase of .012 in the SSI within the AGEB, ignoring spillover effects. This increase in the SSI have an indirect effect in the nearby AGEBS of a .0004 decrease in their SSI.

- The direct effect in the increment by one unit of the variable $density$ is a decrease of 2.17 in the SSI within the AGEB, ignoring spillover effects. This decrease in the SSI have an indirect effect in the nearby AGEBS of a .076 increase in their SSI.

- The direct effect in the increment by one unit of the variable $econ\,units$ is an increase of 1.19 in the SSI within the AGEB, ignoring spillover effects. This increase have an indirect effect of $-.042$, but it is not significant.

- The direct effect in the increment by one unit of the variable gvp is a decrease of 1.18 in the SSI within the AGEB, ignoring spillover effects. This decrease in the SSI have an indirect effect in the nearby AGEBS of a .041 increase in their SSI.

- The direct effect in the increment by one unit of the variable $housing\,deed$ is a decrease of .765 in the SSI within the AGEB, ignoring spillover effects. This decrease in the SSI

have an indirect effect in the nearby AGEBS of a $.027$ increase in their SSI.

- The direct effect in the increment by one unit of the variable $rent$ is a increase of $.184$ in the SSI within the AGEB, ignoring spillover effects. This increase in the SSI have an indirect effect in the nearby AGEBS of a $.006$ decrease in their SSI.

- The direct effect in the increment by one unit of the variable $pmpc$ is a increase of 2.95 in the SSI within the AGEB, ignoring spillover effects. This increase in the SSI have an indirect effect in the nearby AGEBS of a $.103$ decrease in their SSI.

- The direct effect in the increment by one unit of the variable $illegal\, dumps$ is a increase of $.007$ in the SSI within the AGEB, ignoring spillover effects. This increase in the SSI have an indirect effect in the nearby AGEBS of a $.0002$ decrease in their SSI.

Due to these effects on the SSI, I can conclude that the living conditions of people in CDMX can be improved if public policies are applied in the independent variables of this analysis. Well, directly or indirectly, they could affect the SSI in the urbanization process, i. e., public policies can intervene in the independent variables of this model to affect the SSI of the AGEBS in CDMX.

References

Ballesteros, M. M. (2010). “Linking poverty and the environment: Evidence from slums in philippine cities.” *PIDS Discussion Paper Series*.

Chakraborti1, L., y Voorheis, J. (2021). “Are poorer mexicans exposed to worse air quality? long-term evidence from satellite imaging data.” *SSRN*.

Elhorst, J. P. (2014). *Spatial Econometrics. From Cross-Sectional Data to Spatial Panels*. Springer.

Gulyani, S., Bassett, E. M., y Talukdar, D. (2014). “A tale of two cities: A multi-dimensional portrait of poverty and living conditions in the slums of dakar and nairobi.” *Habitat International*(43), 98-107.

Gulyani, S., Talukdar, D., y Jack, D. (2010). “Poverty, living conditions, and infrastructure access.” *The World Bank. Africa Region. Sustainable Development Division*.

Roy, D., Bernal, D., y Lees, M. (2019). “An exploratory factor analysis model for slum severity index in mexico city.” *Urban Studies*.

StataCorp. (2017). *STATA Spatial autoregressive models reference manual. Release 15*. A Stata Press Publication.

UN-HABITAT. (2006). “State of the world’s cities 2006/7.”

UN-HABITAT. (2010). “The challenge of slums: Global report on human settlements 2003. revised and updated version.”

UN-HABITAT. (2020a). “Flagship programme 1: Inclusive, vibrant neighbourhoods and communities.”

UN-HABITAT. (2020b). "Flagship programme 5: Sdg cities."

UN-HABITAT. (2020c). *The New Urban Agenda Ilustrated*. Author.

Velasco, F. J. C. (2019). *Derecho urbanístico mexicano*. UNAM Instituto de Investigaciones Jurídicas.

Venegas, E. (2018). "México, entre los de mayor urbanización del mundo." *La razón de México*.