Revista Portuguesa de Estudos Regionais, nº 63 2023, 127-140

Artigo submetido a 31 de Julho de 2021; versão final aceite a 29 de Março de 2022
Paper submitted on July 31, 2021; final version accepted on March 29, 2022

CoViD-19 During the First Pandemic Wave in Mexico Valle and Mexico City: A Spatial Analysis Approach in Small Areas

CoViD-19 Durante a Primeira Onda Pandêmica no Valle de México e na Cidade do México: Uma Abordagem de Análise Espacial em Pequenas Áreas

Jaime Alberto Prudencio Vázquez
japv@azc.uam.mx
Departamento de Economía, Universidad Autónoma Metropolitana, Unidad Azcapotzalco

José Antonio Huitrón Mendoza
jose_eco71@comunidad.unam.mx
Facultad de Economía, Universidad Nacional Autónoma de México

Abstract

The first cases of the CoViD-19 disease in Mexico came from abroad in February 2020. Communitarian propagation accelerated the infection in the big metropolitan areas of Mexico, such as Valle de México Metropolitan Zone (VMMZ), were located the biggest people concentration in the country. In this study, we evaluate the spatial distribution of the positive cases and deaths in VMMZ at municipality level through a spatial econometric model that include socio demographic and economic variables, besides we explore the active cases in México City at neighborhood level. We found significant spatial effects, most notably in positive cases, that could help to explain the stage of the disease, in both municipality and neighborhood. The model shed light to observe how the CoViD-19 hits harder at the municipalities more densely populated and where the urbanization process was deeper, compared with those peripheral, nevertheless, worst living conditions also exhibit a positive relationship, in both positive cases and deaths.

Keywords: CoViD-19, spatial econometric analysis, Valle de México.

JEL classification: I14, R19, C59

Resumo

Os primeiros casos da doença CoViD-19 no México vieram do exterior em fevereiro de 2020. A propagação comunitária acelerou a infecção nas grandes áreas metropolitanas do México, como a Zona Metropolitana do Vale do México (VMMZ), onde se localizava a maior concentração de pessoas na região. país. Neste estudo, avaliamos a distribuição espacial dos casos positivos e óbitos em VMMZ em nível municipal por meio de um modelo econométrico espacial que inclui variáveis sociodemográficas e econômicas, além de explorarmos os casos ativos na Cidade do México em nível de bairro. Encontramos efeitos espaciais significativos, principalmente nos casos positivos, que poderiam ajudar a explicar o estágio da doença, em ambos os níveis município e bairro. O modelo lançou luz ao observar como o CoViD-19 atinge mais fortemente os municípios mais densamente povoados e onde o processo de urbanização foi mais profundo, em comparação com os periféricos, no entanto, as piores condições de vida também apresentam uma relação positiva, tanto nos casos positivos como nos óbitos.

Palavras-chave: CoViD-19, análise econométrica espacial, Vale do México

Código JEL: I14, R19, C59

1. INTRODUCTION

The vaccine program in Mexico started on December 23rd, 2020, nevertheless, up to June 22nd 2021, almost 2.5 million cases of CoViD-19 and 231.3 thousand deaths were reported in Mexico (Dong, Du and Gardner, 2020). As in the other countries but China, in Mexico the novel coronavirus, SARS-CoV-2, came abroad in February 2020. After some weeks, the communitarian propagation accelerated contagion in the big metropolitan areas of Mexico, such as México Valle Metropolitan Zone (VMMZ). The risk of exposure of the virus and hence the risk of death by CoViD-19 are associated with geographical, social, economic, and even climate and pollution factors (Coccia, 2020; Qiu et al., 2020; Sannigrahi et al., 2020).

The VMMZ is an extremely inequal metropolitan area where the gap between the highest and the lowest municipality per capita income is almost 6 times in 2015 (PNUD, 2019). There is evidence that these sociodemographic disparities have contributed to explain the differences at positive cases and death at a local level. Jaramillo (2021) shows how the poorest settlements in Mexico City have registered the highest number of cases compared to the neighborhoods where people of a higher socioeconomic level live. Hernández (2020) presents evidence of the high rate of infections among the population that shares the following characteristics: low schooling, inadequate nutrition, comorbidities, living with a certain degree of overcrowding and in localities where there is a deficit of access to basic services, all this combined with precarious economic conditions, forced them to have breaking the confinement and exposing themselves to contagions. These ideas about how the population with the greatest economic limitations are those that are most exposed to contagions and their consequences are also present in several studies (Rivero, 2020; Guillón, 2021; Galindo, 2020 and Rivera, 2020).

At the same time, the geographical characteristics of the disease spread have been a research topic in epidemiology a long time ago as well as the use of spatial analytical tools in epidemiology and other disciplines. From the famous cholera map by John Snow (1813-1858), father of modern epidemiology, various methodologies and techniques have been developed where geographic analysis is used to understand the processes of spread of infectious diseases. This is documented by Cliff et al., (1981) where, from a multidisciplinary perspective, a historical review is offered on the centrality of geographical elements in the analysis of the dynamics of epidemics. In the same direction, Smallman-Raynor and Cliff (2004) detailed, from a historical perspective, disease spread patterns in military and civil conflicts. The key argument here is that geography is a node that connects epidemiology with other disciplines, in such a way that its analytical capacities are enhanced.

Regarding the CoViD-19 pandemic, there have been attempts to explore the spatial effects of the transmission. Multiple studies use geographical information systems (GIS), to explore the spatial patterns of concentration of infections associated with the characteristics of the population, as well as multiple variables associated with the urban environment (Ghosh and Cartone, 2020; Cordes and Castro, 2020; Sun et al., 2020; Kang et al., 2020, Franch-Pardo et al., 2020; Molallo et al., 2020). The differences at socioeconomic level like income and poverty (Sannigrahi et al., 2020) and specific occupational sectors (Mongey and Weinberg, 2020) have been mentioned as relevant characteristics to explain the spread of infections and risk of contracting the disease. Moreover, the spatial relations between the infection and socioeconomic characteristics are also part of the discussion (Maroko et al., 2020).

In geographical studies and in other disciplines such as medicine there is a problem called ecological fallacy (Schwartz, 1994; Piantosi et al., 1988; O'Sullivan and Unwin, 2010: 39-40; Harris, 2006). In a nutshell, this problem could be characterized as follows: results from geographic data at a certain aggregation level could not be correct when they are extended to others aggregation levels due to the heterogeneity presented in the data and could carry to inexact inferences.

In this context, there are two objectives in this paper. First, we evaluate the spatial distribution of the positive cases and deaths in VMMZ at municipality level. Secondly, we explore the active cases in México City at neighborhood level. The remaining document is structured as follows. The next section shows a general picture of the first pandemic wave at VMMZ and explores the presence of spatial autocorrelation through classical measures. Then, we propose an econometric spatial model in search of capture the geographical effects of the disease at municipality level and its association with economic and social characteristics of the spatial area units between February and September 2020, when the first pandemic wave took place. The next section is dedicated to the spatial exploration of the active cases at neighborhood level only for México City where we compare and comment on the results of this and the previous section in the context of the ecological fallacy problem. The fifth section is dedicated to the discussion of our results and some recommendations.

2. THE PANDEMIC COURSE OF VALLE DE MÉXICO METROPOLITAN ZONE

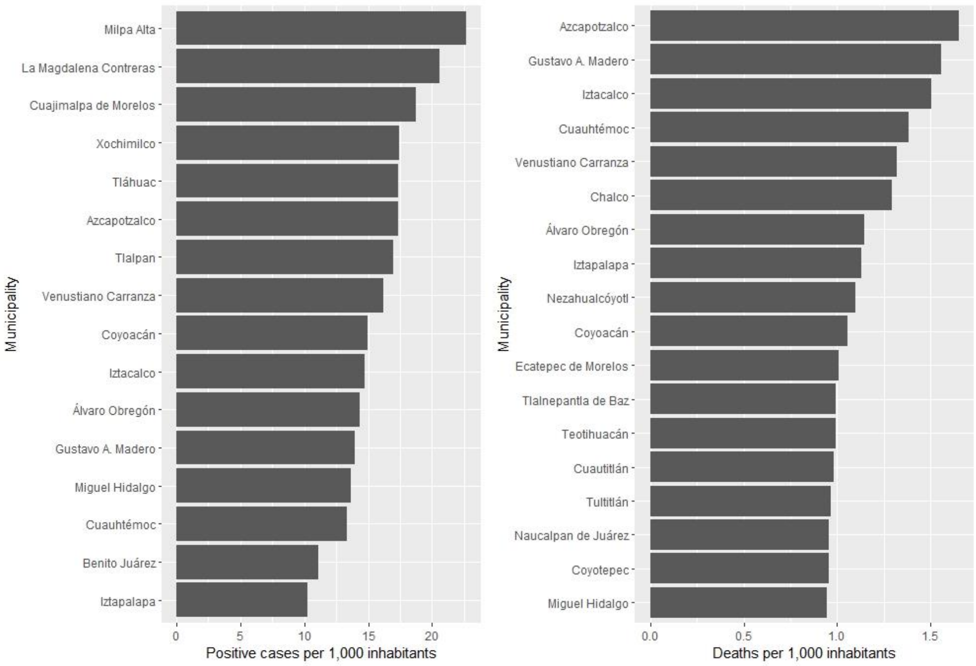
With a population of almost 22 million people by 2020, VMMZ is the biggest metropolitan area in Mexico and one of the most in the world. The VMMZ is integrated by 76 municipalities, a local administrative delimitation in Mexico, that belongs to three different states (intermediate administrative level): México City, Estado de México, and Hidalgo.

Figure 1. Valle de México Metropolitan Zone

Source: Own elaboration based on INEGI (2020). National Geostatistical Framework.

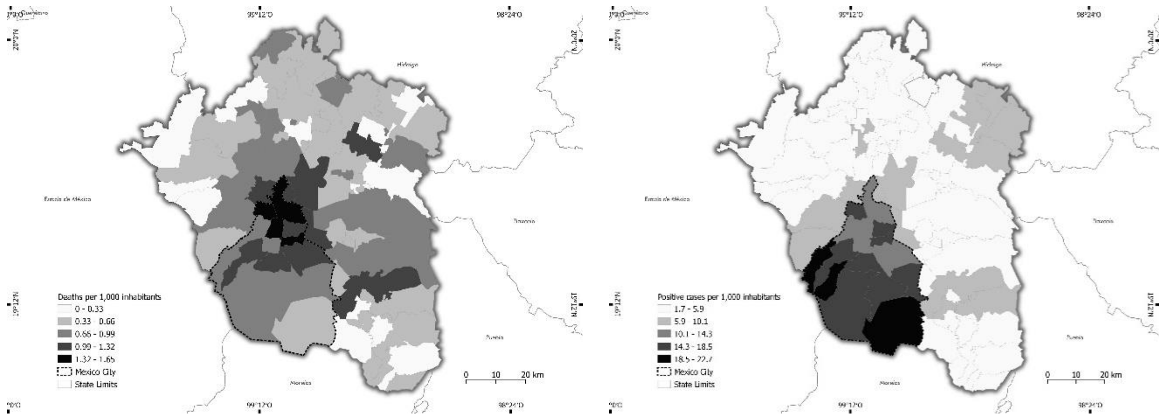
Between March 1st to September 30th 2020, 198 thousand 664 positive cases of CoViD-19 and 20 thousand 497 deaths caused by this disease were recorded at VMMZ, the most in México City. The municipalities of Milpa Alta, La Magdalena Contreras and Cuajimalpa de Morelos, showed the higher number of positive cases while on the other hand, Azcapotzalco, Gustavo A. Madero and Iztacalco, showed the higher number of deaths per thousand people respectively. Moreover, the spatial distribution of both variables differs: whereas the positive cases are in the southern (figure 2 and 3), the deaths are grouped in the center municipalities, with some North and Northeast municipalities in the third classification group.

Figure 2. Positive cases and deaths per 1,000 inhabitants in Valle de México between March 1st to 30th September 2020



Source: Based on National Epidemiological Surveillance System (SINAVE), Mexican Government.

Figure 3. Positive cases and deaths per 1,000 inhabitants in Valle de México between March 1st to 30th September 2020



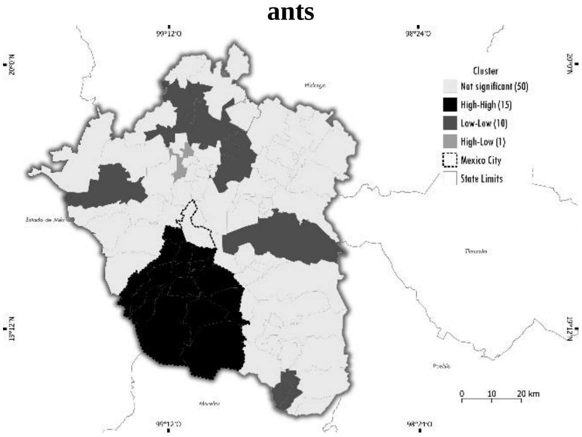
Source: Based on National Epidemiological Surveillance System (SINAVE), Mexican Government.

In order to statistically verify the concentration of data on positive cases and deaths, the Local Indicators of Spatial Association (LISA) were estimated (Anselin, 2020). LISA identifies hot-spots and cold-spots by grouping and evaluating local autocorrelation. This lets us know which municipalities exhibit a spatial relationship with the positive cases and deaths and this can be used to explain the whole process. The results appear in figure 4.

Both positive cases and deaths exhibit a positive statistically significant Moran Index. The null hypothesis (*Ho*) of no spatial autocorrelation is rejected in favor of the alternative hypothesis (*Ha*) of spatial autocorrelation. This is because the values of the estimated index are far from the value of the simulated index, the pseudo p-value is close to zero and the z values are high. Beyond statistical significance, the point to be highlighted in the analysis is that spatial effects are identified in infections. This is evidence of the relevance of capturing those effects on a model, like we tried in the next section.

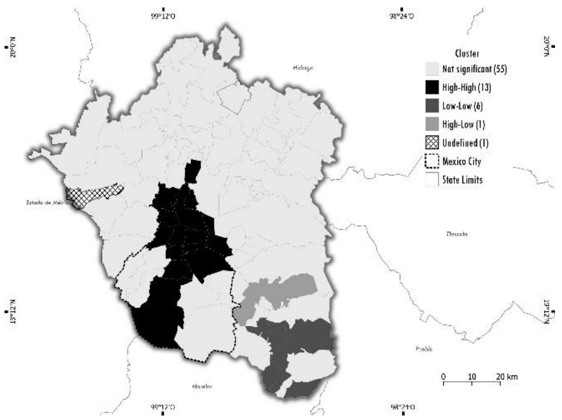
Figure 4. Clusters of positive cases and deaths in the ZMVM

Cluster map of positive cases per 1,000 inhabitants



Moran Index: 0.659 | Simulated Moran Index: -0.0130 | Pseudo-P-value: 0.00001 | Z:8.83| SD: 0.0761 | Permutations: 99,999, Spatial weights contiguity configuration: queen first order.

Cluster map of deaths per 1,000 inhabitants



Moran Index: 0.486 | Simulated Moran Index: -0.0135 | Pseudo-P-value: 0.00001 | Z:6.83 | SD: 0.0771 | Permutations: 99,999, Spatial weights contiguity configuration: queen first order.

Source: Own estimates based on data from the Ministry of Health, Mexico.

3. THE SPATIAL ECONOMETRIC APPROACH AND RESULTS

Spatial econometrics offers some advantages compared to the simple cross-sectional, both technical and theoretical (Baltagui, 2001). On one hand, from the technical point of view, by estimating a spatial model would be possible to solve the problem associated with the errors correlation with an ordinary least squares (OLS) conventional estimation. On the other hand, it is possible to incorporate the spatial nature of such a phenomenon like diseases (Musa, 2013; Rezaeian et al., 2007). Besides the linear model estimated with OLS, we present the results of two different spatial models: lag model (Lag, equation 1) and error model (Err, equation 2). Lag incorporates as an explanatory variable the weighted average of the interest variable, *spatial lag*, through an instrument that captures the spatial relation of the municipalities, spatial weights matrix. Error use the spatial structure to incorporate the autocorrelation though the error terms (Helhorst, 2006: 6):

$$Y = \delta WY + \alpha 1_N + X\beta + \varepsilon \qquad (1)$$

$$Y = \alpha 1_N + X\beta + u \qquad (2)$$
$$u = \lambda Wu + \varepsilon$$

In the Lag model, Y corresponds to the vector of dependent variable, X represents the matrix of explanatory variables, β stands for the estimated coefficients associated with each of the independent variables, W is the spatial structure in both cases, we define the spatial structure through an order one queen spatial weights matrix and WY represents the spatial lag or weighted average of the Y value of the neighbors, δ is the associated coefficient and ε is the error term. Meanwhile, λ refers to the estimated coefficient that captures the spatial autocorrelation through the errors of the model, and ε is the usual pure aleatory error (see Anselin, 2005 and Helhorst, 2006 for details).

We seek to explain the level of the first pandemic wave in VMMZ, positive cases and deaths per thousand inhabitants between March 1st to September 20th, 2020, through four types of variables: i) social, ii) urban and housing, iii) occupational structure and iv) average salaries, such as appears in Table 1.

Table 1. Sociodemographic and economic determinants

		Dependet variables
COVID19	pos	Positive cases per thousand people
	dth	Deaths per thousand people
		Sociodemographic
Social	ss	Population with health services (%).
	hgr	Population >12 yo with higher education (%).
	im	Marginalization 2020 (index).
Urban and housing	den	Population density (people per km^2
	memb8	Population living in households with 8 or more members (%)
	bdr3	Population in dwellings with 3 bedrooms or less (%)
	rm1	Population in dwellings with only one room (%)
		Economic
Occupational structure	poind	Employed personnel in industry (electricity, construction and manufacturing)
	pocom	Employed personnel in commerce (wholesale and retail)
	poss	Employed personnel in services (remaining sectors except aggriculture)
Average salaries	rmind	average salaries in industry (electricity, construction and manufacturing)
	rmcom	average salaries in commerce (wholesale and retail)
	rmss	average salaries in services (remaining sectors except aggriculture)

Positives cases and deaths per thousand people were constructed based on the National Epidemiological Surveillance System (SINAVE) provided for Mexican central government, sociodemographic and economic variables were obtained from the Population Census 2020 and Economic Census 2019, both published by the National Institute of Statistic and Geography (INEGI). We use R software (R core team, 2013) and their packages *spdep* (Bivand, et al. 2018) and *spatialreg* (Bivand et al., 2021) to evaluate spatial dependence and to estimate spatial econometric models, respectively. Based on equation 1 and 2, we propose two different models, the first one for the positive cases and the second one to the deaths, both in logarithmic form:

$$pos_i = \alpha + \delta W pos_i + \beta_1 ss_i + \beta_2 hgr_i + \beta_3 rm1_i + \beta_3 poss_i + \beta_5 rmind_i + u \quad (3)$$

$$u = \lambda W u + \varepsilon$$

$$dth_i = \alpha + \delta W dth_i + \beta_1 im_i + \beta_2 den_i + \beta_3 memb8_i + \beta_3 bdr3_i + \beta_5 poind_i + \beta_6 pocom_i + \beta_7 pos_i + u \quad (4)$$

$$u = \lambda W u + \varepsilon$$

When in equation 3 and 4, $\lambda = 0$ we have a Lag model, and when $\delta = 0$ it is an Err model. Following Anselin (2005), we test a linear model estimate with OLS, Lag and Error models and the spatial diagnosis, as usual. The results, and diagnostics appear on Table 2 and Table 3.

Table 2. Models for positive cases

Variable	OLS	Lag	Err
Intercept	3.939 0.75604 ***	1.382 0.723 *	2.411 0.940 **
ss	1.388 0.68321 **	0.774 0.553	0.874 0.632
hgr	0.508 0.24592 **	0.224 0.204	0.185 0.247
rm1	0.301 0.13658 **	0.107 0.112	0.019 0.146
poss	0.662 0.28587 **	0.315 0.240	0.557 0.245 **
rmind	0.197 0.09208 **	0.175 0.0742 **	0.157 0.071 **
delta (δ)		0.565 0.098 ***	
lambda (λ)			0.678 0.094 ***
sqr R	0.551	0.683	0.695
Adj sqr R	0.519	ND	ND
Akakie C.	105.2	87.0	82.6
N	76	76	76
Spatial Diagnosis OLS			
Moran		0.278	
LM-Lag		17.74 ***	
LM-Err		12.19 ***	
LM-Lag rob		5.54 **	
LM-Err rob		0.000	

Note: *** Significant at the 99% level, ** Significant at the 95% level, * Significant at the 90% level.

Table 3. Models for deaths

Variable	OLS	Lag	Err
Intercept	-34.770 13.204 **	-35.194 12.485 ***	-34.771 12.492 ***
im	9.374 3.375 ***	9.445 3.189 ***	9.377 3.193 ***
den	0.147 0.046 ***	0.155 0.047 ***	0.147 0.044 ***
memb8	0.726 0.232 ***	0.719 0.218 ***	0.728 0.218 ***
bdr3	3.805 1.817 **	3.762 1.716 **	3.816 1.720 **
poind	0.515 0.1603 ***	0.510 0.152 ***	0.515 0.152 ***
pocom	0.497 0.287 *	0.483 0.271 *	0.499 0.271 *
poss	1.397 0.416 ***	1.394 0.394 ***	1.399 0.393 ***
delta (δ)		-0.054 0.122	
lambda (λ)			0.006 0.167
sqr R	0.663	0.664	0.663
Adj sqr R	0.627	ND	ND
Akakie C.	69.0	70.8	69.0
N	76	76	76
Spatial Diagnosis OLS			
Moran		0.002	
LM-Lag		0.183	
LM-Err		0.001	
LM-Lag rob		0.428	
LM-Err rob		0.245	

Note: *** Significant at the 99% level, ** Significant at the 95% level, * Significant at the 90% level.

3.1 Positive cases

All the proposed variables in the OLS model, equation 3, for positive cases are positive and significant and the explanatory variables can explain nearly 50% of the variability of them. The OLS model presents autocorrelation (Moran equal to 0.27), which is an indicative for the relevance of a spatial model. The spatial diagnostic reveals that the Lag model is the best as is shown by the Lagrange multiplier test in their robust form.

The model reveals a positive relation between people with access to health services (ss) and positive cases. This fact perhaps reflects that greater access to social security is linked to an effective detection of cases, which could prevent them from evolving into death. Besides, this variable has the biggest coefficient of all explanatory variables.

It is possible to note a positive association between the places with populations with higher education (hgr) and positive cases, nevertheless it is relevant to keep in mind the fact that there are several municipalities in the VMMZ with people living in semirural areas at peripheral places. So, the identified relation could be explained because this first wave took place fundamentally in larger urban areas, as it is shown in Figure 3, where the people usually have a higher formal education.

The Mexico central government, as in many countries around the world, recommended keeping physical distance to prevent the contagion but it is not always possible to keep adequate physical distance at home when one of the family members catches the virus if there are not enough rooms at the house. So, our model reveals that the greater the proportion of people that live in houses with only one room (rm1), the higher the number of positive cases is.

Like many other big cities, the VMMZ is predominantly a service city with 55% of the occupied people laboring in this sector. The proportion of people occupied in the service sector (poss) has the second bigger coefficient associated with positive cases. Again, the central municipalities exhibit a higher proportion of occupied persons at services and are more densely populated.

Finally, between the economic structure the average salaries at industry (rmind) evince a direct relation with positive cases. While the spatial distribution of the industry salaries does not follow the same pattern as positive cases, this is indicative of the relevance of the economic structure.

Regarding the spatial models, not only the sign of the estimated coefficient associated to the spatial lag, delta, is positive and highly significative, besides it helps to elevate the explanatory power of the model, nevertheless, the other variables are not significant anymore, except the industrial salaries. Something similar occurs with the Error model, where the coefficient lambda is positive and highly significant, too. The spatial diagnosis, according to the Lagrange multipliers, makes clear that between Lag and Error models, we would accept Lag. This result is consistent with Gosh and Cartone (2020), and with our exploratory analysis and shows the importance of considering the spatial correlation present in the context of local transmission of the disease.

3.2 Deaths

The picture revealed by the deaths model, equation 4, is something different, particularly about the spatial models. Like in the positive cases model, all the explanatory variables of the linear model estimated with OLS are positive and significant. Marginalization index (im) shows that high values of the index correspond to low degree of marginalization: the higher the index is, the better the living conditions are. The positive relation between deaths and this index indicates that were in the center places, where the living conditions are better compared with the semirural areas, were where the virus hit stronger.

In the deaths model, we have three different variables linked to urban and housing conditions: population density (den), population living in households with 8 or more members (memb8) and population in dwellings with 3 bedrooms or less (bdr3). These three variables show that the inner home living conditions and the urban context have remarkable consequences in the outcome of the disease because all of them are associated with deaths in a positive way.

We could compare the different importance of the three considered sectors in the course of the pandemic. The results show that the municipalities with the higher proportion of occupied persons in services have a higher number of deaths per thousand persons, followed by the proportion of people in industry sectors, and finally people in commerce.

The spatial diagnosis shows that there is no autocorrelation in the OLS estimation, even more, the coefficients linked to the spatial versions of the models are close to zero and not significant as Lag and Error models reveal. So, it seems that the model that explains in a better way the proportion of deaths in the first pandemic wave at VMMZ is a linear model estimated with OLS.

Through these two models, equation 3 and 4, we incorporate the spatial autocorrelation and point out the relevance of the local transmission of CoViD-19 in VMMZ, remarkably with the positive cases, in the relatively most densely populated places. Notwithstanding, it is important to consider the transmission dynamic linked to the people's mobility patterns, the possibility to work from home, and many other factors that, of course, could enrich these results and we do not incorporate here.

In the introductory section, we refer to the so-called ecological fallacy. It is time now to consider this problem in the next section, not with a model because of the data availability, instead of that we present some explanatory analysis at neighborhood scale.

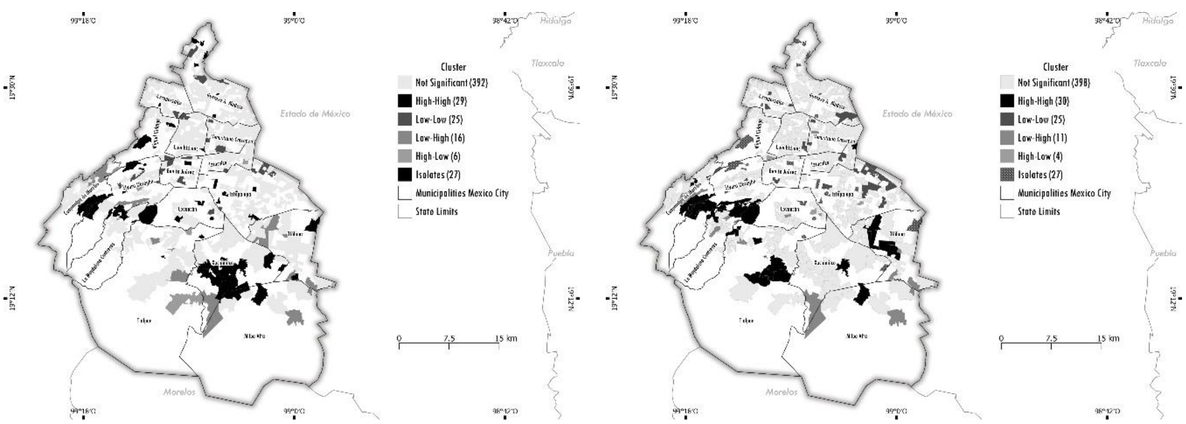
4. ECOLOGICAL FALLACY: HETEROGENEITY OF DATA IN MEXICO CITY

We will turn our attention now to a lower desegregation level conducting an exploration of spatial data on a smaller scale than the municipal one. The government of Mexico City has systematized the CoViD-19 active cases data by neighborhood. Active cases are defined as cases that meet the following: i) test definitive result is positive, ii) symptoms start in less than 15 days, iii) deaths are discarded. Additionally, for privacy issues, the data set considers neighborhoods that have more than 5 cases. We consider an estimation of total active cases by neighborhood for two periods: a) June 29th to July 29th, 2020, and b) between August 31st to September 29th 2020 due the availability information. This allows us to construct a sample of 495 neighborhoods for which there are also sociodemographic data from the Population Census 2020 and economic data from the National Statistical Directory of Economic Units (DENUE), both published by INEGI.

The exploratory analysis at this level allows us to observe from a quantitative and qualitative point of view the spatial patterns of average active cases on a smaller geographical scale, associating them with some social and economic characteristics. We do not propose a model at this scale for the following reasons: i) there is not complete information for all the polygons at the neighborhood level of the VMMZ, and ii) the set of variables used at the municipal level are not the same at this scale of analysis. However, the available information allows us to treat some ideas regarding with ecological fallacy problem.

Figure 5. Clusters (LISA) of active cases for a sample of 495 neighborhoods in Mexico City

Period A: June 29 to July 29, 2020 | **Period B: August 29 to September 31, 2020**



Moran Index: 0.212 | Simulated Moran Index: -0.0021 | Pseudo-P-value: 0.00007 | Z:5.37 | SD: 0.0399, | Permutations: 99,999 | Spatial weights contiguity configuration: queen first order.

Moran Index: 0.235 | Simulated Moran Index: -0.0021 | Pseudo-P-value: 0.00001 | Z:5.90 | SD: 0.0402, | Permutations: 99,999 | Spatial weights contiguity configuration: queen first order.

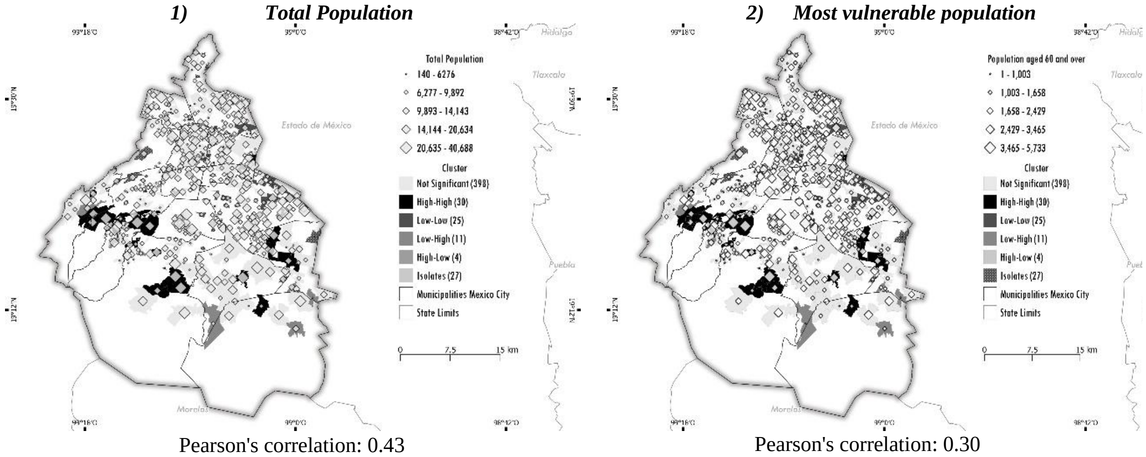
Source: Own estimates based on data from Mexico City Government.

At the municipal level, we found that almost all the spatial units of Mexico City form a high-high cluster for positive cases and deaths. However, when increasing the scale of analysis at the neighborhood level, what is observed is that the hotspots correspond only to some areas in the south of the city. About low-high and high-low clusters, given the spread pattern of the infections, these clusters could be source of spread of the disease at any pandemic moment.

Now, in Figure 6, we present some insights about patterns and spatial relations between active cases between August 31st to September 29th, 2020, and demographic and socioeconomic characteristics.

Figure 6. Clusters (LISA) of active cases for a sample of 495 neighborhoods in Mexico City and socio-demographic characteristics

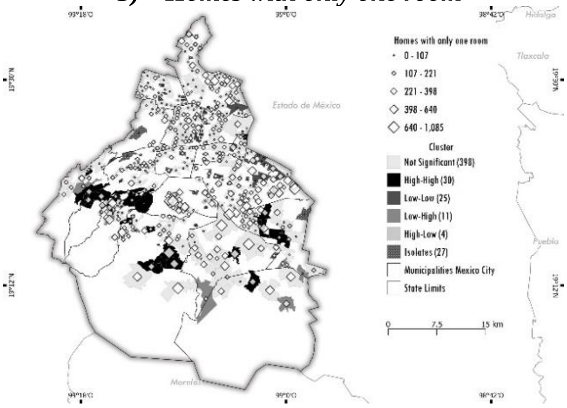
1) Total Population | ***2) Most vulnerable population***



Pearson's correlation: 0.43

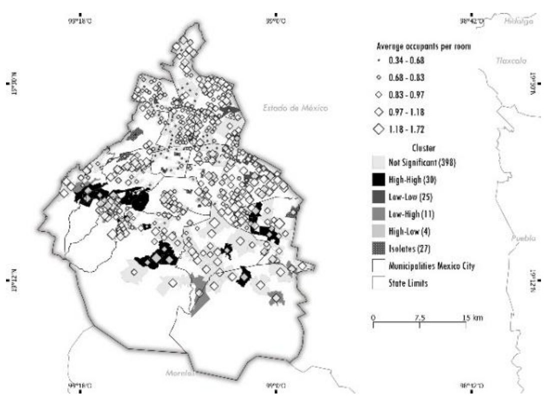
Pearson's correlation: 0.30

3) Homes with only one room



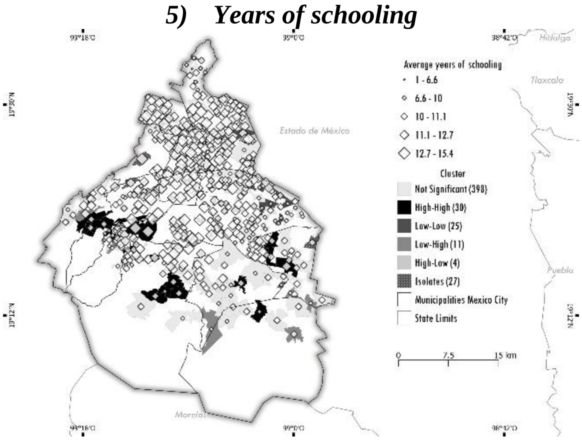
Pearson's correlation: 0.32

4) Average occupants per room



Pearson's correlation: 0.15

5) Years of schooling



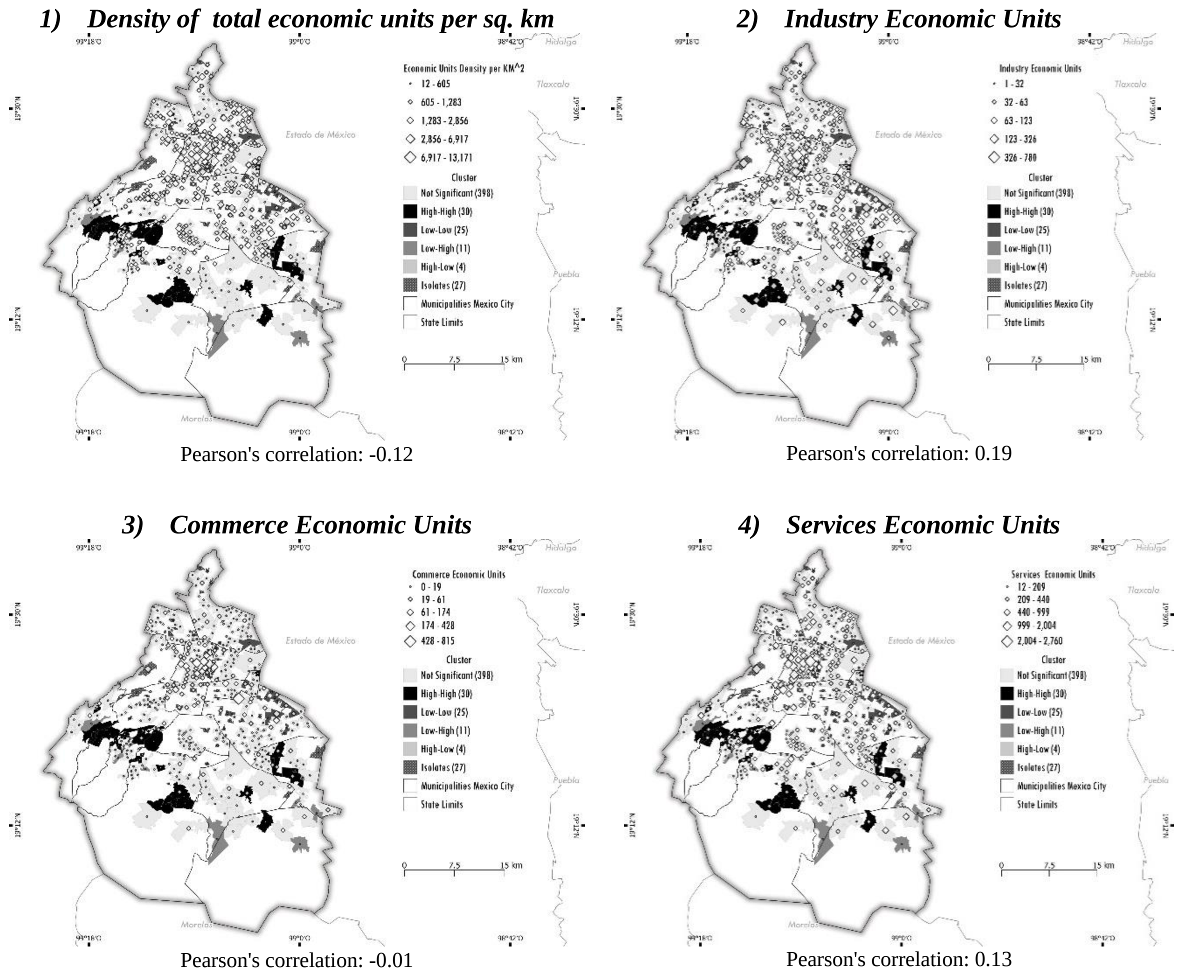
Pearson's correlation: -0.12

Source: Own elaboration based on data from Mexico City Government and INEGI. Population Census 2020. The correlation coefficients are constructed within active cases between August 31st to September 29th and the corresponding point variable in each panel figure.

The patterns observed at Figure 6 could be summarized as: i) for each neighborhood, the greater population is, the more probability of positive cases have, which could require some actions by the policy makers; ii) the relationship between the active cases and the population aged 60 years and over could also be considered strong, particularly because for the period of analysis there were not yet available vaccines; iii) the neighborhoods where there are dwellings with only one bedroom influence largely for the increase in active cases, as we observe in the model for positive cases, equation 3; iv) the average members living in households exhibit a lower association, probably because of the heterogeneity; v) the small negative association between formal education and active cases contrast with the results obtained in the models, this could mean that population with a higher educational level has sometimes better material conditions, more information and respect the rules of confinement, among other advantages.

The results in Figure 7, according now to economic conditions, describe in some sense how the confinement rules act at Mexico City, that is, the closure of non-essential activities. The density of economic units per square kilometer implies in this analysis a low, but negative, incidence in the appearance of contagion cases, which means that the closure measures due to confinement had a moderate impact. In the case of industrial and service economic units, among which are many of those considered essential activities, they had a positive impact on infections, not so in the case of commercial activities, which have a negative relationship. In short, to a greater or lesser extent, the closure of the economy had the expected effects in terms of containing the pandemic in the city. Although it would be relevant to analyze data related to informal activities such as street markets, which could have an inverse impact to the one analyzed here.

Figure 7. Clusters (LISA) of active cases for a sample of 495 neighborhoods in Mexico City and economic characteristics



Source: Own elaboration based on data from Mexico City Government and INEGI. National Statistical Directory of Economic Units. The correlation coefficients contemplate the active cases in period B and the corresponding point variable in each map.

5. DISCUSSION AND FINAL REMARKS

The tracking of this global pandemic is very complex: each wave has their own characteristics on specific areas. Of course, the sociodemographic and economic factors affect that path, but there are many other influence factors like mobility patterns, specific occupations, and specific medical conditions. All these factors are part of the social extraction: catching CoViD-19 when people can afford adequate hospital expenses could have a different ending if not. In this work we present an exploration of the sociodemographic and economic factors that could explain the first pandemic wave in the most populated area in Mexico and explore their spatial characteristics.

The model results depict evidence of spatial association in our two dependent variables: positive cases and deaths. In this first pandemic wave, the most deaths per thousand people took place at north center municipalities (Azcapotzalco, Gustavo A. Madero and Iztacalco), but the places with the higher living conditions were far from the first places (Miguel Hidalgo and Benito Juárez, with high-income neighborhoods took the 18th and 25th place in deaths rank).

We incorporate the spatial relations identified in the exploratory analysis through spatial econometric models. The best model to explain the positive cases in VMMZ at municipality level is the Lag model. The model shed light in order to see how the CoViD-19 hits harder at the municipalities more densely populated and where the urbanization process was deeper, compared with those peripheral: worst living conditions exhibit a positive relationship, both positive cases and deaths, as we observed in relation to urban and housing.

On the other hand, the results could help to explain that economic structure plays a relevant role in the state of the pandemic, like the service sector and industrial occupation reveal. A more detailed study about the specific occupation structure and the jobs that can be conducted at home could help to understand this in depth.

Nevertheless, it is not all a specific territory where the pandemic took place, like a municipality, so we did an exploratory analysis in order to identify the neighborhood's positive cases and their socioeconomic conditions. Although the municipality and neighborhood results are not directly comparable, the last one could help to understand the state of the pandemic at a very low geographical scale.

At the neighborhood level, it was possible to observe how the main cluster of infections and deaths at the municipal level presents a heterogeneous behavior measured with active cases. The geographical conditions of population concentration, overcrowding and some other living conditions have a positive relation on the active cases. Local economic structure measured by proportion of sectoral economic units show that the measures for the closure of non-essential activities is negative associated to active cases.

The results of our econometric models go against what was found in other studies (Jaramillo, 2021; Hernández, 2020), which may be linked to the level of disaggregation used, so they should be taken with caution until reliable information is available with a higher level of detail that allows contemplating the proposed models.

In addition, in Mexico and particularly in urban areas, there is a labor informality rate of close to 50% (INEGI, 2020), which suggests the need to incorporate these elements in our study, to the extent that these types of activities imply a greater risk of exposure due to the greater interaction. A more detailed study of the occupational structure and the jobs that can be done from home can help to understand this in greater depth.

Another aspect that is completely outside what is contemplated by our model is the mobility patterns in this metropolitan zone, present in other studies like Badr, et al (2020) for the USA: incorporating this information in the future is a pending task.

We hope that the results of this paper could shed light to contribute with some ideas in the designing of policies against the economic and social consequences of this and future pandemics, mostly if they consider the necessary presence of spatial effects. We thought that a Comprehensive National Epidemiological Information System with the most possible disaggregation level, as mentioned in Huitrón-Mendoza and Prudencio-Vázquez (2020) not only for CoViD-19, but also for other possible infectious diseases is a key element to construct policies with a multidisciplinary content to save lives.

REFERENCES

Anselin, L. (2005). Exploring Spatial Data with GeoDa: A Workbook. Center for Spatially Integrated Social Science. < https://www.geos.ed.ac.uk/~gisteac/fspat/geodaworkbook.pdf>

Anselin, L. (2020). Local Spatial Autocorrelation (1) LISA and Local Moran. GeoDa. An Introduction to Spatial Data Analysis. Documentation.< https://geodacenter.github.io/workbook/6a_ local_auto/lab6a.html>

Badr, H. S., Du, H., Marshall, M., Dong, E., Squire, M. M., and Gardner, L. M. (2020). Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *The Lancet Infectious Diseases*, 20 (11). https://doi.org/10.1016/S1473-3099(20)30553-3</div>

Baltagi, B. H. (2021). Econometric analysis of panel data. Springer Nature.

Bivand, Roger S. and Wong, David W. S. (2018). Comparing implementations of global and local indicators of spatial association TEST, 27(3), 716-748. URL https://doi.org/10.1007/s11749-018-0599-x.

Bivand, R., G. Millo, and G. Piras (2021). A Review of Software for Spatial Econometrics in R. Mathematics 9 (11):1276. https://doi.org/10.3390/math9111276.

Cliff, A. D., Ord, J. K., Haggett, P., & Versey, G. R. (1981). Spatial diffusion: an historical geography of epidemics in an island community. Cambridge University Press, Cambridge.

Coccia, M. (2020). “Factors determining the diffusion of COVID-19 and suggested strategy to prevent future accelerated viral infectivity similar to COVID”. Science of the Total Environment, 729, 138474. < 10.1016/j.scitotenv.2020.138474>

Cordes, J. and Castro, M. C. (2020). “Spatial analysis of COVID-19 clusters and contextual factors in New York City”. Spatial and Spatio-temporal Epidemiology, 34, 100355. < https://doi.org/10.1016/j.sste.2020.100355>

Dong E., Du H., Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Inf Dis. 20(5):533-534. doi: 10.1016/S1473-3099(20)30120-1. Accessed June 22th 2021.

Franch-Pardo, I., Napoletano, B. M., Rosete-Verges, F. and Billa, L. (2020). Spatial analysis and GIS in the study of COVID-19. A review. Science of The Total Environment, 739, 140033< https://doi.org/10.1016/j.scitotenv.2020.140033>.

Galindo, Jorge (2020). “La pandemia sigue el rastro de la desigualdad en México”, El País, accessed July 05, 2021 <https://elpais.com/sociedad/2020-05-19/la-pandemia-sigue-el-rastro-de-la-desigualdad-en-mexico.html>.

Ghosh, Pritam and Alfredo Cartone, (2020), A Spatio-temporal analysis of COVID-19 outbreak in Italy, Regional Science Policy & Practice, 12(6), pp. 1047-1062 < https://doi.org/10.1111/rsp3.12376>.

Gobierno de México (2021). Casos históricos asociados a COVID-19. Subsecretaría de Prevención y Promoción de la Salud Dirección General de Epidemiología. Accessed June 26th 2021 <https://www.gob.mx/salud/documentos/datos-abiertos-152127>.

Guillén, Pedro (2021). “El COVID-19 sí entiende de clases sociales”, The Conversation, accessed July 05, 2021 <https://theconversation.com/la-COVID-19-si-entiende-de-clases-sociales-163443>

Harris T.M. (2006) Scale as Artifact: GIS, Ecological Fallacy, and Archaeological Analysis. In: Lock G., Molyneaux B.L. (eds) Confronting Scale in Archaeology. Springer, Boston, MA <https://doi.org/10.1007/0-387-32773-8_4>

Hernández, Héctor (2020). “Mortalidad por COVID-19 en México. Notas preliminares para un perfil sociodemográfico”, Notas de Coyuntura del CRIM, N° 36, accessed June 29, 2021 <https://web.crim.unam.mx/sites/default/files/2020-06/crim_036_hector-hernandez_mortalidad-por-COVID-19_0.pdf>

Huitrón-Mendoza, J. A. and Prudencio-Vázquez, J.A. (2020) Gestión Geoespacial de la Pandemia SARS-CoV-2 en México. Boletín Ciudades y Regiones. <https://www.boletinciu dadesyregiones.org/2020/11/gestion-geoespacial-de-la-pandemia-sars.html>.

INEGI (2020). Encuesta Nacional de Ocupación y Empleo, Nueva Edición, ENOEN.

Jaramillo, Máximo (2021). “La pandemia contra los pobres: la Ciudad de México y COVID-19”, Blogs LSE, accessed July 15, 2021 <https://blogs.lse.ac.uk/latamcaribbean/2021/03/02/la-pandemia-contra-los-pobres-la-ciudad-de-mexico-y-COVID-19/>

Kang, D., Choi, H., Kim, J. H. and Choi, J. (2020). Spatial epidemic dynamics of the COVID-19 outbreak in China. International Journal of Infectious Diseases, 94, 96-102. < https://doi.org/10.1016/j.ijid.2020.03.076>

Maroko, A., Nash, D. y Pavilonis, B. (2020). "COVID-19 and Inequity: a Comparative Spatial Analysis of New York City and Chicago Hot Spots", Journal of Urban Health, Vol. 97, 461-470 < https://link.springer.com/article/10.1007/s11524-020-00468-0>.

Mollalo, A., Vahedi, B. and Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. Science of the total environment, 728, 138884 < https://doi.org/10.1016/j.scitotenv.2020.138884>

Mongey, Simon and Alex Weinberg, (2020), Characteristics of workers in low work-from-home and high personal-proximity occupations, Becker Friedman Institute for Economic, White Paper < https://bfi.uchicago.edu/working-paper/characteristics-of-workers-in-low-work-from-home-and-high-personal-proximity-occupations/>.

Musa, G. et al. (2013). “Use of gis Mapping as a Public Health Tool-From Cholera to Cancer”. Health Services Insights, (6), 111-116 < https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4089751/>.

O’Sullivan, D. and Unwin, D. (2010). Geographic Information Analysis, Second Edition, John Wiley & Sons, Inc., New Jersey

Piantadosi, S., Byar, D. P. and Green, S. B. (1988). “The ecological fallacy”, American journal of epidemiology, 127(5), 893-904 < https://doi.org/10.1093/oxfordjournals.aje.a114892>.

Programa de Naciones Unidas para el Desarrollo, PNUD (2019). Informe de Desarrollo Humano Municipal 2010–2015. Transformando México desde lo local. Accessed June 22th <https://www.mx.undp.org/content/mexico/es/home/library/poverty/informe-de-desarrollo-humano-municipal-2010-2015--transformando-.html>.

Qiu, Y., Chen, X., & Shi, W. (2020). “Impacts of social and economic factors on the transmission of coronavirus disease 2019 (COVID-19) in China”. Journal of Population Economics, 33(4), 1127-1172 < https://link.springer.com/article/10.1007/s00148-020-00778-2>.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rezaeian, M. et al. (2007). “Geographical Epidemiology, Spatial Analysis and Geographical Information Systems: A Multidisciplinary Glossary”. Journal of Epidemiology Community Health, (61), 98-102. doi: <10.1136/jech.2005.043117>.

Rivera, Alejandra (2020). “El COVID-19 y las desigualdades sociales”, CLACSO, Observatorio Social del Coronavirus, accessed July 05, 2021 <https://www.clacso.org/wp-content/uploads/2020/04/Alejandra-Rivera-Alvarado.pdf>.

Rivero, Paulina (2020). “El coronavirus tiene clase social”, Revista Nexos, accessed June 30, 2021 <https://www.nexos.com.mx/?p=47908>

Sannigrahi, S., Pilla, F., Basu, B., Basu, A. S., and Molter, A. (2020). “Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach”. Sustainable cities and society, 62, 102418 <https://doi.org/10.1016/j.scs.2020.102418>.

Sannigrahi, Srikanta, Francesco Pilla, Bidroha Basu, Arunima Sarkar Basu, and Anna Molter, (2020), Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach, Sustainable cities and society, 62, 102418 < https://doi.org/10.1016/j.scs.2020.102418>.

Schwartz, S. (1994). The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. American journal of public health, 84(5), 819-824 < https://pubmed.ncbi.nlm.nih.gov/8179055/>.

Smallman-Raynor, M. & Cliff, A. (2004). War epidemics: an historical geography of infectious diseases in military conflict and civil strife, 1850-2000. Oxford University Press, Oxford < https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2289967/>.

Sun, F., Matthews, S. A., Yang, T. C. and Hu, M. H. (2020). “A spatial analysis of the COVID-19 period prevalence in US counties through June 28, 2020: where geography matters?”. Annals of epidemiology. 52, 54-59 < https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7386391/>.