Presented at the DLSU Research Congress 2016
De La Salle University, Manila, Philippines
March 7-9, 2016

Spatial Regression Analysis of Violent Crimes in the National Capital Region

Adelbert W. Choi[1], Vio Jianu C. Mojica[1,*], Frumencio F. Co[1], and Robert Neil F. Leong[1]
[1] Mathematics Department, De La Salle University
**Corresponding Author: vio_mojica@dlsu.edu.ph*

Abstract: Throughout the years, crime analysts continue to regard the incorporation of spatial analytics in crime analyses. This interest, alongside various technological advancements with geographic information systems (GIS), in crime visualization plays a vital role in understanding crime dynamics with the intention of aiding in strategic planning. Utilizing Empirical Bayes smoothed crime rates (per 100,000 population) in 2010, this paper aims to identify the presence of spatial autocorrelation within index crime rates observed in the National Capital Region (NCR) using Moran's I. Spatial units were taken to be the 16 cities and one municipality in NCR. Results revealed that among the seven index crimes, only two, murder and physical injury, were able to exhibit significant spatial autocorrelation. Accounting for the presence of spatial autocorrelation, a regression model which takes spatial dependence into account was used. In particular, spatial lag models were fitted to study the relationships of murder and physical injury with certain demographic variables across cities. It was observed that population density and percentage of male individuals aged 15-24 are found to be positively correlated to the incidence of both of these crimes. On the other hand, factors such as the percentage of married adult population, percentage of foreign immigrants, percentage of the population with a high school diploma, and the crime solution efficiency correlate negatively to the incidence of these crimes. These observed factors could serve as indicators of crime incidences and, thus, monitoring them could aid in crime prevention and security management.

Key Words: spatial autocorrelation; Moran's I; spatial lag model; index crimes

1. INTRODUCTION

Peace and order is a crucial element in sustaining a country's economic development, social order, and political stability. A state of peace and order within a nation can assist in the growth of investments, in the increase in employment opportunities, and in the attraction of foreign visitors (National Economic and Development Authority, 2004). Hence in this context, decreasing crime incidences is recognized in the Philippines Development Plan 2011-2016 as a strategy to promote and maintain public order and internal stability (Senate of the Philippines, 2013).

The Senate of the Philippines (2013) defines crime as "an act committed or omitted in violation of a law forbidding or commanding it and for which a punishment is imposed upon conviction". The Revised Penal Code (RPC) is the fundamental law enumerating the definitions of criminal offenses and their corresponding penalties for the act of such within the Philippines. Reported crimes are sorted into two categories, namely index crimes and nonindex crimes. Index crimes are identified as wrongdoings that occur regularly and that would be serious in nature. They are further subdivided into two categories, crimes against persons and crimes against properties (National Statistical Coordination Board, 2011).

Near the end of the 20th century, along with the advancement of spatial analyses, a growing interest in incorporating spatial properties in crime analysis has been observed. Furthermore, with technological advancements such as computer mapping applications where geographic information systems (GIS) are developed, they have paved the way for efficient measurement and representation of spatial relationships in data. As recently as 2013, utilization of GIS data in the spatial analysis of crime has already been done in the Philippines. Using crime data from 2010-2012, a research by Barrera, Cagang, and Capistrano (2013) had successfully generated spatio-temporal heat maps of reported crimes in Dumaguete City. Particularly, their results showed that crimes tend to be concentrated at or near the central business district.

In the context of crime modeling in particular, a study by Omotor (2010) suggests that certain demographic and socio-economic characteristics are significantly associated with the levels of reported crimes. The results of the study indicate that lagged crime rate, per capita income, and population density are significantly and positively correlated to all forms of crime.

Based on the aforementioned literature, this research is conducted to possibly explore whether the same patterns can be observed in the Philippine setting, specifically in the National Capital Region (NCR). The objectives of the study are (1) to observe spatial patterns of crime in the NCR, and (2) to determine possible association between crime and demographic variables in the cities of NCR while taking possible spatial autocorrelation structure within the variables into account.

From the findings of the study, various policy recommendations may be implemented based on observed spatial patterns and crime determinants. Variables determined to have a tendency to increase crime rates may be monitored accordingly to improve crime control. However, the study is limited only to reported index crimes in NCR in the year 2010. Also, possible crime determinants are limited to mostly demographic variables found in the 2010 Census of Population and Housing. In addition, due to the small number of spatial units (n=17), the regression part of the study will be limited to modeling one independent variable at a time for each crime.

2. METHODOLOGY

2.1 Data

The data used in this study was provided by the National Capital Region Police Office (NCRPO). It consisted of reported counts of index crimes in 2010. To allow comparability of crime figures across different spatial units, counts were expressed as rates per 100,000 population. Then, Empirical Bayes (EB) smoothing was applied to the raw crime rates to account for the variation of population sizes among spatial units. These smoothed rates were used as the response variables.

On the other hand, possible covariates of crimes were obtained from a collection of demographic data from the 2010 Census of Population and Housing. Using total household population as the base, these covariates were expressed as percentages for each spatial unit. In addition, Crime Solution Efficiency (CSE), a measure of effectiveness of the NCRPO's implemented strategies during the periods covered, was also added to the list of possible covariates.

2.2 Analysis

To define neighborhood, binary distance-based weights were used due to the dynamic nature of the population units. The threshold used for distances between the centroids of each areal domain was set as 0.12, which is the median of all possible pairwise distances of the centroids of the spatial units. Hence, a binary weights matrix, ***W***, was then computed as follows:

$$w_{ij} = \begin{cases} 1, \text{ if } d_{ij} \leq 0.12 \\ 0, \text{ if } d_{ij} > 0.12 \end{cases} \quad \text{(Eq. 1)}$$

where:

d_{ij} = distance between centroids of spatial units (*i*) and (*j*)

Spatial autocorrelation was assessed using Moran's *I* indices for all EB smoothed index crime rates. Significance of spatial autocorrelation was based on a permutation test with 99,999 permutations. Of the seven crimes considered, only two registered positive spatial autocorrelation (murder and physical injury). Hence, only the two crime rates were used in the proceeding study.

The selection of covariates used in modeling was initially tested for their significance with both crime rates using Pearson's correlation coefficient with SAS®9.3. Covariates resulting to have a significant correlation with either one of the crime rates were considered as independent variables in modeling both murder and physical injury. Furthermore, only one covariate was chosen amongst covariates that were observed to be highly correlated with one another. From an initial list of 21 possible covariates, only six covariates were used in the spatial regression modeling (see Table 1).

Table 1.Determinants of Crime

Variable	Description
POPDEN	Population Density(Population/Area in square kilometers)
PMALE1524	Percentage of males between 15-24 years of age
MARRIED20	Percentage of married adult population (ages ≥ 20)
HSDIPLOMA	Percentage of the population with a high school diploma
FOREIGN	Percentage of the population residing in the Philippines but previously resided in another country in the past 5 years
CSE	Crime Solution Efficiency (Total Solved Crimes/Total Reported Crimes)

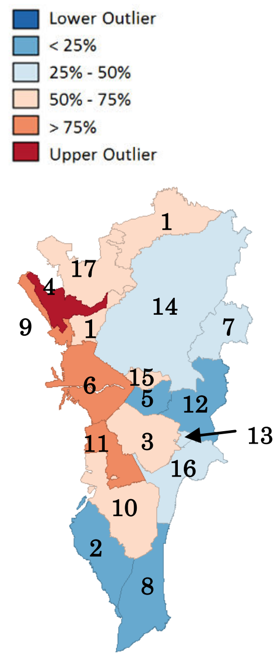
This study followed a spatial regression model decision process as suggested by Anselin (2005). The suggested process selects either the spatial lag model or the spatial error model based on the significance of Lagrange Multiplier (LM) statistics, LM-lag or LM-error, generated after fitting an Ordinary Least Squares (OLS) regression model. In some situations, robust counterparts of these statistics were used in to selecting the appropriate spatial regression model. These statistics. along with spatial regression analysis were done using *GeoDa*™.

3. RESULTS AND DISCUSSION

3.1 Descriptive Statistics

In 2010, murder rates in the NCR were identified to be more prevalent within cities located to the west of the region. These rates were observed to be less within the eastern and southern areas of the region (see Fig. 1). The cities of Malabon, Manila, Navotas, and Pasay are among the four cities with high murder rates. These cities were reported to have murder rates that are greater than the 3rd quartile (6.0042/100,000). On the other hand, cities observed to have murder rates lower than the 1st quartile (3.2537/100,000) consists of Pasig, Mandaluyong, Muntinlupa, and Las Piñas.

Most of the cities that were mentioned to have high murder incidences are observed to be alongside the coast of Manila Bay. Clustering of spatial units with values above the median rate (3.5408/100,000) and those with values lower than the median rate could be visualized through the map, as shown in Fig. 1. From this observation, it could be seen that positive spatial autocorrelation is evident in murder rates.

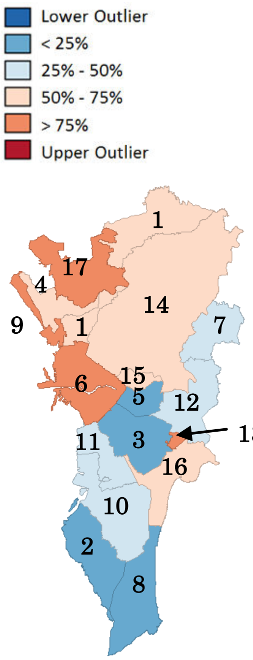


	City	Incidence Rate (per 100,000)
1	Caloocan	6.0042
2	Las Piñas	1.6932
3	Makati	3.5408
4	Malabon	10.8911
5	Mandaluyong	2.2652
6	Manila	9.7884
7	Marikina	3.2537
8	Muntinlupa	1.7198
9	Navotas	8.0019
10	Parañaque	4.5372
11	Pasay	7.9057
12	Pasig	2.9021
13	Pateros	3.4486
14	Quezon City	3.4027
15	San Juan	4.4353
16	Taguig	3.4027
17	Valenzuela	4.1360

Fig. 1. Quartile maps of EB smoothed murder rates in 2010

Physical injuries were observed to be more prevalent in the northern cities of NCR while southern cities tend to exhibit lower physical injury rates (see Fig. 2). The cities of Manila, Navotas, Valenzuela, and the municipality of Pateros were reported to have physical injury rates greater than the 3rd quartile (74.3388/100,000). On the other hand, cities with rates below the 1st quartile (27.0550/100,000) include Mandaluyong, Makati, Muntinlupa, and Las Piñas.

Clustering of spatial units with values above the median rate (40.2528/100,000) and those with values lower than the median rate could be visualized through the map, as shown in Fig. 2. From this observation, it could be seen that positive spatial autocorrelation is also evident for physical injury rates.



	City	Incidence Rate (per 100,000)
1	Caloocan	40.2528
2	Las Piñas	10.3739
3	Makati	16.1144
4	Malabon	74.3388
5	Mandaluyong	23.6638
6	Manila	127.4777
7	Marikina	27.0550
8	Muntinlupa	12.2964
9	Navotas	108.4770
10	Parañaque	27.9187
11	Pasay	34.1231
12	Pasig	28.6620
13	Pateros	88.5715
14	Quezon City	72.1431
15	San Juan	73.8992
16	Taguig	45.4789
17	Valenzuela	106.3117

Fig. 2. Quartile maps of EB smoothed physical injury rates in 2010

The coefficient of variation for murder rates (0.5792) was observed to be lower than that of physical injury rates (0.6895). A high level of variability suggests the presence of large and dispersed rates, indicating possible existence of crime hot spots. Based on the coefficients of variation for the two crime rates, hot spots for murder rates were observed to be divided entirely into only two clearly visible clusters across areas in NCR (see Fig. 1), contrary to physical injury rates, where hot spots do not tend to cluster into a single area in the region (see Fig. 2).

3.2 Spatial Regression

Both LM-lag and LM-error were not found to be significant for murder with every covariate. Hence, ordinary least squares (OLS) regression models were used for murder.

As suggested by the p-values of the coefficient estimates in Table 2, only four covariates were found to be significantly associated with murder rates. Specifically, it was observed that population density and the percentage of male individuals between 15-24 years of age are positively associated with murder rates. These suggest that more densely populated cities or cities with more male individuals aged 15-24 are more likely to have higher murder incidences. Inverse relationships were observed between murder rates and covariates such as the percentage of married adult population aged 20 or older and the percentage of the population with a high school diploma. These covariates are related with lower murder rates if their corresponding proportions are high.

Positive coefficients such as the one exhibited by population density with murder could be interpreted as follows: a unit increase in the covariate contributes an increase in the murder rate in a certain spatial unit equivalent to the regression coefficient. For instance, an increase of 1,000 in the population density of a spatial unit could contribute one unit increase in the murder rate of that spatial unit. On the other hand, negative coefficients are interpreted inversely.

Table 2. Coefficient estimates of OLS regression model for murder

Covariate	Coefficient	Standard Error	p-value
POPDEN	0.0001	<0.0001	0.0159**
PMALE1524	340.1825	172.9259	0.0679*
MARRIED20	-168.7715	35.7618	0.0003***
HSDIPLOMA	-29.3645	12.7276	0.0357**
FOREIGN	-667.7392	391.7984	0.1089
CSE	-3.3490	3.8856	0.4023

*** significant at α=0.01
** significant at α=0.05
* significant at α=0.10

For physical injury, the spatial lag model was fitted for every covariate since it was found to be the most appropriate as indicated by LM diagnostics.

As suggested by the p-values of the coefficient estimates in Table 3, only the percentage of married adult population aged 20 or older and the percentage of the population residing in the Philippines but previously resided in another country in the past 5 years were found to be significant for physical injury at the specified significance levels. Both covariates exhibited negative coefficients, which indicate that high values for these covariates are associated with lower physical injury rates. Interpretations for these coefficient estimates are as follows: a unit increase in the covariate contributes a decrease in physical injury rates equivalent to the magnitude of the corresponding coefficients. For instance, a 1% increase in married adult population aged 20 or above in a spatial unit could contribute a decrease of about 1,289 units in the physical injury rate of that spatial unit. It should be noted that the obtained coefficient estimates were duly adjusted for the presence of high spatial autocorrelation in physical injury rates.

Table 3. Coefficient estimates of spatial lag model for physical injury

Covariate	Coefficient	Standard Error	p-value
POPDEN	0.0008	0.0005	0.1311
PMALE1524	1,645.5990	2,061.5340	0.4247
MARRIED20	-1,288.7780	545.9958	0.0183**
HSDIPLOMA	-234.2658	153.4054	0.1267
FOREIGN	-8,870.9120	4,199.3000	0.0347**
CSE	-56.0235	41.3769	0.1757

*** significant at α=0.01
** significant at α=0.05
* significant at α=0.10

Also, it must be noted that the lag coefficients, ρ, which represent the added neighborhood effect on y, are all positive and significant for all covariates. Their significance suggests an association between physical injury rates in a certain city with physical injury rates of its neighboring cities. For instance, with the covariate defining the percentage of married adult population aged 20 or older, the estimated value of 0.4924 for ρ indicates that a city's physical injury rate is expected to increase by 4.924 if the average physical injury rates of its neighboring cities increased by 10 per 100,000 population.

Table 4. Lag coefficient estimates of spatial lag model for physical injury

Covariate	Lag Coefficient	Standard Error	p-value
POPDEN	0.6216	0.2306	0.0070***
PMALE1524	0.6092	0.2383	0.0106**
MARRIED20	0.4924	0.2764	0.0749*
HSDIPLOMA	0.5346	0.2690	0.0470**
FOREIGN	0.5306	0.2651	0.0454**
CSE	0.5566	0.2607	0.0328**

*** significant at α=0.01
** significant at α=0.05
* significant at α=0.10

3.3 Discussion

Based on the results obtained, some of the covariates considered are significantly associated with murder rates. According to Cole and Gramajo (2009), high population density could highlight inequality and therefore, provide more opportunities for crime. In the same study, it was also noted that a higher level education, although only particular to male education, could be associated to lower crime rates. Virola (2002) mentioned that family members not being victimized by crimes against persons are associated with improved quality of life or reduction of poverty. Both population density and educational attainment or lack thereof could be related to poverty and unemployment, which in turn, could be associated with increased incidence of crimes against people, particularly, murder.

The relationship between sex and age with murder was observed to be positive, as shown by the association of percentage of male individuals aged 15-24 with murder rates. Various studies have also suggested this positive association due to different possible reasons. One possibility is that people within the scope of the variable are more exposed to lifestyles which influence drug and alcohol use. In a study by Njord, Merrill, Njord, Lindsay, and Pachano (2010), drug use was characterized in Filipino adolescents, particularly among street children. Additionally, Dawkins (1997) stated that substance use, mainly marijuana and alcohol, are continuously present in violent crime offenses among adolescents in the United States of America.

Some of the covariates were considered to be significantly associated with physical injury rates. The percentage of married adult population aged 20 or older appeared to be the only covariate significant for both murder and physical injury. It is negatively related with the two crimes. This result validates the findings from a study by Sampson, Laub, and Wimer (2006), which showed that being in a state of marriage is associated with a reduction in the probability of crime incidence. The study also suggested that this finding might be influenced by the opportunities provided by marriage to invest in new relationships that offer social growth.

Lastly, the percentage of the population currently residing in the Philippines but previously residing in a foreign country for the past five years was found to be negatively associated with physical injury only. This finding is an interesting addition to the results of a study by Bianchi, Buonanno, and Pinotti (2008), where it was shown that immigration has been able to significantly affect murder and robberies in Italy. The observed association of immigration and physical injury suggests that the former is a justifiable determinant of crime. Also, since this covariate includes Filipinos who may have worked abroad then came back to the Philippines, it may be associated with improved quality of life or lower poverty rates which may consequently lead to lower crime rates.

4. CONCLUSIONS

Based on EB smoothed rates, murder and physical injury have exhibited significant spatial autocorrelation. These results suggest that crime incidences in a certain city are correlated with the incidences of nearby areas. Thus, information regarding the incidences in nearby cities can aid in alerting a rise or fall of incidences in a particular city.

Results from the fitted regression models revealed that some covariates have significant relationships with murder rates and physical injury rates. These covariates can be used to assist in determining future crime incidences across cities in NCR. Hence, constantly monitoring them could help in the strategic planning of policies and procedures for the prevention of these crimes.

The following recommendations may be considered for future researches related to this study: (1) more granular spatial units may be considered (e.g., barangay level), to allow researchers to better observe spatial patterns and make better inferences to the individuals across these areas; (2) panel data obtained through different time periods may be utilized to observe possible temporal patterns of the distribution of crimes; and (3) other variables that may exhibit significant association with crime incidences may also be examined.

5. ACKNOWLEDGMENTS

The researchers would like to extend their sincerest gratitude to the National Capital Region Police Office's Regional Investigation and Detective Management Division for providing the essential data utilized in this particular study.

6. REFERENCES

Anselin, L. (2005).*Exploring Spatial Data with GeoDa™: A Workbook.* Urbana, IL: Center for Spatially Integrated Social Science.

Barrera, D.J., Cagang, S., & Capistrano, D. (2013). Spatial and Temporal Maps of Reported Crimes in Dumaguete City, Negros Oriental, Philippines. *PRISM 2013, 18,* 11-24.

Bianchi, M., Buonanno, P. and Pinotti, P. (2012), DO IMMIGRANTS CAUSE CRIME?.*Journal of the European Economic Association*, 10: 1318–1347.doi: 10.1111/j.1542-4774.2012.01085.x

Cole, J.H., & Gramajo, A.M. (2009). Homicide Rates in a Cross-Section of Countries: Evidence and Interpretations. *Population and Development Review*, 35: 749–776.doi: 10.1111/j.1728-4457.2009.00307.x

Dawkins, M.P. (1997). Drug use and violent crime among adolescents. *Adolescence,* 32: 395-405.

National Statistical Coordination Board.(2011). Defense, Crime and Delinquency. *Philippine Yearbook 2011.* Retrieved from https://psa.gov.ph/sites/default/files/2011PY_Defense%2C%20Crime%2C%20and%20Delinquency.pdf.

National Economic and Development Authority. (2004). Medium-Term Philippine Development Plan 2004-2010. Retrieved from http://www.philconstruct.com/docs/downloads/NEDA%20DOCS/Medium%20Term%20Development%20Plan%202004-2010.pdf.

Njord, L., Merrill, R.M., Njord, R., Lindsay, R., & Pachano, J.D. (2010).Drug use among street children and non--street children in the Philippines. *Asia Pacific Journal of Public Health, 22,* 203-211. doi: 10.1177/1010539510361515.

Omotor, D.G. (2010). Demographic and Socio-Economic Determinants of Crimes in Nigeria: A Panel Data Analysis. *Journal of Applied Business & Economics, 11,* 181-195.

Republic Act No. 3815, An Act Revising the Penal Code and Other Penal Laws [Philippines], Retrieved August 30, 2015 from http://www. un.org/Depts/los/LEGISLATIONSAND TREATIES/PDFFILES/PHL_revised_penal _code.pdf.

Sampson, R.J., Laub, J.H., &Wimer, C. (2006). Does Marriage Reduce Crime? A Counterfactual Approach to Within-Individual Causal Effects. *Criminology, 44,* 465-508.

Senate of the Philippines. (2013). Crime Statistics At A Glance. Retrieved from https://www.senate.gov.ph/publications/AAG%202013-05%20-%20Crime%20Statistics.pdf.

Virola, R.A. (2002). *Statistical Challenges on Poverty Reduction in the Philippines.* Retrieved from http://www.nscb.gov.ph/technotes/poverty/partii/annex4_2.pdf