Understanding unsolved crimes hotspots: a spatial approach

Cabral Juan Andrés

Work in progress

Abstract

In this paper, unsolved crime hotspots are investigated through a spatial approach. Using data from more than 52,000 homicides, throughout the last decade, in the largest cities in the United States, common tests are carried out to detect the presence of spatial dependence, and later spatial mapping and modeling tools are used. The analysis carried out is intended to be useful to understand (and prevent) those crimes that are not solved.

Introduction

Over the past few years, numerous studies have investigated the determinants and effects of crime hotspots. Many studies suggest that crime is not evenly distributed throughout space, but rather is concentrated in certain areas called "hotspots" (L. W. Sherman, Gartin, and M. E. Buerger 1989, D. Weisburd et al. 1992 and Eck and D. L. Weisburd 2015).[1] As defined by L. W. Sherman 1995 hotspots are small places in which the occurrence of crime is so frequent that it is highly predictable, at least over one year. Given this definition, no doubt knowing which areas are hotspots can be used to fight crime. Therefore the analysis of these areas is very useful to reduce crime.

An important part of the literature on hotspots stick to the following structure: mapping crimes, identifying hotspots through a particular method, using some type of spatial dependency test and later using a particular approach, such as trying to model spatial dependency (see Baller et al. 2001). On the other hand, another part of the literature uses OLS models to see the impact of hotspots on some variable (see, for example, Ceccato and Wilhelmsson 2020 where they see the impact of hotspots on property prices). Finally, there are those studies that try to see which variables can explain crime without explicitly including spatial dependence. An example is that of Favarin 2018, where he shows that in Milan Italy there are hotspots, that they are stable and try to explain them through different factors using negative binomial regressions models.

Since hotspot phenomena are essentially spatial phenomena, it is difficult to find a study that omits tools of spatial econometrics entirely. Furthermore, all the studies cited here that have performed a Moran's I or similar test obtained results that suggest spatial dependence, so ignoring spatial effect in this type of analysis is not very common.

As commented before, one of the reasons why investigating crime hotspots is relevant is because there is abundant evidence on the effectiveness of police interventions in hotspots (see for example Kuo, Lord, and Walden 2013 and for a summary of the available evidence see Braga, Papachristos, and Hureau 2014).

In this work, the hotspots of unsolved crimes are investigated through spatial analysis. As said before, over the years, hundreds of papers have been devoted to analyzing the

[1] Some authors write "hotspots" as "hot spots".

hotspots of different types of crimes and their effects on other variables, but to date, no study has looked specifically at hotspots of those crimes that are not solved.

This paper contributes mainly to two pieces of literature. On the one hand, it analyzes crime hotspots like many other works. It will seek to replicate some stylized facts, such as showing that crimes are not distributed randomly but that there are clusters. On the other hand, this work contributes to crime prevention literature. Much of the criminology work focuses on crimes without discriminating whether they were solved or not. Focusing on unsolved crimes can be useful to dedicate resources precisely to avoid those crimes that for various reasons cannot be solved.

Spatial Econometrics background

As Herrera Gómez et al. 2017, p. 2 comments, the birth of spatial econometrics is located around 1979. In its early years, the models that explicitly incorporated space were focused on issues of geographic or urban economy, however, over the years, spatial econometric models began to be applied in other areas. Among the areas that were influenced by spatial econometrics is the study of crime.

Given that the phenomenon of hotspots is related to non-random distributions in space, the use of spatial tools to identify clusters is commonplace (for example, many of the examples used in spatial econometrics textbooks are related to criminology, see Sarrias 2020, LeSage 2015 and Elhorst 2014). One of the first steps in studies that analyze hotspots is to map these hotspots, for this, there are various methods as dot maps, line maps, ellipse maps, etc. Depending on the type of data and what you want to show, different methods are used. Line maps, for example, are useful when hotspots are along streets whereas Dot Maps, are used when hot spots are at specific addresses, corners, and other places. Ellipse and choropleth maps are used when the designated hotspots share the same risk level, so a specific location inside that area is irrelevant (see Table 1 for more methods and examples).

As previously stated, hotspot analyzes begin through the use of some spatial dependency test, whether local or global. Some of these tests are Moran's I and Geary's C. These tests are generally computed through particular softwares, as an example, the CrimeStat spatial software includes the above tests (Levine 2006, p. 44). Local Moran's I tests are common in the literature of crime hotspots; Kerry et al. 2010 investigates car theft, Murray et al. 2001 investigates property crimes, Baller et al. 2001 turned their attention to homicides and Cohen and Tita 1999 attempted to identify clusters in violent crimes.

When looking for jobs that include spatial analysis, it is convenient to review the most recent jobs because these tend to use particular tools of spatial econometrics. Ahmed and Salihu 2013 investigate hotspots for various categories of crime in Nigeria through interpolation methods (particularly using the Inverse Distance Weighted method). In most papers, crimes are mapped, however, not many stop to analyze the predictive capacity of the method used. Chainey, Tompson, and Uhlig 2008 show that it is not only relevant to be able to map hotspots but also to evaluate the predictive capacity of crime, since it is useless to find hotspots if they will disappear in the future. By comparing various methods, it shows how effective each one is in predicting crime.

As mentioned above, many works perform spatial dependency tests, Chakravorty 1995,

Table 1: Tools used in crime concentration investigation

Methods	Example
Mapping	
Dot maps	Ceccato and Wilhelmsson 2020
Line maps	Telep and Hibdon 2019
Ellipse maps	Zhang et al. 2010
Grid thematic maps	Chainey, Tompson, and Uhlig 2008
Choropleth/thematic maps	Chainey, Tompson, and Uhlig 2008
Isoline maps	Telep and Hibdon 2019
Interpolation and continuous surface smoothing methods	Chainey, Tompson, and Uhlig 2008
Test for spatial auto correlation	
Moran's I	Baller et al. 2001
Geary's C	Eck, Chainey, et al. 2005
Local Indicators of Spatial Association (LISA) statistics	
Gi and Gi*	J. H. Ratcliffe and McCullagh 1999
Local Moran's I	Rogerson and Kedron 2012
Local Geary's C	Anselin 2019
Commonly used software	
CrimeStat	Levine 2006
GeoDa	Leitner and Brecht 2007
ArcView	Brunsdon, Corcoran, and Higgs 2007
Regression models	
Spatial lag model (SLM)	Kubrin and Weitzer 2003
Spatial error model (SEM)	Miles-Doan 1998

Table 1 shows a summary of different tools used to detect and analyze hotspots and spatial interactions in crime related research. The table organization is mainly based on the works of Chainey, Tompson, and Uhlig 2008 and Chainey and J. Ratcliffe 2013.

p. 57 through an analysis of Philadelphia, recommends that to locate hotspots it is advisable to perform LISA tests instead of global tests because LISA tests identify small local clusters when large heterogeneous areas are used. On the other hand, Rogerson and Kedron 2012 suggests that due to the problem of using different weighting matrices it is necessary to adjust for the multiple tests performed, in the work a local Moran's I test is developed that adjusts by testing with multiple weighting matrices.

As mentioned before, there are works that explicitly model special effects. What is interesting about the work of Baller et al. 2001 is that they use models that explicitly take spatial effects into account and compare them. The authors find that there is residual spatial autocorrelation in all the investigated periods, in addition to the fact that there are diffusion effects in which certain crimes in some counties of United States have an influence on crimes in others.

Law, Quick, and Chan 2015 investigated about hotspots for violent crime in Toronto (2006-2007), through a Bayesian approach, they identify hotspots based on the trend from 2006 to 2007. An interesting aspect of this work is that it adopts a non-frequentist approach that according to the authors is convenient from a law enforcement perspective. This type of study is not very common.

It is also relevant to highlight that the works that analyze hotspots do so by selecting a certain time window (commonly months or years), these works then have panel data so it is also possible to perform spatio-temporal analyzes. Among those works that study crime hotspots with a spatio-temporal approach is He, Páez, and Liu 2017 who study the temporal persistence of hotspots. The study shows that certain socio-economic factors manage to explain the presence of violent crimes in hot zones.

Data

The data comes from a Washington Post work. The database contains more than 52,000 homicides in the past decade in the 50 largest cities in the United States. The database includes the location (latitude and longitude), if the case is closed and if there were any arrests among other basic data of the victim (race, sex, age). This database was built through data that comes from different sources. As the repository from Washington Post said, reporters worked for months to clean up the data and standardize it. In comparisons with FBI data, it was verified that the database is consistent with other sources.

A homicide is considered to be closed by arrest when police reported that to be the case. Cases were counted as closed without arrest if they were reported to be "exceptionally cleared". This means that there is enough evidence but an arrest is not possible, an example of this is when the suspect has died. All other cases were classified as having no arrest.

Some analysis of the database can be found at Washington Post. More information on the database can be found in a Github Repository.

Descriptive statistics

Some descriptive statistics can be seen in Table 2. Most of the variables are dummies. It is possible to see that the victims tend to be relatively young (around 30 years old) or

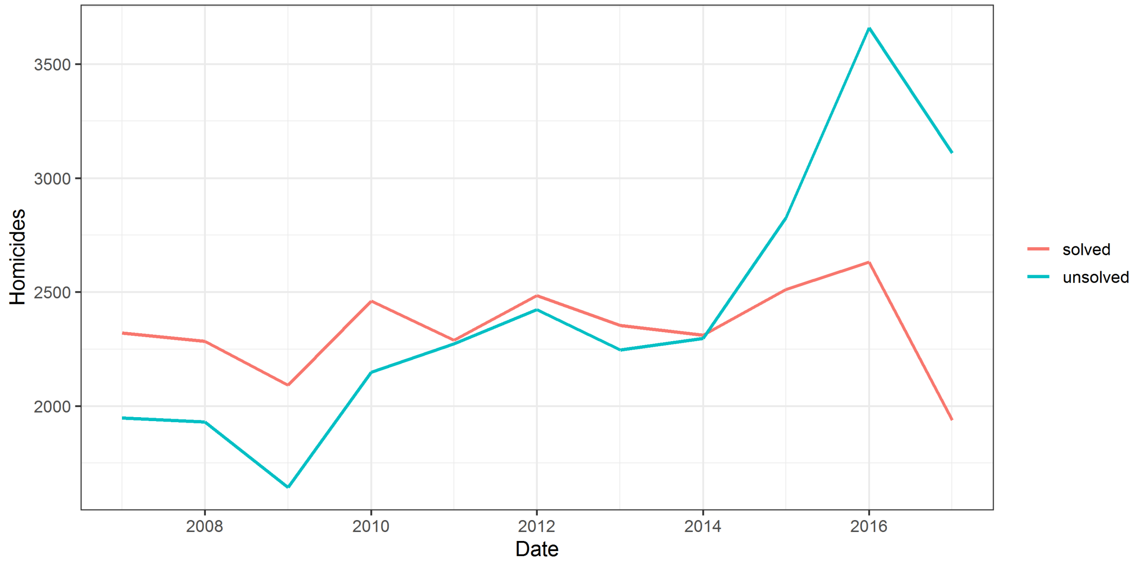

Figure 1: Evolution of solved and unsolved homicides

that the cases that are closed and not resolved are only 5%.

Some relationships that can be observed through this dataset is that the proportion of unsolved cases (open or closed without arrest) is approximately 51%. On the other hand, through the data of the victims we can see that in the case that the victim is a black person, it is more likely that the case is not resolved (54% for black people, compared to 37% for people white). Something similar happens with men, men's cases tend to be resolved less (53% compared to 42% in the case of women).

Table 2: Descriptive statistics

Statistic	N	Mean	Min	Max
open_or_no_arrest	52,179	0.508	0	1
closed_without_arrest	52,179	0.056	0	1
arrest	52,179	0.492	0	1
black	52,179	0.639	0	1
white	52,179	0.121	0	1
male	52,179	0.781	0	1
age	49,180	29.120	1.000	101.000

A temporal analysis shows us that over time, there was a greater increase in unsolved crimes than in solved crimes (Figure 1).

Methodology

Results

Conclusions

References

Ahmed, M and RS Salihu (2013). "Spatiotemporal pattern of crime using geographic information system (GIS) approach in Dala LGA of Kano State, Nigeria". In: *American Journal of Engineering Research (AJER)* 2.3, pp. 51–58.

Anselin, Luc (2019). "A local indicator of multivariate spatial association: extending Geary's C". In: *Geographical Analysis* 51.2, pp. 133–150.

Baller, Robert D et al. (2001). "Structural covariates of US county homicide rates: Incorporating spatial effects". In: *Criminology* 39.3, pp. 561–588.

Braga, Anthony A, Andrew V Papachristos, and David M Hureau (2014). "The effects of hot spots policing on crime: An updated systematic review and meta-analysis". In: *Justice quarterly* 31.4, pp. 633–663.

Brunsdon, Chris, Jonathan Corcoran, and Gary Higgs (2007). "Visualising space and time in crime patterns: A comparison of methods". In: *Computers, environment and urban systems* 31.1, pp. 52–75.

Ceccato, Vania and Mats Wilhelmsson (2020). "Do crime hot spots affect housing prices?" In: *Nordic Journal of Criminology* 21.1, pp. 84–102.

Chainey, Spencer and Jerry Ratcliffe (2013). *GIS and crime mapping.* John Wiley & Sons.

Chainey, Spencer, Lisa Tompson, and Sebastian Uhlig (2008). "The utility of hotspot mapping for predicting spatial patterns of crime". In: *Security journal* 21.1-2, pp. 4–28.

Chakravorty, Sanjoy (1995). "Identifying crime clusters: The spatial principles". In: *Middle States Geographer* 28, pp. 53–58.

Cohen, Jacqueline and George Tita (1999). "Diffusion in homicide: Exploring a general method for detecting spatial diffusion processes". In: *Journal of quantitative criminology* 15.4, pp. 451–493.

Eck, John, Spencer Chainey, et al. (2005). "Mapping crime: Understanding hotspots". In:

Eck, John and David L Weisburd (2015). "Crime places in crime theory". In: *Crime and place: Crime prevention studies* 4.

Elhorst, J Paul (2014). *Spatial econometrics: from cross-sectional data to spatial panels.* Vol. 479. Springer.

Favarin, Serena (2018). "This must be the place (to commit a crime). Testing the law of crime concentration in Milan, Italy". In: *European Journal of Criminology* 15.6, pp. 702–729.

He, Li, Antonio Páez, and Desheng Liu (2017). "Persistence of crime hot spots: An ordered probit analysis". In: *Geographical analysis* 49.1, pp. 3–22.

Herrera Gómez, Marcos et al. (2017). *Fundamentos de Econometrıa Espacial Aplicada [Fundamentals of Applied Spatial Econometrics].* Tech. rep. University Library of Munich, Germany.

Kerry, Ruth et al. (2010). "Applying Geostatistical Analysis to Crime Data: Car-Related Thefts in the Baltic States." In: *Geographical analysis* 42.1, pp. 53–77.

Kubrin, Charis E and Ronald Weitzer (2003). "Retaliatory homicide: Concentrated disadvantage and neighborhood culture". In: *Social problems* 50.2, pp. 157–180.

Kuo, Pei-Fen, Dominique Lord, and Troy Duane Walden (2013). "Using geographical information systems to organize police patrol routes effectively by grouping hotspots of crash and crime data". In: *Journal of Transport Geography* 30, pp. 138–148.

Law, Jane, Matthew Quick, and Ping W Chan (2015). "Analyzing Hotspots of Crime Using a B ayesian Spatiotemporal Modeling Approach: A Case Study of Violent Crime in the Greater Toronto Area". In: *Geographical Analysis* 47.1, pp. 1–19.

Leitner, Michael and Henrike Brecht (2007). "Software review: Crime analysis and mapping with GeoDa 0.9. 5-i". In: *Social Science Computer Review* 25.2, pp. 265–271.

LeSage, James (2015). "Spatial econometrics". In: *Handbook of research methods and applications in economic geography*. Edward Elgar Publishing.

Levine, Ned (2006). "Crime mapping and the Crimestat program". In: *Geographical analysis* 38.1, pp. 41–56.

Miles-Doan, Rebecca (1998). "Violence between spouses and intimates: Does neighborhood context matter?" In: *Social Forces* 77.2, pp. 623–645.

Murray, Alan T et al. (2001). "Exploratory spatial data analysis techniques for examining urban crime: Implications for evaluating treatment". In: *British Journal of criminology* 41.2, pp. 309–329.

Ratcliffe, Jerry H and Michael J McCullagh (1999). "Hotbeds of crime and the search for spatial accuracy". In: *Journal of geographical systems* 1.4, pp. 385–398.

Rogerson, Peter A and Peter Kedron (2012). "Optimal weights for focused tests of clustering using the Local Moran statistic". In: *Geographical Analysis* 44.2, pp. 121–133.

Sarrias, Mauricio (2020). "Notes on Spatial Econometrics". In:

Sherman, Lawrence W (1995). "Hot spots of crime and criminal careers of places". In: *Crime and place* 4, pp. 35–52.

Sherman, Lawrence W, Patrick R Gartin, and Michael E Buerger (1989). "Hot spots of predatory crime: Routine activities and the criminology of place". In: *Criminology* 27.1, pp. 27–56.

Telep, Cody W and Julie Hibdon (2019). *Understanding and Responding to Crime and Disorder Hot Spots*. Department of Justice.

Weisburd, David et al. (1992). "Contrasting crime general and crime specific theory: The case of hot spots of crime". In: *Advances in criminological theory* 4.1, pp. 45–69.

Zhang, Xiang et al. (2010). "Detecting and mapping crime hot spots based on improved attribute oriented induce clustering". In: *2010 18th International Conference on Geoinformatics*. IEEE, pp. 1–5.