Testing Determinants of Crime in Philadelphia: A Statistical Spatial Analysis Using Geographic Weighted Poisson Regression

Bridget Burgos

Abstract

This paper examines the geography of property crime in Philadelphia, Pennsylvania. Determinants of crime were assembled based on modern Broken Windows Theory literature that emphasizes social disorder, physical disorder, and other ecological factors. Though there have been many studies linking ecological factors to crime in Philadelphia, few have included multiple environmental and social factors to get at a more complete model of crime determinants. Even less has been studied on how determinants can vary in influence across neighborhoods (i.e., whether relationships are nonstationary). This research tested an annual average of non-violent crime counts from the years 2015-2019 aggregated at the census block group level against social and ecological determinants of crime. General linear modeling (GLM) assuming a Poisson distribution will provide a well-specified global, but aspatial, baseline model. Using a Geographically Weighted Poisson Regression model (GWPR) will capture these relationships while considering spatial heterogeneity. Based on Akaike’s information criterion (AIC) and reduced spatial autocorrelation among residuals, results show that the GWPR model is a better fit than the GLM. The model indicates ecological determinants have a strong influence in areas of north-eastern, central, and southern Philadelphia.

1. Introduction and Background

In 2015 Weisburd published his article The Law of Crime Concentration and the Criminology of Place. In it he emphasizes the need for studying places at the microgeographic scale to better understand causal theories of crime. Since then, there have been several studies that have revealed relationships between crime and the places they happen using more modern spatial analysis techniques (Stein & Sandoval, 2019; Konkel et al., 2019; Wo, 2019). This research builds on the findings of previous studies by incorporating modern variables into traditional concepts. Drawing upon determinants of crime from multiple theories and incorporating it into a modern model that accounts for spatial dependence helps gain insight into where crime happens and why.

In 1982 George Kelling and James Wilson explained the broken windows theory in an article published in The Atlantic. The theory is that community disorder leads to fear of crime, resulting in community withdrawal which then results in a high crime rate (Kelling & Bratton 1998). They called it broken windows theory due to the following analogy: “if a window in a building is broken and is left unrepaired, all the rest of the windows will soon be broken” (Wilson & Kelling, 1982). The first window is a symptom of disorder, and the subsequent broken windows perpetuate said disorder. Many cities have adopted this theory formally and informally into their policing methods, with New York as a notable example. Kelling and NYPD commissioner William Bratton recognized the broken windows theory as a “guiding idea” for police reform in the 1990’s (Kelling & Bratton, 1998). One issue that arose from this evolution of policing was that it led to profiling specifically towards communities of racial minorities (Jefferson, 2016). Due to the problematic policing methods that were borne from the broken windows theory, it has been subject to much research. Some studies measure disorder through non-serious crime offenders to demonstrate how their presence reflects neighborhood violence. (Jia, 2018; Vilalta, 2019). However, this is just one aspect of the theory. Broken Windows theory often goes hand in hand with other crime theories, which can make it difficult to measure specific aspects of it.
What crime studies often do is take the idea of broken windows and break down its into smaller components that draw from older or more contemporary theories. Examples include the collective efficacy theory, social disorganization theory, and physical disorder (Sampson & Morenoff, 1997; Shaw & McKay 1942). Social disorder is often characterized by the economic disadvantages that occur within the community such as poverty, and lack of education (Sampson et al., 1997). Collective efficacy theory describes the degree of social cohesion within communities. This can be represented through number of renters in a neighborhood or overall community attitudes towards their neighborhood (Raleigh & Galster, 2015). There is evidence to believe that socially disadvantaged neighborhoods have high rates of violence, whereas communities with indicators of higher collective efficacy have more property crimes (Schreck et al., 2009).

Physical disorder manifests through symptoms such as graffiti and urban decay (Fagan and Davies, 2012). In a contemporary study by Konkel et al., (2019), they define physical disorder or “Public Space Disorder” through litter, graffiti, and abandoned cars. This specific measure of disorder was found to predict neighborhood non-violent crime rates. Though many studies have looked at specific indicators of disorder, not many have incorporated other known crime predictors. A study by Rybarczyk et al., (2015) tested a wide spectrum of ecological attributes against crime, however he tested all crime types as one aggregate. It is important in crime determinant studies to differentiate crime type since different crimes have proven to have stronger associations with one type of disorder over another (Schreck et. al, 2009). One interesting finding of the Rybarczyk et al. study was the association of crime with sidewalk density. Though this does not specifically fall under the umbrella of any disorder theory, it may be useful incorporating it as a control for a crime model to test types of disorder. Another control that may heavily influence the distribution of crime is land use. Wo (2019) found a strong association between land used for retail and crime rates.

Not only are crime data spatially dependent but relationships can vary across space (Vilalta, 2013). Therefore, to properly test whether factors can predict where crime is likely to occur, it is necessary to incorporate a spatial model. Geographic Weighted Regression (GWR) offers a way to observe which variables’ influence are most significant within local areas. The works of Cahill & Mulligan (2007) pioneered the use of GWR to prove that ecological determinants of crime can vary over space. From then, the use of GWR has been emulated in other crime studies (Arnio & Baumer 2012; Rybarczyk et al., 2015; Yoo, 2017). However, it is often difficult to interpret the findings of these studies because the natural distribution of crime rates is rarely normal, and GWR is most commonly an extension of the OLS model which implies a normal distribution of the test variable (Feng, 2014; Piza 2012). Geographic Weighted Poisson Regression (GWPR) has not only been used by other crime studies, but it can be used to model crime counts, which could make interpreting the results less ambiguous (Tavares & Costa, 2021; Andreason et al., 2020).

2. Goals and Objectives

The goal of this project is to test the effects of social and physical disorder on number of property crimes while controlling for other potential environmental traits that could influence the spatial distribution of property crimes. A city such as Philadelphia that can vary widely in neighborhood traits across a relatively small spatial scale is a good candidate for space-dependent crime study. Specific questions this project examines are the following: do factors from the broken windows theory correlate with property crime in Philadelphia? Do physical factors play a stronger role than social ones? Can the
GWPR model provide insight as to how these factors vary over space?

3. Methodology

3.1 Data

To test the relationship of crime to potential predictors, property crime counts were used. Property crimes are defined in the state of Pennsylvania as, burglary, arson, theft, motor vehicle theft, forgery, fraud, embezzlement, stolen property, and vandalism (PA Mental Health and Justice Center of Excellence, n.d.). These crime types are available as points in shapefiles through the city of Philadelphia open data portal and were aggregated based on census block group polygons. Crime shapefiles from the years 2015-2019 will be downloaded and an annual average count per block group will serve as the test variable. Since Philadelphia exhibits intense urban density over a small geographic space, I choose census block groups as the unit of spatial analysis. This is the smallest unit that contains the most recent social data associated with it. The smaller the unit of analysis, the better since this allows variation in crime clusters to be examined (Weisburd, 2015), and minimizes the modifiable areal unit problem (MAUP) (Rybarczyk et al., 2015). The shapefile containing polygons of the census block groups in Philadelphia were downloaded from the US Census data explorer.

3.2 Independent Variables

Independent variables that contain population data will be derived from the Census Bureau’s American Community Survey (ACS) 5-year estimates from 2015 to 2019. Tables that contained data for percent of individuals living below the poverty line, percent of households receiving public assistance, percent of female-headed households, percent unemployed, and percent less than age 18 were downloaded. These attributes for each block group were used to calculate the concentrated disadvantage index developed by Sampson in 1997. The concentrated disadvantage index provides a robust measure for public health and has been accepted as being a useful indicator for crime (Konket et al., 2019). Total Collective efficacy was also incorporated into the model and characterized through residential mobility. Therefore, the model also includes the 2015-2019 ACS 5-year estimates for total number of populations that was in different house within the US 1 yr ago.

Traditional measures of physical disorder include but are not limited to the presence of trash (Stein, 2015). A unique dataset used in this research that represents physical disorder is the Philadelphia Litter Cabinet’s Litter Index from 2019, which is available as a shapefile from the city’s data portal. This provides a street-by-street rating of litter present on the streets that was aggregated into averages by block group.

To control for other spatial characteristics that could influence where crime is likely to take place, other environmental factors were considered. Though there are a multitude of factors that could be associated with crime, this paper examines the EPA’s Walkability Index available on the EPA’s website as a shapefile for group blocks. This is because crime has been found to be associated with areas of high physical mobility, such as bus routes and sidewalks (Gallison & Andresen, 2016; Rybarczyk et al., 2015). Another factor examined is percent of land used for commercial purposes, which was derived by the most recent land use data (2016) the city had available on the data portal. Research by Wo (2019) and Vilalta (2019) shows land in proximity to retail areas have higher crime rates. Lastly, population density was also included to account for population size and distribution.
3.3 Analysis

By drawing different variables from all these theories, a comprehensive crime model that will be able to explain a large variation in crime counts was created. This section will explain how these variables were modeled using Poisson regression, and GWPR. All tables containing population data were downloaded as CSV files and Excel was used to join separate tables and calculate concentrated disadvantage index as a percentile score per block group. This data was then aggregated by joining features to the block group polygons using ArcGIS Pro. Since this model was using population data as a predictor, only block groups with a population over zero were considered.

To begin exploring the relationships, the global model was run using the linear regression tool within ArcGIS Pro. This type of regression calculates a predicted value from each data point and compares it to the actual value at the point. Since over-dispersion was evident in the crime count data, one outlier was removed, making the final block group sample size 1326. The results that were most important from this test are the pseudo R-squared value or percent deviance explained, the coefficients, and the residuals. The residuals were tested for spatial autocorrelation using Moran’s I test.

The weighting equation at each data point, then estimate coefficients are determined based on a weighing matrix. The weighing matrix is determined by the kernel function, distance, and bandwidth. The kernel function determines bandwidth by calibrating surrounding data points at each location. The distance decay function from each neighbor can be calculated using either a binary or continuous scheme (Vilalta, 2013). The GWR process only uses one bandwidth, which it calculates based on the lowest AIC value it produces from trying different bandwidths. The result is that estimated parameters are influenced by their neighbors. The lower the AIC value, the better the model is (Yoo, 2017). However, Guo et al. (2008) found that basing the model off lowest AIC value may not produce the best results. GWR often chooses the model that produces the lowest AIC value which typically corresponds to the lowest bandwidth. Guo et al.’s (2008) research stresses that having a smaller bandwidth for the GWR can cause issues such as over-fitting, and high levels of coefficient variability. Moreover, small bandwidth values can potentially cause extremely high multicollinearity among local regressions, which could invalidate the results of the model (Fotheringham & Oshan, 2016). Given this, the large sample size and spatial clustering patterns of variables, a predetermined bandwidth of 100 was used for the GWPR. The GWPR was run using the GWR4 program developed by Nakaya in 2009 to retrieve parameter estimates and t-values from each covariate at each sample location.

4. Results

Initial exploration of crime counts began with running the Cluster and Outlier Analysis (using a Euclidean distance of 8 nearest neighbors) tool on the crime dataset. Fig. 1 reveals that block groups with high crime counts and low crime counts exhibit some spatial clustering. Distribution of significant positive spatially autocorrelated clusters of crime were around Center City, which is synonymous with the commercial downtown area of the city. Significantly negative spatially autocorrelated clusters of crime are in the more suburban areas of Northwest and Northeast areas of the city.
The results of the general linear model at the global level indicated that all relationships with crime determinants were positive, except population density. No variables were subject to multicollinearity based on VIF levels (<10). All variables were also significant at the 1 percent level (P<0.01). Variables that appear to have the most influence on crime are commercial land use and litter scores (table 1). After running the GWPR model and comparing the goodness-of-fit indicators, it was revealed that not only is spatial autocorrelation significantly reduced but also the AIC is lowered (table 2). This means that the GWPR model is a better choice for the data.

**Table 1. Summary of variables from Poisson regression output**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>P-Value</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking Index Score</td>
<td>0.139</td>
<td>&lt;0.01</td>
<td>1.37</td>
</tr>
<tr>
<td>Litter Score</td>
<td>0.125</td>
<td>&lt;0.01</td>
<td>1.09</td>
</tr>
<tr>
<td>Concentrated Disadvantage Percentile Score</td>
<td>0.061</td>
<td>&lt;0.01</td>
<td>1.10</td>
</tr>
<tr>
<td>% Of population in a different home 1 yr ago</td>
<td>0.148</td>
<td>&lt;0.01</td>
<td>1.04</td>
</tr>
<tr>
<td>% Commercial land use</td>
<td>0.266</td>
<td>&lt;0.01</td>
<td>1.36</td>
</tr>
<tr>
<td>Population Density</td>
<td>-0.115</td>
<td>&lt;0.01</td>
<td>1.04</td>
</tr>
</tbody>
</table>

n= 1326

Percent of deviance explained: 54
Overall, the percent of deviance explained went from 54 percent in the global model to 75 percent after introduction of the GWPR model. The strongest correlates were commercial land use, walkability, and litter score. Interestingly, population density had a negative association in both models. This could be due to incorrect model fitting of over-dispersed crime data (Chen et al., 2020) or, that population density does not necessarily correspond to where crimes are being committed, given that they take place on properties and not on the people that make up the population. Figure 2 shows that the GWPR model has the strongest fit in Center City, portions of Northeast Philadelphia, and a small portion of West and South side.

Table 2. Goodness-of-fit tests to compare global and spatial model

<table>
<thead>
<tr>
<th>Measure</th>
<th>Global</th>
<th>GWPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>14220</td>
<td>8160</td>
</tr>
<tr>
<td>Moran’s I</td>
<td>0.111 (p=0.000)</td>
<td>-0.009 (p=0.522)</td>
</tr>
<tr>
<td>Percent Deviance Explained</td>
<td>54</td>
<td>75</td>
</tr>
</tbody>
</table>

Figure 2. Local percent deviance explained. Overall % deviance explained: 75
Figure 3. Significant parameter estimates among local coefficients for crime determinants
4.1 Analysis of each covariate

Non-stationarity is evident among all covariates based on differing spatial patterns of their local coefficients. Figure 3 depicts where all significant parameter estimates for each covariate were mapped based on where pseudo t-values that had an absolute value more than 1.96. All covariates experienced significant negative and positive estimates. Positive values are where the relationship between the covariate and crime are stronger, whereas in the negative areas, the effects of the covariate are not as strong. However, again this pattern could also be the effect of incorrect model fitting.

Walking index scores, and percent of land used for commercial purposes are strongest contributors to the model. Each having widely significant, positive parameter estimates throughout the city. This is in line with the works of Vilalta (2019), and Wo (2019). Walking index and commercial areas greatly control the opportunity of property crimes because transportation to target is easier, and commercial areas are an easier target for theft, which is the third most common reported incident (Philadelphia Police Dept., 2015-2019). Patterns shown in parameter estimate maps are expected for the most part. Unexpected walking index relationship patterns in Center City may be due to an inappropriate model choice, bandwidth, and omission of the outlier.

Litter index scores have a general positive relationship with crime. This is consistent with Skogan’s (1990) idea that physical disorder is associated with property crime and theft. There is one area in downtown that is not consistent with patterns immediately surrounding it. This portion of block groups with non-significant litter index estimates could be because this area is immediately adjacent to an area of higher crime counts, however contains relatively low litter block scores. The low litter scores of this area (about 1.4, compared to the city average of 1.8) could be because this area of the city has one of the highest concentrations of large compacting public waste bins (Philadelphia Streets Dept., 2015). Litter distribution in Center City may not be consistent with the rest of city, which could contribute to the pattern of the coefficient map.

Overall, concentrated social disadvantage and percentage of population who lived in a different home 1 year ago exhibit a positive effect on crime counts. This is consistent with traditional social disorder theory. However, social disadvantage does not have a strong effect in Center City. A couple of reasons for this pattern may be that this area of Philadelphia is not as socially disadvantaged and has the greatest ease of transportation. Thus, social disadvantage would not have as much of an influence in this area. Though the traditional works of Shaw & McKay (1942) outlined how lack of social cohesion and mobile populations were tied to higher social disadvantage, the results of the GWPR do not necessarily tell the same story. Fig 4 depicts a graph of predicted crime counts at varying levels of concentrated disadvantage and walking index (values within 1 standard deviation of the mean). Trendlines of the predicted crime counts reveal that crimes in areas that are highly walkable (+1SD) are less influenced by social disadvantage. Whereas in areas of low walkability (-1SD) crime is more driven by social disadvantage. This parallels the works of Wo (2019) which found that retail is found to be a disruptive force in neighborhoods that have social advantage. In this case, walkability seems to be a more disruptive force. Fig. 4 also shows predicted crime counts at varying levels of concentrated disadvantage and mobile populations. The trendline of crime counts at high levels of population mobility (+1SD) compared to the trendline of crime counts at low levels of population mobility (-1SD) indicate that crime is more likely to be driven by social disadvantage where the population is less mobile. This contradicts the traditional literature; the effect of high population mobility in conjunction with social disadvantage should exhibit a direct positive relationship. As opposed to a stable number of crimes counts regardless of concentrated disadvantage. A 2013 study by Jiang et al. in China found that poverty had no effect on...
social cohesion and stability, it appears those results are emulated here. One reason for this may be that social disorder is often used to predict where violent crime occurs (Sampson & Morenoff, 1997; Wang et al., 2019). Property crime on the other hand is more complex, due to the fact it often does not take place where people live. The areas of strong relationship with mobile population and crime are also areas where there are potentially high numbers of student populations given the proximity to colleges and universities. Thus, areas that exhibit this relationship may have more to do with the localization of the victim pool rather than perpetrators and victims.

5. Discussion

This is one of the first crime correlate studies to use an objective database containing trash ratings to quantify physical disorder. Based on this analysis the presence of physical disorder does correlate to the number of neighborhood property crimes. The findings support prior works of Konkel et
al. (2019), Chen & Rafail (2021), and O’ Shea (2006). The main difference however is how these studies chose to quantify physical disorder. Chen & Rafail used different types of 311 service calls, Konkel et al. used an index based on study-collected data of trash and housing conditions. Though the findings are similar, the results may differ if litter index scores were used in conjunction with other measures of physical disorder such as housing conditions or other property crimes such as graffiti or vandalism. In terms of including social disorder into the model that predicts property crimes, the complex interaction of property crimes with populations must be considered. Almost a quarter of Philadelphia residents live at or below the federal poverty line, whereas 8.2 percent make $150,000 a year or more (Cineas, 2018). Meanwhile, Philadelphia also has one of the largest college student populations in the US (Florida, 2016). Thus the stratification of population makes it complex to quantify social disorder and use it to predict where property and non-violent crimes are likely to occur. The results of the population mobility map and its interaction with concentrated disadvantage suggest that the mobile population may be a suitable victim pool, however not necessarily the perpetrating and or victim population, as suggested in other crime studies.

The limitations in this research mainly arise from the ill-fitting nature of the Poisson model. The chi-square test would be one way to test the goodness of fit for a Poisson model. The overdispersion of crime counts would suggest that a negative binomial regression model is more appropriate for crime count data (Piza, 2012). However, if using this method with geographic weighted regression, it would also be prudent to test for spatial variability of covariates. Since most programs that run GWR do so using a single bandwidth not all relationships may be accurately described. Some may operate on a global scale, whereas others on a local one. Either incorporating a semiparametric model or a multiscale geographic weighted regression would produce more accurate results (Fotheringham et al., 2017; Nakaya et al., 2005). By using a larger bandwidth in this project, some relationships may have been muted or over-stated.

In all, findings conclude that presence of litter can be a useful indicator for where property crime is likely to occur. Though this is not to be confused with causation, there is more to be explored with the presence of trash on the streets, physical disorder, and property crime. Is it merely a coincidence? Broken windows theory has been studied in and out, however the layperson’s decision on where to park a car can be simply decided by a cursory glance at the “neighborhood”. If avoidance of a community due to fear of crime is a self-fulfilling prophecy, as the broken windows describes, would simply aesthetically cleaning up the streets reduce said crime? The works of Kondo et al. (2016) and Wolfe (2012) suggest that increased vegetation and greening of vacant lots reduce violent crime. Kondo et al., however found that motor vehicle thefts increased after lots were greened. From a policy-oriented perspective, simply improving physical urban conditions may not improve property crime. Yet, there is some evidence that property crimes interact with violent crimes (Konkel et al., 2019; Chen & Rafail, 2021), and it is known that neighborhood violent crime is more driven by social disadvantage. Policy focused on bottom-up oriented methods that focus on improved housing such as rent control, or prioritizing community services for everything from public sanitation to education may be a way to systematically reduce all types of crime in the long term.
References:


Florida, R. (2016, September 18). The Reality of America’s College Towns. Bloomberg. https://www.bloomberg.com/tosv2.html?vid=&uuid=1ab70af3-9b13-11ec-84f7-4f6a474e6e74&url=LZStd3MvYXJ0aWNsZXJ0MjAxNi0wOS0wOC9hbWVyayWWhlXMTYmNzZ2VzC1jb21lZWhdLXRv


Piza, Eric L. "Using Poisson and negative binomial regression models to measure the influence of risk on crime incident counts." Rutgers Center on Public Security (2012).


