A Spatial Analysis of Violent Crime Cold Spots: Testing the Capable Guardian Component of the Routine Activity Theory

by

Brian David Jopp

A Thesis Presented to the
Faculty of the USC Graduate School
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
Master of Science
(Geographic Information Science and Technology)

December 2016

Dedication

I dedicate this document to my wife for all of the times she has filled in for me as a parent, as I studied by night and by weekend. Without Marnie's help, there is no way possible that I could have had the time to gain the necessary knowledge and skills to complete a Master's Degree. Thank you.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I want to give a special thanks to the entire committee: Dr. Kemp for many key ideas and ways to improve spatial analysis and organization of content within the thesis, with a focus on quantitative analysis; Dr. Lee for helping me to turn a null hypothesis into a working model; and Dr. Loyola for dealing with many revisions, always maintaining a professional and positive demeanor! For the initial work on my thesis, I owe a lot to Dr. Sedano, who kept me headed in the right direction and provided counsel to advance my thoughts into a thesis. Finally, I want to thank Dr. Oda for cartography and visualization advice and all of the professors at USC as being exceptional in guiding me along my path to earning a master's degree.

List of Abbreviations

DC	District of Columbia
CPED	Crime Prevention Through Environmental Design
GWR	Geographically Weighted Regression
GIST	Geographic Information Science and Technology
LAPD	Los Angeles Police Department
LBACH	An explanatory variable: log of percent of the total number of the population who are twenty-four years of age and over and earned a bachelor's degree or higher
LDET	An explanatory variable: the log of percent of the total number of homes that are detached
LMAROWN	An explanatory variable: the log of the percent of the total number of homes owned by married couples
LV1M	An explanatory variable: the log of percent of total number of homes with values over $1,000,000
MPD	Metropolitan Police Department
MAUP	Modifiable Areal Unit Problem
OCTO	Office of the Chief Technology Officer
OLS	Ordinary Least Squares
PHDCN	Project on Human Development in Chicago Neighborhoods
USC	University of Southern California
VIF	Variance Inflation Factor

Abstract

According to the routine activity theory, violent crime may be deterred by a capable guardian. Cohen and Felson’s routine activity theory asserts three conditions need to be met for a crime to take place: a likely offender; a suitable target; and the absence of a capable guardian (Cohen and Felson 1979). A hot spot analysis of violent crimes for Washington, DC shows a divided city. In northwest DC, the census block groups correlate with low violent crime rates. To understand why northwest DC has low crime rates, a quantitative spatial analysis uses housing characteristics as proxies for capable guardianship to test whether a correlation exists between capable guardianship and the deterrence of violent crime. The rationale behind using housing and homeowner characteristics in a model relies upon fusing capability with perception of success. Accordingly, if the criminal perceives a capable guardian to be present, then the criminal will not commit the crime. Following this logic, neighborhoods displaying capable guardianship through housing characteristics ought to have lower violent crime rates. Using exploratory regression, Ordinary Least Squares, and Geographically Weighted Regression the construction of a guardianship model with significant explanatory variables suggests a relationship between capable guardianship and areas with lower violent crime rates do exist. Furthermore, quantitative spatial analysis suggests a strong relationship between low violent crime rates and obtaining higher levels of education exists.

Chapter 1: Introduction

Homicide and other violent crimes cost the government and its citizens a significant amount of money and cause a severe amount of emotional pain for the victim, and the victim's relatives and friends. Various sources state the cost to be between $2 and $22 million per homicide (Delisi, Kosloski and Sween 2010, McCollister, French and Fang 2010). These figures usually break down as follows: fifty percent applies to the lost quality of life for the victim's family and other loved ones affected by the homicide, for example, costs associated with the mental state of those who will need to come to terms with the death which includes grief and potential depression; twenty-five percent in tangible victim costs such as income and payments because of the homicide; twenty percent in criminal justice costs; and five percent in productivity losses. The National Institute of Health provided a narrower range from $4.1 to $11.4 million per homicide. Using this mean, each homicide costs $7.8 million. In 2015, there were 162 homicides in Washington, DC (Bowser 2016), meaning the cost of homicide in Washington equaled approximately $1.3 billion. Homicide related criminal justice costs by the government in 2015 were $260 million, or roughly fifteen percent of DC's income tax revenues (DeWitt 2015) (this is a relative comparison not meant to suggest homicide costs are paid by income tax dollars alone or at all).

Additionally, more resources ought to be expended toward Geographic Information Systems (GIS) to cross-examine, question, verify, or nullify existing theories and qualitative analyses. Criminal analysts use theories to define criminal behavior, to aid in selection of data, in interpretation of their results, and to construct models to predict locations where a crime will likely occur next. "There is a strong body of evidence to support the theory that crime is predictable (in a statistical sense) – mainly because criminals tend to operate within their comfort zone. That is, they tend to commit the type of crimes that they have committed successfully in the past, generally close to the same time and location" (Perry, McInnis and Price 2013, 2). Perry et al are in agreement with Cohen and Felson's routine activity theory, whereby a crime takes place due to a convergence of the victim and criminal in time and space. If crimes take place in similar locations, then the activity can be studied using GIS.

Homicide and other violent crimes are rarely quantitatively studied due to the complex nature of the crime, such as the dynamic relationships between the perpetrator and the victim along with the mental state of the perpetrator (Nicolaidis, Curry and Ulrich 2003, Bozeman 2014). Although these factors needs to be considered, quantitative spatial analysis, by its nature, can be applied without the need to dig into the mind of the perpetrator. Tobler's first law of geography, "Everything is related to everything else, but near things are more related than distant things" (Tobler 1970) can be applied to the study of violent crimes in Washington, DC. For example, when the violent crime "assault with a dangerous weapon" is clustered in an area, there is a strong likelihood that "homicide" will also be clustered in the same area. Thus, one of the most important risks for the occurrence of homicide is spatial proximity to "assault with a dangerous weapon." More details on these relationships are provided in chapter four.

In 2014, Washington, DC's area measured 68.3 square miles and the population estimate was 659,836, so the population density per square mile equaled 9,661. However, with the federal park land subtracted (twenty-five percent of DC is federal park land) then the proper area to use in the formula would be 51 square miles of land. The more realistic population density per square mile is therefore 12,937: using this figure, Washington's population density ranks fourth in the United States amongst cities with a population above 500,000, behind New York City (28,056), San Francisco (18,187), and Boston (13,586); just above Miami (11,997), Chicago

(11,959), Philadelphia (11,635), and Long Beach, California (9,416). The population densities above include federal land. However, removing the federal land and recalculating the population density would not change the numbers nearly to the extent that it did for DC. Although the whole city is considered for spatial analysis, the concentration is on northwest DC because violent crime cold spots remained in this area throughout the study period. This area is located north and west of Rock Creek Park, as shown in Figure 1 below. A more detailed description of the study area follows in Chapter 3.

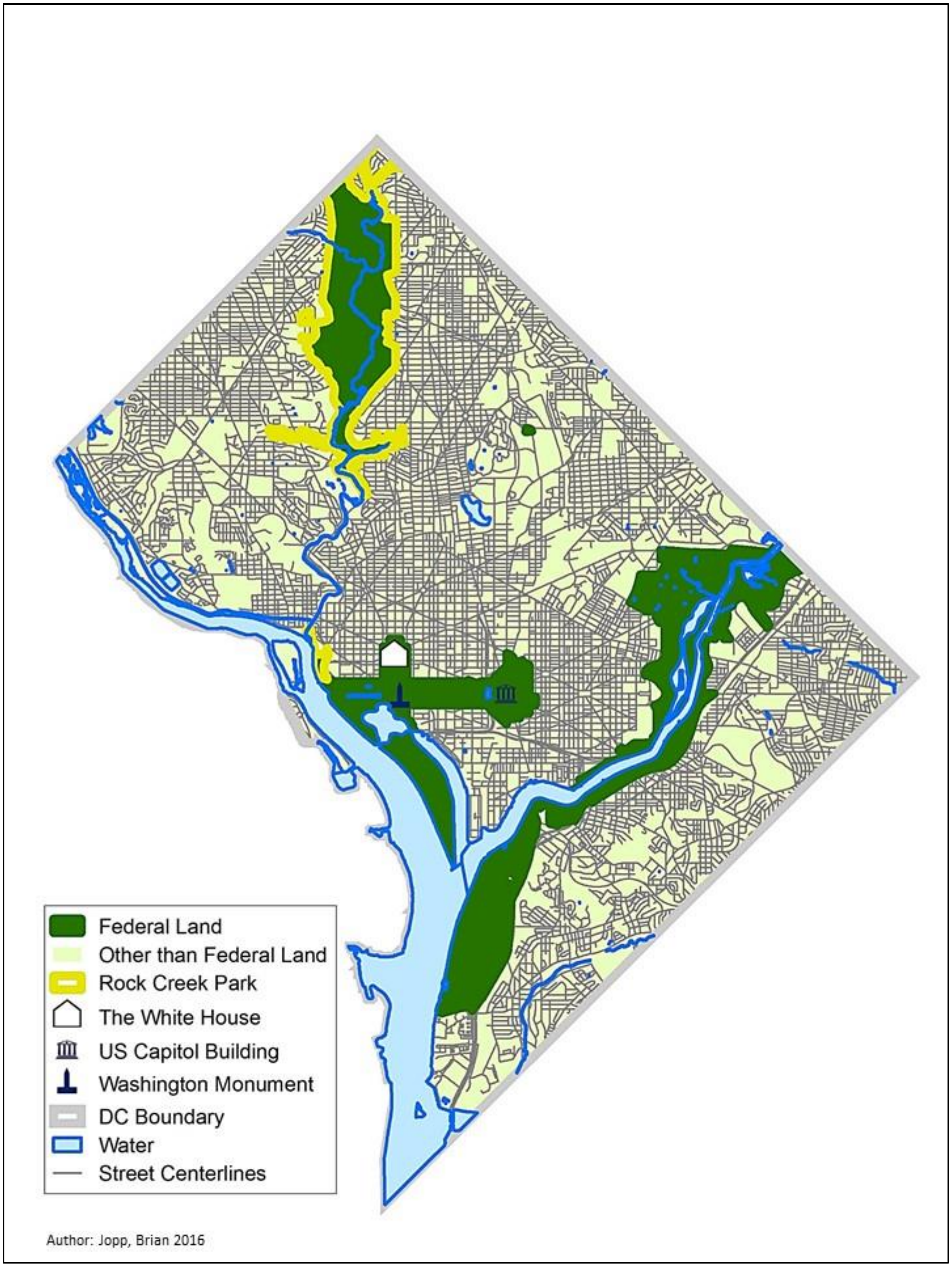

Figure 1: Map of study area: Rock Creek Park and federal land

1.1 The Rationale for Using the Routine Activity Theory

Routine activity theory states that "criminal acts require convergence in space and time of likely offenders, suitable targets and the absence of capable guardians against crime" (Figure 2) (Cohen and Felson 1979, 588). The focus for this thesis centers on the third component of the routine activity theory.

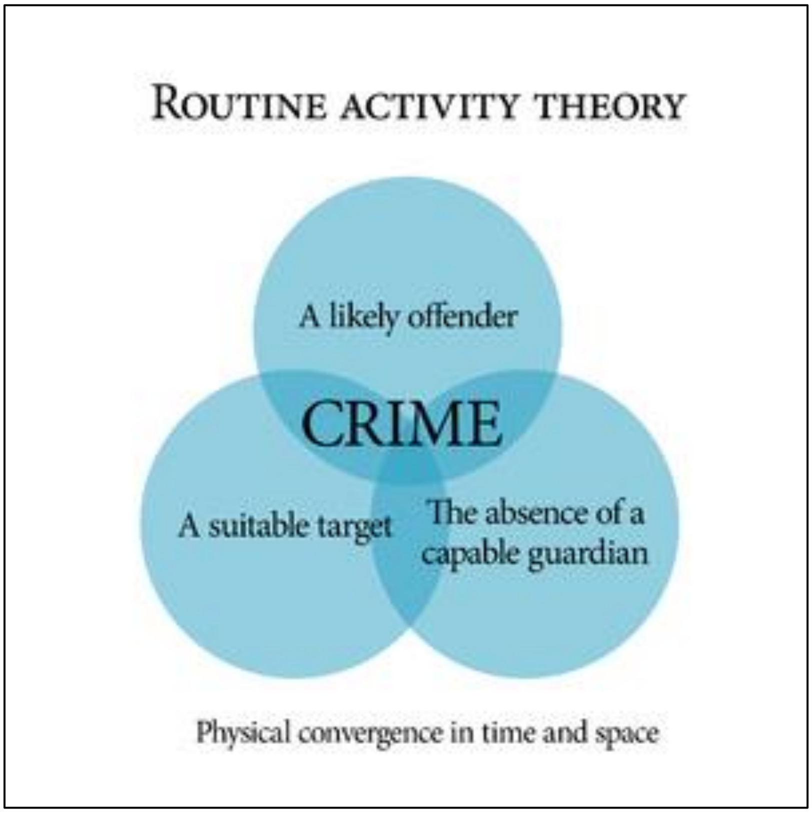

Figure 2: Venn diagram of the routine activity theory. Felson 2006.

However, testing the absence of a capable guardian may prove difficult, so the inverse will be used. In the same article, Cohen and Felson suggest that the routine activity approach will work in the inverse (1979, 589), that a perceived presence of a capable guardian will deter crime. Therefore, the aim of this research will be to test if correlations exist between the perceived presence of a capable guardian and the deterrence of violent crime, which stays within the parameters of the theory. A later article defining the capable guardian reaffirms that a person will choose not to commit a crime if he "feels" somebody is watching him (Hollis, Felson and Welsh 2013, 66 - 67). This paramount implication suggests the physical setting itself may deter crime. If so, then a physical environment with a set of features aligned with a capable guardian ought to show different crime rates than an environment lacking these features.

The authors provide a specific definition: "A guardian is any person and every person on the scene of a potential crime that may notice and intervene (whether they intend to or not)" (Hollis, Felson and Welsh 2013, 73). Working through the above definition, the point-of-view for the criminal hinges on potential assumptions of both presence and capability of the guardian. The criminal makes a decision based on the concept of being observed or not observed by a capable guardian. The guardian does not necessarily need to be present, only the assumption of being present need exist for the deterrence of crime.

Although Hollis et al did not go beyond defining capability "as the presence of a human element of intervention" (2013, 73-74), a well-accepted reality is that humans, even criminals make judgments. These judgments would be based on perceptions of the environment. If the criminal thinks he can commit a crime and get away, then there is a much stronger likelihood that he will commit the crime in comparison to the opposite. Criminals who think they will be caught, due to the environment, will not commit the crime.

Under this theory, criminals make intentional decisions based on assessment of risk, and these decisions take the environment into consideration. For example, detached homes provide physical buffers. The criminal would need to move over this buffered private space, allowing more time for a capable guardian to notice. Detached homes ought to provide a stronger element of guardianship, or at least the perception of the presence of a guardian. More physical space attaches with it a higher risk of being observed in a private space.

Given the above reasoning, housing characteristics may represent the presence of a capable guardian. Housing characteristics, to some degree, define the occupier. In other words, the value of the home, whether or not the structure is occupied, and whether or not it is an apartment or detached home, to some extent show an element of capability. If the criminal even perceives the presence of a capable guardian due to the physical attributes of a neighborhood, then the criminal will not commit the crime, according to the routine activity theory.

1.2 Research Question

According to Cohen and Felson's routine activity theory, capable guardianship may deter crime. If a criminal perceives the presence of a capable guardian, then the criminal will decide not to commit the crime because of the increased risk of being caught. If this this true, then areas where housing characteristics suggest capable guardianship should correlate to low violent crime areas or cold spots. The following spatial analysis uses housing characteristics as proxies to represent capable guardianship to test whether a negative correlation exists between housing characteristics and violent crimes in DC: Using quantitative spatial analysis, are there negative correlations between housing characteristics and low crime areas in the Washington, District of Columbia area?

1.3 Thesis Structure

The thesis consists of five chapters. Chapter One introduced the routine activity theory, the thesis question, and the study area. Chapter Two will take a closer look into the application of criminal theory historically and narrow the focus to the routine activity theory. Furthermore, a literature review will be given to emphasize the need for more quantitative spatial analysis of violent crime and the study of negative correlations between capable guardianship and violent crime. Chapter 3 will describe the data sets, a hot spot analysis, the method of linear regression analysis, and an explanation for each step taken to arrive at the results from Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR). Chapter Four will show the results of the analyses and provide a suggested model derived from exploratory regression. Also, the results for the model will be shown using Esri's ArcMap OLS and GWR tools. Chapter Five will provide potential explanations for the spatial relationships between housing characteristics and violent crime cold spots. Furthermore, a discussion concerning future quantitative spatial analysis when studying violent crimes will be provided.

Chapter 2: Theory and Literature Review

This chapter's two primary concerns are theory and the review of literature. A closer look at five theories for understanding violent crimes will be provided: 1) broken windows theory; 2) conflict theory; 3) social disorganization theory; 4) collective efficacy theory; 5) and the routine activity theory. Additional focus will be given to the routine activity theory, specifically capable guardianship and housing characteristics. Then, a review of literature will be provided to emphasize that quantitative spatial analysis and capable guardianship may benefit researchers when trying to understand violent crime.

Different crime types require different theories to understand the motivation of the criminal. For example, "stranger crimes," those committed by an unknown offender to the victim, such as burglaries and thefts, involve a motivation based on money, and seldom occur more than one time between the offender and the victim (Perry, McInnis and Price 2013). However, in more than half of all homicides, the victim and the murderer know each other (US Department of Justice–Federal Bureau of Investigation 2012). These relationships provide a basis for the victim and criminal to converge in time and space. On the other hand, the victim may cross into a location where criminals lurk. If so, then a spatial analysis may identify these locations and would be a useful tool for researchers when attempting to understand violent crimes.

Police use GIS more and more as software becomes available and as people understand how to use it (Perry, McInnis and Price 2013). However, for the most part, geospatial analysts who work for police departments mostly focus on non-confrontational crimes rather than violent crimes. Non-confrontational crimes consist of crimes such as theft, grand theft auto, and burglary.

Homicide may be different than other crimes in other ways as well. In most homicides, research suggests there is an escalation of an existing crime and most of the murderers show a pattern of committing other violent crimes (Bozeman 2014). In a study where twenty-seven murderers were interviewed and qualitative analysis was applied, 100% of the murders were the result of an escalation of violence, sixty-six percent confessed to either previously committing robbery or the homicide escalated from a robbery, and forty-four percent had either previously committed a form of assault or the homicide escalated from an assault (Bozeman 2014). These statistics provide valuable information for studying violent crime, specifically homicide. Homicide occurs as an escalation within a crime, and usually the offender has committed a violent crime in his past. These salient points heighten the potential of spatial analysis between violent crimes. If assault with a dangerous weapon precedes homicide, then areas with high incidences of assault with a dangerous weapon may likely be high in homicide rates as well. For this reason, studying all four violent crimes may lead to strong correlations between these crimes to be used in future analysis.

Within *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations* (2013), the routine activity theory, rational choice theory, and crime pattern theory were combined to construct a blended theory. The author admits, "This blended theory best fits stranger offenses [non-confrontational crimes] such as robberies, burglaries and thefts…and does not fit well [when applied to violent crimes] due to the break from criminal rational choice framework" (Perry, McInnis and Price 2013). Although the blended theory may not be properly applied to homicide, routine activity theory does seem to apply to homicide. However, before moving into the examination of the routine activity theory, a brief analysis of four other theories will be examined: broken windows theory, conflict theory, social disorganization theory, and social efficacy theory.

Although portions of each of these four theories may seem valid when applied to homicide, other portions of these theories do not apply. Below, specific details are provided to show all of the theories apply to some extent, and a rationale is provided for choosing the routine activity theory as the most relevant theory within which to study violent crime.

2.1 Broken Windows Theory

Broken windows theory gained ground in the early part of the twenty-first century. Both physical and social disorder in neighborhoods indirectly result in higher crime rates (St. Jean 2007). To counteract the negative results, intervention needs to happen early. Although intervention to prevent crime seems intuitive, the theory implies the movement of people in and out of an area at a high rate precludes this. There is also a strong emphasis placed on the physical condition of the buildings such as condemned buildings as a basis for crime, thus broken windows theory.

> "Society cannot ignore these problems, or there looms a strong likelihood that the conditions will get worse. If left unchecked, neighborhood disorder will continue to increase, petty crimes will increase, and residents will perceive that more serious crimes are also on the increase. Fearful of crime, law-abiding citizens will then refrain from using public spaces, become less attached to the neighborhood, and eventually move out of the area only to be replaced by less attached people. Serious crimes will then follow" (P. K. St. Jean 2007, 2).

Although there are elements within this theory that parallel results from qualitative studies such as Bozeman's outcome of 'escalation of violence' and 'low income areas' as factors that contribute to homicide (Bozeman 2014), by itself, was insufficient to explain high levels of drug dealing, robbery, and battery on neighborhood street blocks" (P. K. St. Jean 2007, 195).

The theory concentrates on physical buildings being vacant, abandoned, and the overall environment dissipating to signs of disorder, but these probably work like symptoms of a disease, and do not necessarily motivate criminals. In other words, fixing up buildings, cleaning parks, demolishing abandoned buildings, or any other activity to literally change the physical environment from being unkempt to tidy may only mask the criminal activity. This may enable criminals in these areas because an unaware person may let his or her guard down. Most likely, the changes named above do not actually influence the criminal's reasoning for committing a crime. Broken windows theory does not address the motivations or the reasons for the disorder in the neighborhoods. To find the motivations of criminal behavior, researchers will need to look elsewhere (Sampson and Raudenbush 2004).

Broken windows theory uses poverty, large movements of people in and out of communities, and deterioration of buildings indicators of social deterioration within a community and as a basis for crime. A possible indicator for this change would be high percentage of foreclosures with an area of crime. However, in a study released by Indiana University, where 142 metropolitan areas were analyzed using weighted regression analysis, the results led to conclusions that higher levels of housing-mortgage stress did not result in higher levels of violent crimes (Jones and Pridemore 2012). The research used a multilevel model with individual, familial, and neighborhood levels. Amongst the dependent variables, at least two were violent crimes: assault and robbery. The main explanatory variables were negative equity, loan-to-value ratio, and cost-to-income. Foreclosure does not support broken windows theory as a fully applicable.

Some parallels to broken windows theory to this thesis were captured. Although the theory will not be employed, some of the underlying assertions concerning physical environment, whereby criminals may feel more comfortable committing a crime are in line with the rationale for this thesis. However, this finding is slight. Overall, broken windows would not provide enough substantial input to employ as a theoretical basis for this thesis.

2.2 Conflict Theory

Conflict theory (Marx and Engels 1848) asserts crime is caused by political leverages used by the upper socio-economic class against the lower socio-economic class to promote their interests. Under conflict theory, the powerful construct policy for the upper socio-economic classes to maintain their power, thus, supplying the cause of conflict. The effect from their leverage creates a class conflict, whereby some of the people in the lower socio-economic classes commit crimes. Research efforts to validate the conflict approach, however, have not produced significant findings (Siegel 2000).

Essentially, the central focus of the conflict theory is the conflict between the wealthy and the poor. Although income may both inhibit and encourage homicide and violent crime, low income by itself does not lead to homicide or violent crime. Even though Pratt and Lowenkamp support that an inverse relationship between poor economic conditions and crime exist, they state, “conflict theorists often specify an inverse relationship between economic conditions and crime. Empirical support for this contention in time-series analyses, however, has been inconsistently revealed in the literature, where positive, inverse, and null results have all been found” (2002, 61). The inverse relationship between economic downturns and violent crime is not consistent throughout the decades. For example, in the1960s and 1970s the economy grew along with the escalation of crime rates, but the economic boom in the 1990s showed a drop in crime rates, and these crime rates reached all-time lows in the early 2000s (Scheider, Spence and

Mansourian 2012, 3). The inverse relationship between economic conditions and crime is sporadic when the researcher takes into consideration decades of work.

If wealth and poverty were the main driving mechanisms of crime, then one would expect to see a strong relationship between the economy and unemployment with violent crime. However, in 2015, homicide spiked by fifty-six percent in Washington, DC, yet there were no major economic downturns in the economy. In fact, the unemployment rate went down and the average wages went up (United States Department of Labor 2016). Most likely, an economic attribute could be applied in all theories concerning crime, but the idea of economic change as directly increasing the number of homicides was not the case in the DC area. For the most part, crime analysts include socio-economic variables within their analysis, in one way or another. However, when studying violent crime, economics should not be the only variable.

2.3 Social Disorganization Theory

Social disorganization theory developed in 1942 from mapping juvenile delinquency, whereby Shaw and McKay plotted crime patterns on a land use map. Shaw and McKay gathered the data and recognized certain areas show high numbers of crimes, for example, certain neighborhoods exhibited high crime rates in a continuous spatial pattern. Their findings showed crimes happen where negative social change occurs, for example, neighborhoods with large numbers of transients.

Within the theory, the community loses its moral consensus leading to the deterioration of social control (Anderson 2014). If this theory were employed to explain violent crime, then one would expect to find correlations between high violent crime areas and the following characteristics: poverty, high population density, ethnic diversity, close proximity to industrialized areas, high percentage of immigrants, with a high percentage of transients (Sampson and Groves 1989, Regoeczi and Jarvis 2011).

In 2015, homicide spiked by fifty-four percent in Washington, DC (Bowser 2016). At the same time, net domestic migration was 82% less than the annual average of the prior three years (Bowsner and DeWittt 2015). Given the data, immigration most likely did not play a major factor for the homicide spike nor violent crimes. Therefore, the social disorganization theory may no longer be applicable to study violent crimes in DC. However, the idea that emphasizes place appears to be valid. Hot spot crime areas and spatial analysis is routine for the Los Angeles Police Department to a large extent, which can be seen as easily as going to their website LAPDonline.org (Departmet 2016). Although the element of place in both the social disorganization theory and broken windows theory may apply, other parts of the theory may no longer apply. The routine activity theory incorporates place, so there would be no reason to include social disorganization theory as a theoretical basis.

2.4 Collective Efficacy, an Extension of the Social Disorganization Theory

An off-shoot of the social disorganization theory, the collective efficacy theory may provide a more modern approach and deserves some attention. Collective efficacy occurs in neighborhoods where people are willing to intervene because they are connected to one another through social cohesion (Browning 2002). Neighborhoods with higher levels of collective efficacy will show lower levels of crime (Sampson, Raudenbush and Earls, 1997). Within the above study, the results showed that collective efficacy was correlated with a reduced rate of homicide by 39.7%. However collective efficacy only provided a partial explanation.

The article “Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy” provides statistical legitimacy for the concept of collective efficacy (Sampson, Raudenbush and Earls 1997). However, the data may be biased due to being survey data and asking participants to predict crimes in the future. Each participant answered questions according to his or her opinion about crimes and intervention of crimes that may happen in the future. The data was tallied by the Project on Human Development in Chicago Neighborhoods (PHDCN). The results showed collective efficacy was negatively related to violence. However, the method relied upon qualitative analysis, which leaves open the potential for bias and expressions of uncertainty. “Overall, humans do not do well with predicting probability when using subjective or personal probability, the belief that a certain explanation or estimate is correct; it is comparable to a judgment that a horse has a three-to-one chance of winning a race” (Heuer, Jr 1999, 152). Rather than using human opinions as a way to measure collective efficacy, this study uses quantitative spatial analysis to measure a similar component, where collective efficacy will be replaced by guardianship, for reasons discussed in the next section.

However, in another study, it is suggested that collective efficacy does lower crime rates. Browning assessed that collective efficacy is negatively associated with intimate homicide (including people who are or were in a relationship and consider themselves to be partners, cohabitating, or dating) for homicides from 1994 through 1995 in Chicago (Browning 2002). A potential reason Browning et al findings showed a negative relationship may be due to the study being limited to intimate homicide and not all homicide.

Collective efficacy, measured using qualitative analysis yields mixed results as noted above, as well as others. A potential reason for the mixed results may hinge on the groups chosen for the analysis. When groups who state their purpose is to proactively stop crime are entered into the analysis, then there is a stronger likelihood a correlation will be measured, in comparison to selecting groups with collective efficacy but not necessarily with a direct goal to stop crime (Sampson, Raudenbush and Earls 1997).

Collective efficacy did not measure as being significantly correlated with homicide, according to a study in Chicago in 1995 (Morenoff, Sampson and Raudenbush 2001). Social institutions and neighborhood groups were not significantly correlated with a lower rate of homicide. Instead, Morenoff et al highlight spatial proximity as the major factor to consider when attempting to understand homicide.

A major difference between collective efficacy theory and the capable guardian component relies on the players involved. The collective efficacy theory targets social groups and organizations, and the routine activity theory limits capable guardianship to an individual. In fact, Felson et al make it clear not to use social groups and organizations when applying routine activity theory (Hollis, Felson and Welsh 2013).

2.5 Routine Activity Theory, Spatial Analysis, and the Capable Guardian

The Routine Activity Theory relies on a convergence of a likely offender and a suitable target along with the absence of a capable guardian (Cohen and Felson 1979), as shown in Figure 2. Although no single accepted theory explaining the behavior of homicide exists (Bozeman 2014), a different approach combining the routine activity theory with quantitative spatial analysis may provide a framework of understanding where homicides do not occur. In a review of criminal studies, based on thirty-three articles written between 1995 and 2005, Spano and Freilich limit the third component of the routine activity theory to the absence of a capable guardian (2009). Some of the theories wrongfully equate guardianship to carrying a weapon, using cameras, and social groups (Felson 2006); these types of analysis are not discussed. This study suggests the routine activity theory can also be used to show where violent crimes are less likely to occur. Later in this section, the definition for capable guardian is given.

There are many qualitative analysis studies that could benefit greatly from spatial analysis. For example, in a study published in the Journal of General Internal Medicine, qualitative analysis involving interviews of women who survived an attempted homicide by an intimate partner, revealed twenty-eight of the thirty women "had previously experienced physical violence, controlling behavior, or both from the partner who attempted to kill them" (Nicolaidis, Curry and Ulrich 2003). A spatial analysis of the women's locations of these attempted murders may have shed some light on other factors associated with the environmental conditions where the attempted crime took place. Place could be a vital component to understand the problem set. Although these violent crimes may be complex, crime analysts ought to use all available tools to understand the problem. The routine activity theory emphasizes the importance of time and place, but neither were looked at. This is just one example where the application of the basics within the routine activity theory may be able to enhance a qualitative analysis.

Spatial analysis may shed some light on such a dark topic, even though the answers may be more complex than spatial analysis can provide, merely asking the questions can open up a discussion leading to possible solutions. Did these victims come from a low income area? Did the victims live in areas where violent crimes were more prevalent? What did the victim's neighborhood look like? What are the elements of the physical surroundings? However, spatial analysis on its own may lead to more questions than answers. When studying violent crime, it is fundamental to include theory. By utilizing spatial analysis in a theoretical framework, data, methods, and results can be better understood. Through this understanding, solutions may surface. The routine activity theory seems to be an excellent candidate for examining violent crime.

The main focus in this study is the third component of the routine activity theory, capable guardianship. Being capable may be synonymous with other characteristics such as success, achievement, and individual autonomy. Therefore, explanatory variables such as housing characteristics could be chosen for quantitative spatial analysis.

2.5.1 Definition of a Capable Guardian

"Guardianship can be defined as the presence of a human element which acts – whether intentionally or not – to deter the would-be offender from committing a crime against an available target" (Hollis, Felson and Welsh 2013, 76). Hollis et al reinforce that the guardian must be a human element and not an official such as the police. However, dogs could be guardians, but cameras and other tools only reinforce an already present guardian. This thesis works within the author's intent that the capable guardian is a person. The physical properties inherent with housing characteristics portray information about the owner, which suggests a measurement of capability. Criminals are able to identify these neighborhoods as more likely to contain capable guardians, and this perception may deter crime. This thesis asserts the perception of the capable guardian, regardless of whether or not the guardian was present, correlates with less crime. If so, then this ought to be measurable if the proper explanatory variables are chosen.

2.5.2 Housing and Homeowner Characteristics as a Proxy for the Capable Guardian Component

Housing and homeowner characteristics gained from the United States Census Bureau are used to act as potential proxies representing a capable guardian. Along with housing characteristics, education level attained and financial variables are considered for the capable guardian model. The explanatory variables ought to be defensible by quantitative statistics.

When placed into a model, the variable is expected to show a negative linear relationship to violent crime. In no way is the intention of this model to predict violent crime.

Crime Prevention through Environmental Design (CPTED) studies the spatial characteristics of environments that may enable or deter criminal activity (Cozens and Love 2015). Similar to a camera being an aid to the capable guardian, defensible space may enable people to be more capable. There are four design elements for defensible space that apply to the above housing characteristics being used as variables in a model meant to represent capable guardianship: 1) Perceived areas of clearly defined ownership of space (detached homes); 2) Image and milieu built for the perception of space promoting well maintained and orderly places (high-valued homes); 3) Opportunities for surveillance for residents (occupied); 4) Geographical juxtaposition being the capacity for surrounding spaces to influence adjacent areas (potentially many characteristics) (Cozens and Love 2015, 393-395; Taylor and Harrell 1996; Reynald 2015). Therefore, housing characteristics may be a good variable to examine for a negative linear relationship to violent crime.

2.6 A Review of Literature that Used the Routine Activity Theory

Historically, spatial science researchers used the routine activity theory to study risk associated with time and space, whereby a likely offender crosses paths with a suitable target, and without a guardian present (Cohen and Felson 1979). A major focus of researchers who base their rationale on the routine activity theory, is on the victim’s movements, whereby the parents (guardians) are not present (Vazsonyi, Belliston and Hessing 2002; Lauritsen and Quinet 1995). Many of the studies place the crimes during travel times, for example, traveling to and from school. The other focus is on places where capable guardians are not present (Garofalo, Siegel and Laub 1987). A couple of articless cite juveniles or college students straying from their normal routes as increasing the risk and exposure to likely offenders (Nofziger and Kurtz 2005, Tewskbury and Mustaine 2003). By moving away from the safeguards of a home, neighborhood, or place where capable guardians are present, the students enter areas with greater risk to criminal activity.

Two common themes used by most researchers of the routine activity theory are victimization and deviance. Spano and Freilich (2009) published an accredited review of research articles on the routine activity theory from 1995 to 2005. In the titles of the thirty-three articles cited, “victimization” appears thirty-one times and “deviant” shows up four times. In this assessment, the dependent variable crime/deviance showed a negative linear relationship to capable guardianship in seventeen out of eighteen of these studies (Spano and Freilich 2009). Throughout the research cited, strong evidence supports the absence of a capable guardian to be a significant component linked to crime.

As stated above, the bulk of articles concentrate on travel and being away from a capable guardian. To capture these crime events, the researchers generally gain access to surveys and interviews for analysis. In “Personal Criminal Victimization in the United States: Fixed and Random Effects of Individual and Household Characteristics,” Tseloni attempts to use a multilevel model to “disentangle the unexplained heterogeneity between individuals and between households by linking surveys to explanatory variables concerning personal crimes (assault, purse snatching, rape, sexual assault, and robbery are examples)” (2000, 415). Tseloni used several household variables such as income, number of household members, number of vehicles, education, and marital status. Essentially, the results of negative linear relationships depended on the proximity to the crime areas. As households with similar characteristics became closer to the crime areas, the impact of the household characteristics lessened.

Tseloni asserts that the results for housing characteristics were difficult to interpret without considering the lifestyles of the occupants. Tseloni ends the journal article by asserting more research needs to occur at the census tract level, building on various comments concerning the lack of research using explanatory variables (such as household explanatory variables) to attempt to better understand victimization and crime (2000).

Some studies cite the routine activity theory but they do not use the capable guardian component in accordance with the originator's definition. The results of a cross-sectional study of Bogota, Colombia showed an overlap between victims and perpetrators, whereby one-third of the sample of 3,007, engaged in activity of being both the victim and the perpetrator over the course of a year (Klevens, Duque and Ramirez 2002). Place appears to be a salient feature for crime, as a cross-analysis between victim/perpetrators and victims-only showed both answered questions such as "Avoids going out at night alone," "Stays home at night," and "Avoids dangerous neighborhoods" with a range of less than two percent, according to the interview answers. This study minimized the potential intervention of the capable guardian.

In another article, geo-located 911 calls in Minneapolis, Minnesota showed violent crimes were clustered: All robberies were committed in 2.2% of the area within the city, and all rapes were in 1.2% of the area in Minneapolis, sometimes tied to a specific building or lot (Sherman, Gartin and Buerger 1989). Within the study, strong evidence is provided that confirms these areas lacked a capable guardian. Sherman et al sum up the argument for capable guardianship on page forty-six, "If the distribution of crime hot spots was determined -solely by the concentration of offenders, then how can we explain the complete 1-year absence of predatory crimes from 73% of the places in high-crime crime areas in Minneapolis (compared with the expected absence from only 57%)?"

Although, “Violent Disorder in Ciudad Jaurez: a spatial analysis of homicide,” used social disorganization theory, the study revealed negative linear relationship correlations between explanatory variables and homicide (Vilalta and Muggah 2014). This study was placed in this chapter to highlight homicide can be better understood using regression analysis aggregated by police district (similar to census tract), and that significant negatively linear-related explanatory variables are attainable. Six negative linear explanatory variables for homicide were cited: Female population between six and eleven that do not attend school, Population with employment, Population ascribed to Seguro popular (state funded health insurance), Population over twelve that is married, Number of people in temporary housing, and Occupied home units with land floor. Within the areas where these variables show a negative relationship to homicide, Vilalta and Muggah suggest possible reasons as being wide socioeconomic and socio-behavioral dividends due to family support, social ties such as marriage, supportive welfare programs, and employed populations (Vilalta and Muggah 2014). Regardless of explanation, the point is that explanatory variables with a negative relationship to homicide were found using regression analysis, even in an exceedingly complex environment such as Ciudad Juarez. Even though Ciudad Juarez towers Washington, DC in homicide with 6,436 homicides between 2007 and 2010, and the socio-economic situation and culture of violence is much more complicated, cold spots were identified along with six significant linear regression explanatory variables.

In the aforementioned research, violent crimes are spatially clustered. However, most of the studies use qualitative analysis in the form of interviews, and all of the analysis employed the absence of a capable guardian, even though routine activity theory includes the inverse to be part of the theory: the presence of a capable guardian deters crime. However, one article was found that used the presence of a capable guardian as having a negative linear relationship to violent crime. Below, the main ideas and similar details are provided, along with a justification for using Geographically Weighted Regression (GWR) within this analysis.

In “Using Geographically Weighted Regression (GWR) to Explore Local Crime Patterns,” (2007) Meagan Cahill and Gordon Mulligan used regression analysis to study violent crimes in Portland Oregon. Similar to DC, high crime and low crime spatial clusters were identified in the city. Although the regression analysis used variables with positive and negative linear relationships to violent crimes (homicide, sexual assault, robbery and aggravated assault), two guardianship variables were highlighted that had a negative linear relationship to violent crime: residential stability, and percent of married families. In addition, the aggregation method was census block group and the data used consisted of a five-year study period for the years 1998 – 2002. Given the similarity of the dependent variable, aggregation, inclusion of the routine activity theory, and some of the independent variables directly showing negative linear relationships to violent crimes, the results and conclusions for this study were noted for similar application in this study.

“The application of GWR to a model of violence rates and its comparison to an OLS base model has yielded several striking results” (Cahill and Mulligan 2007, 190). Four of the eight parameters showed non-stationarity, and the measure of affluence produced a counterintuitive result, being positively related to violent crime. Through the use of GWR, Cahill and Mulligan identified 20% of the census block groups as being affluent and positively related to violent crime. Furthermore, single-person households, married families, and population density were not highly correlated to crime. Through the use of GWR, an exploratory regression analysis, more information can be better understood and applied in future analysis. For example, Cahill and

Mulligan suggested using a higher level of income for the affluent variable, moving the benchmark of $50,000 to $75,000, which may be enough to change the counterintuitive results.

Given that Cahill and Mulligan's article holds many similarities to this study: violent crime as the dependent variable, census block aggregation, the inclusion of the routine activity theory, in a similar-sized city being Portland, GWR may be another tool to use in the study of violent crimes in DC. A possible reason for heterogeneity when studying violent crime may be that "While structural characteristics of neighborhoods influence crime, it can also be said that crime influences the structural characteristics of neighborhoods" (Hipp 2010, 205). With regards to other studies suggesting violent crimes may be non-stationary, researchers who employ OLS may want to consider using GWR as well to double check the variables used for heterogeneity.

Chapter 3: Methods

This chapter describes the methodology developed to test the capable guardian component of the routine activity theory. The main objectives are to conduct a spatial analysis of violent crime cold spots and to build a model of housing characteristics as the explanatory variables. Many problems within the study area needed to be resolved prior to finding a model to test whether housing characteristics indicate the likelihood or perceived likelihood of a capable guardian being present, thereby deterring crime. This chapter is broken down into six main sections: 1) Data sets; 2) The study area, federal land, and aggregation choice; 3) Hot spot analysis; 4) The explanatory variables; 5) Exploratory regression and Ordinary Least Squares (OLS); and 6) Geographically Weighted Regression (GWR).

3.1 Data Sets

Eight data sets are used to complete the analysis, as shown in Table 1. The first data set consists of census block group boundaries, formatted as polygons. The second data set consists of four violent crimes, made up of point data. Data sets three through eight consist of twenty-five variables relating to homeowner or housing characteristics in the form of count data.

The dependent variable is made up of four violent crime datasets (homicide, assault with a dangerous weapon, robbery, and sexual assault) from 2012 through 2015. These datasets are provided by the Metropolitan Police Department, District of Columbia. Given the data source consists of police reports, the accuracy of the points is not being challenged. The georeferenced crime points are plotted at either end or in the middle of the block on which the crime occurred, as shown in Figure 3. Being within a city block provides enough spatial accuracy to be analyzed at the spatial scale of data that is aggregated by census block group, which is the aggregation used throughout the study. In addition, all of the points are clearly located within a single census block group, none occurring on the boundaries.

All of the explanatory variable data sets were obtained on March 5, 2016 from the US Census Bureau. The data sets are 2010 – 2014 American Community Survey 5 – Year Estimates, extracted online via the American Fact Finder search tool: http://factfinder.census.gov/. The accuracy for each data set differs greatly for each data set and for each census block group. The details for sample size, data quality measures, data accuracy and statistical testing can be found on the American Community Survey website in the Data and Documentation section.

Table 1: Data sets and variables

Data Set	Type	Variable	Description
Boundary	**Polygon**		**Aggregation**
		N/A	Census Block Group 2010
Violent Crime	**Point**		**Dependent Variables**
		Hom	Homicide (2012 - 2015)
		Dan	Assault w Dan Weapon (2012 - 2015)
		Rob	Robbery (2012 - 2015)
		Sxa	Sexual Assault (2012 - 2015)
US Census	**Survey**		**Explanatory Variables**
Edcuation	**Count**		**Highest Education Level Obtained**
		NoDip	No High School Diploma
		HSDip	High School Diploma
		SColl	Some College
		Bach	Bachelor's Degree or Higher
Home Value	**Count**		**Value of Home**
		V199	Under $199,999K
		V999	$200,000K - $999,000K
		V1M	Over $1M
Income	**Count**		**Household Income**
		I39	Under $39,999K
		I74	$40,000 - $74,999
		I199	$75,000 - $199,999
		I200	Above $200K
Occupancy	**Count**		**Percentage of Occupied Homes**
		Occ	Occupied
		Vac	Vacant
Housing Type	**Count**		**Type of Home in Relation to Other Homes**
		Det	Detached
		A4	Attached 1 - 4 Units
		A9	Attached 5 - 9 Units
		A10	Attached 10 Units or More
Marital and Ownership	**Count**		**Marital and Ownership Status**
		MarOwn	Married and Owns the Home
		MarRen	Married and Rents the Home
		ManOwn	Man Without a Wife, Owns the Home
		ManRen	Man Without a Wife, Rents the Home
		WomOwn	Woman Without a Husband, Owns the Home
		WomRen	Woman Without a Husband, Rents the Home
		NFOwn	No Family, Owns the Home
		NFRen	No Family, Rents the Home

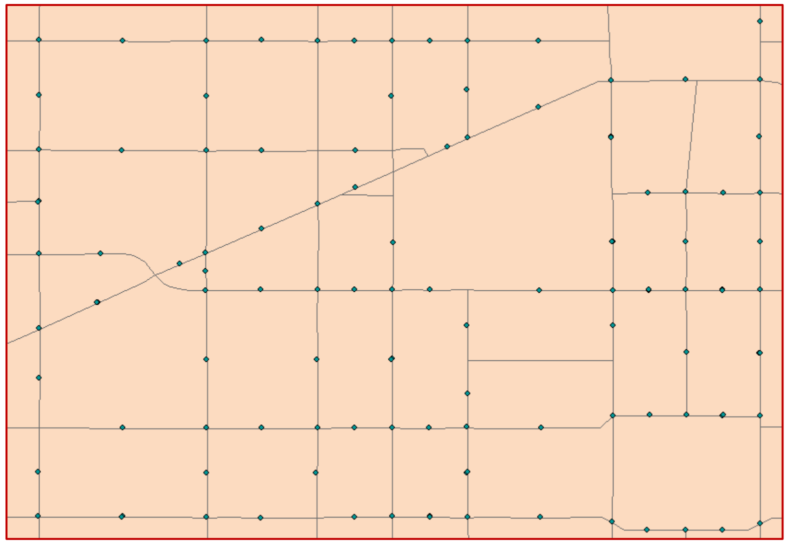

Figure 3: Crime points plotted to one of three positions on the block. Source: Metropolitan Police Department.

Each violent crime is georeferenced as a point. Along with the coordinates, additional information such as date, type of offense, method, and data at different scales is provided, as shown in Table 2 below.

Table 2: Violent crime variable examples

ID	X	Y	Report Date	Offense	Method	Block Site Address
1	-77.03143305	38.91112616	2/3/2013	Robbery	Knife	1330 - 1399 Block of Q Street NW
2	-76.93432072	38.88311158	2/3/2013	Robbery	Gun	Benning Road SE and 46th Street SE
3	-76.98928103	38.90020276	2/3/2013	Robbery	Others	1200 - 1299 Block of H Street NE

ID	X Coord	Y Coord	W	N	D	P	NC	BG	CT
1	397274	138140	2	2F	Third	307	7	005001 2	5001
2	405698.82	135031.7	7	7F	Sixth	608	33	009907 2	9907
3	400930	136927	6	6A	First	104	25	008402 1	8402

W = Ward, N = Neighborhood, D = District, P = Police Service Area,
NC = Neighborhood Cluster, BG = Block Group, CT = Census Tract

In 2013, there were 104 homicides, including twelve people who were murdered in the Navy Yard on September 16, 2013 (Bowser, Muriel 2013). I assessed this to be an outlier event, so these twelve homicides were deleted from the study lowering the number of homicides to ninety-two (92) for 2013, as shown in Table 3. Without deletion, the event would skew the study area as well as the results of a hot spot analysis. No other outlier events are known concerning violent crimes for the years included in this study.

Esri's nearest neighbor tool is used to show that each of the violent crimes for each category and year are clustered at a 99% confidence level, as shown by the Z-Scores in Table 3 below. The lowest Z-Score found is homicide in 2012, Z=-4.5481 (well below the critical value -2.58), meaning there is a less than one percent likelihood that these clustered patterns could be the result of random chance.

Table 3: Violent crime Z-scores for nearest neighbor analysis of each violent crime by year

	2012		2013		2014		2015		2012-15	
Crime	N	Z-Score	n	Z-Score	n	Z-Score	n	Z-Score	N	Z-Score
Homicide	87	-4.5481	92	-6.3119	105	-6.0076	156	-9.8856	440	-17.1721
Assault w Dan Weapon	2,358	-51.7878	2,393	-53.836	2,467	-54.9157	2,385	-52.9673	9,603	-136.099
Robbery	4,209	-69.0641	3,994	-67.9034	3,269	-58.408	3,352	-59.7989	14,824	-166.838
Sexual Assault	258	-8.6571	292	-12.4553	311	-12.061	275	-10.5234	1,136	-30.1684

	N
Total	26,003

Since there are only 440 homicides included in the regression analysis, homicide is given a fraction of measurement (less than two percent weight) when the analysis shifts to regression analysis in comparison to assault with a dangerous weapon (n = 9,603; thirty-seven percent weight), robbery (n = 14,824; fifty-seven percent weight), and sexual assault (n = 1,136; four percent weight). If the violent crimes generate violent crime hot and cold spots along with showing a relationship in regression analysis, these differing sample counts, most likely, will have minimal impact. The overarching purpose of the study is to show the negative correlation to violent crime overall, not the deterrence of a specific violent crime.

All data are converted to the geographic coordinate system NAD 1983 NSRS 2007 State Plane Maryland FIPS 1900 (US Feet), and the projection Lambert Conformal Conic is used throughout the study.

3.2 The Study Area, Federal Land, and Aggregation Choice

As stated in Chapter 1, Washington, DC consists of 68.3 square miles, but twenty-five percent of this land is federal land, under federal jurisdiction. Since this land is not covered by the Metropolitan Police Department of Washington, DC (MPD), the violent crime data for these areas may not be accurate (Perry, McInnis and Price 2013). To ensure these areas do not skew the analysis, as these would show up as cold spots, the largest census block groups are removed from the study area as highlighted in red, displayed in Figure 4. The number of census block groups are reduced from 450 to 446.The census block groups outlined in orange in the top center of Figure 4 are retained because these areas include park land as well as populated areas with proper census data.

After the omission of these areas, the remaining areal units are checked for spatial autocorrelation of violent crimes. Since, Esri's incremental spatial autocorrelation tool did not provide a meaningful distance to apply for conceptualization of spatial relationships, two other options are used to ensure spatial autocorrelation exists for violent crime when aggregated by census block group. Inverse Distance and Inverse Distance Squared are chosen to represent the crime data over Contiguity Edges Corners because the borders for blocks and block groups are not cultural and did not employ a meaningful relationship to violent crime.

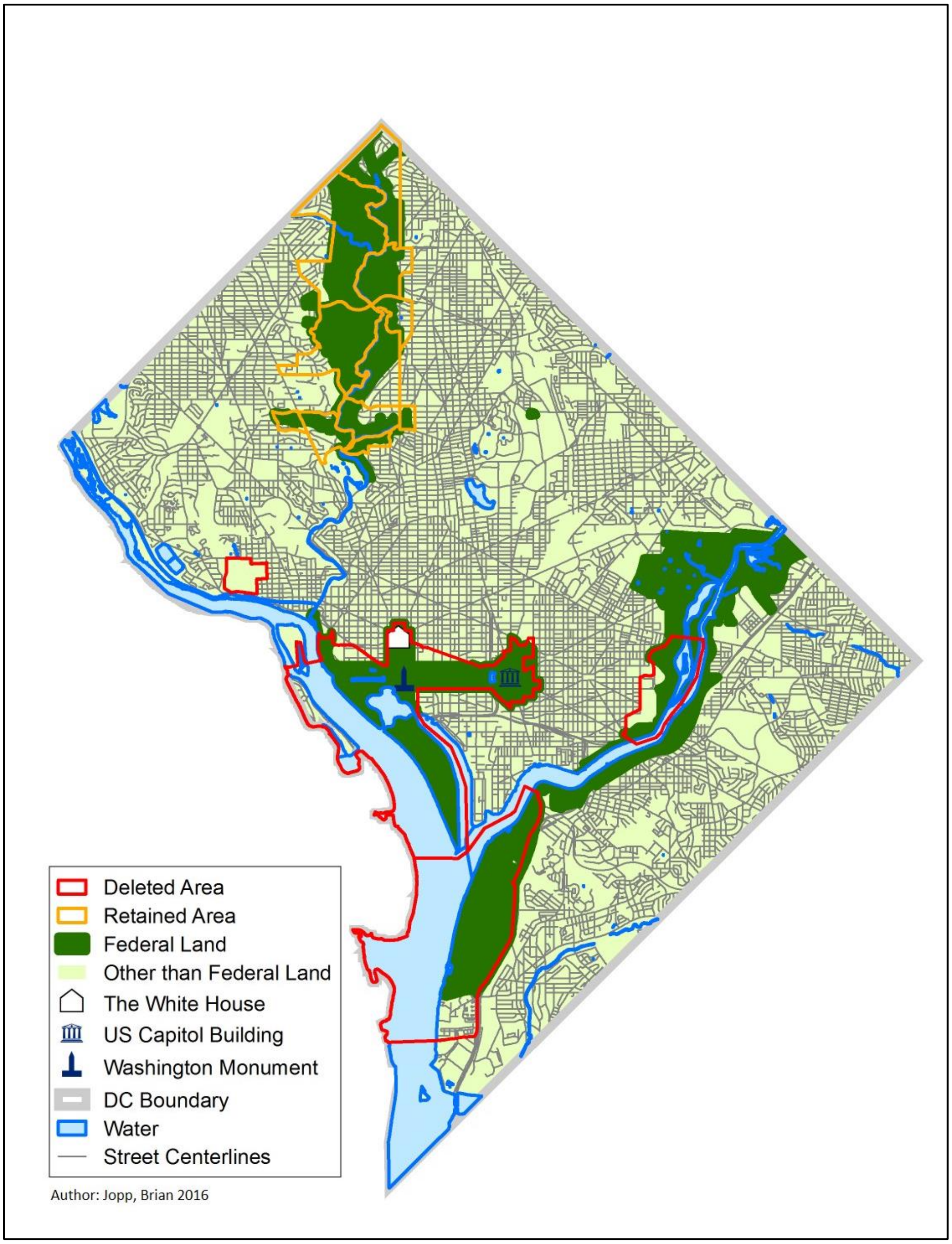

Figure 4: Modified study area

A well-known problem called the Modified Areal Unit Problem (MAUP), results when inappropriate spatial units are used in a spatial analysis (Bolstad 2012). Although MAUP cannot be completely avoided in this study because the polygons are not constructed specifically for the analysis, a finer spatial aggregation may reduce the number of different types of homes and homeowner characteristics. Therefore, I chose census block groups as a better scale than census tracts to minimize the inclusion of multiple types of housing. Although census block groups may still contain differing neighborhoods, the crossover ought to be less in comparison to census tract aggregation.

More than likely, due to the census block groups being divided up without regard to housing, heterogeneity or non-stationarity may cause misspecification to some degree and the model may be missing a key variable (Wilson and Fotheringham 2008). To fix this problem would require new polygons to be drawn according to housing type and then a survey be applied to fit the aggregation, which is beyond the scope of this study. Although census block may refine the analysis even more than census block group, the census bureau does not provide housing data at this level of aggregation.

Esri's spatial autocorrelation tool is used to make sure violent crimes are significantly clustered, as this would need to be the case to apply a spatial study to violent crimes. According to the Z-Scores, violent crimes are clustered in all three forms of aggregation measured: census block, census block group, and census tract, as shown in the Table 4.

Table 4: Spatial autocorrelation and Z-scores when violent crime is aggregated

Inverse Distance				
Aggregation	n	mod n	Z-Score	Index
Block	6,507	6,094	96.0043	0.1426
Block Group	450	446	18.472	0.2363
Census Tract	179	175	8.7657	0.2203

Inverse Distance Squared				
Aggregation	n	mod n	Z-Score	Index
Block	6,507	6,094	8.4172	0.1797
Block Group	450	446	11.4657	0.252
Census Tract	179	175	6.2667	0.2344

Z-Scores at the census block group level are well above the accepted level of 2.58, meaning this aggregation method, when using violent crime data, ought to show significant hot and cold violent crime spots when analyzed via a hot spot analysis. All of the violent crime datasets are normalized using log transformation.

3.3 Hot Spot Analysis

Each violent crime data set (homicide, assault with a dangerous weapon, robbery, and sexual assault) is independently tested for hot and cold spots using Esri's Optimized Hot Spot Analysis tool. The resulting graphics are overlaid as well as combined in a separate analysis to demonstrate that violent crimes occur in the same areas, suggesting a strong relationship between violent crimes exists. Therefore, I assess that all four violent crimes can be combined together to form a single dependent variable providing a larger sample size. This allows for a more defined aggregation method, such as census block group instead of census tract. The results of the analysis benefits with a higher probability, which may lead to more conclusive results concerning finding areas with capable guardians. Areal units with less space may better represent individual neighborhoods, so homeowner and housing characteristics ought to be more similar.

Census block group is closer to the ideal than census tract. This may highlight certain areas and may lead to higher adjusted R-squared values when regression analysis takes place. It is very important to show that the violent crimes are related and occur in the same areas. Hot spot analysis is used to show this relationship exists prior to choosing any potential explanatory variables.

Given that the Optimized Hot Spot Analysis tool automatically applies scale and spatial dependence to its formula to generate a statistical measurement, both saves time and ensures human error is reduced (Esri 2016). Other hot spot analysis tools require the analyst to choose parameters that can change the results drastically.

Since no distance can be justified through the evaluation of violent crime in concerns of scale of analysis, the automatic strategies built into Esri's tool are used. In the first strategy, the tool uses incremental spatial autocorrelation to measure the intensity of clusters using Z-scores to identify a peak to establish a distance (using Global Moran's I statistic). However, when no peak is found, then the distance is determined by computing the average distance that yields K neighbors for each feature, when K is computed as 0.05 * N (N is the number of features in the Input Features layer). The tool automatically adjusts so K is no less than three and no greater than thirty neighbors are used. Outliers are identified and not included in the analysis. Finally, the hot and cold violent crimes for census block group aggregation is displayed at high confidence levels: 90%, 95%, and 99%.

3.4 The Explanatory Variables

The main purpose for using linear regression analysis is to find variables with a negative linear relationship to violent crime, so a model can be built using housing and homeowner characteristics. Given prior research highlighted in Chapter 2, variables chosen are rationalized, as to why these are proxies for the perceived presence of a capable guardian. In other words, random variables without rationale for being related to a capable guardian are not considered. A seventy-five percent negative linear relationship figure is used as the benchmark for advancing to the next phase of testing, for inclusion in the final model, with a goal to show 100% negative linear relationships for all variables included in the final model. This study is mostly concerned with showing a negative linear relationship between the proxy explanatory variables and violent crimes. A benchmark for a minimum R-squared score is not set. However, each explanatory variable needs to measure as being significant in the final model.

All of the explanatory variables are changed to a ratio by computing a percent using the raw number divided by the total number within each dataset. The Census Bureau provides the count as well as the total number used within the sample, so computing a percent can be easily done using Esri’s field calculator. Prior to analysis, all of the dependent variables are normalized using the log or arcsin function. Also, the explanatory variables, when applicable are normalized using the log function when the variable displayed a positive skew and arcsin when the variable displayed a negative skew. The transformations for the explanatory variables can be found in Table 5.

Table 5: Normalizing the explanatory variables

Variable	Skew	Pmean	Pmedian	Diff	Pkurtosis	CompTr	Tmean	Tmedian	Diff	Tkurtosis
NoDip	Pos	0.11235	0.09349	0.01887	3.2492	Log	0.102	0.0892	0.01273	2.679
HSDip	Pos	0.2044	0.17863	0.02577	2.499	Log	0.1733	0.1636	0.00976	2.0103
SColl	Pos	0.18087	0.17057	0.0103	3.4361	Log	0.1603	0.1568	0.00355	2.6014
Bach	Pos	0.58672	0.51128	0.07544	1.6973	Log	0.3912	0.3983	-0.00708	1.65
V199	Pos	0.12423	0.03648	0.08775	11.167	Log	0.1009	0.0358	0.06506	4.8804
V999	Neg	0.71136	0.82342	-0.11206	3.4763	ArcSin	0.8931	0.9674	-0.0743	2.5011
V1M	Pos	0.09174	0	0.09174	9.1998	Log	0.0775	0	0.0775	7.2554
I39	Pos	0.32596	0.28585	0.04011	2.5325	Log	0.2707	0.2514	0.01927	2.2225
I74	Pos	0.21102	0.2028	0.00822	2.915	Log	0.1883	0.1847	0.00368	2.7265
I199	Neg	0.33284	0.34798	-0.01514	2.3496	ArcSin	0.3437	0.3554	-0.01177	2.4156
I200	Pos	0.13018	0.07237	0.05782	4.728	Log	0.1146	0.0699	0.04474	3.7644
Occ	Neg	0.89053	0.90507	-0.01454	4.1143	ArcSin	1.1555	1.1316	0.0239	2.6326
Vac	Pos	0.10947	0.09493	0.01455	4.1143	Log	0.1008	0.0907	0.01006	3.423
Det	Pos	0.16264	0.04649	0.11616	6.0797	Log	0.1314	0.0454	0.08593	4.9582
A4	Pos	1.2254	1.2076	0.0178	3.1054	Log	6.5935	6.6012	-0.0077	2.6446
A9	Pos	0.0633	0.02511	0.03819	3.313	Log	0.0576	0.0248	0.03276	2.6696
A10	Neg	1.4196	1.4253	-0.0057	2.106	ArcSin	0.4196	0.4253	-0.0057	2.106
MarOwn	Pos	0.24678	0.21217	0.03461	4.1004	Log	0.2126	0.1924	0.02019	3.2646
MarRen	Pos	0.06961	0.05762	0.01199	4.1924	Log	0.0659	0.056	0.00984	3.7011
ManOwn	Pos	0.01592	0	0.01592	2.2325	Log	0.0155	0	0.01549	2.123
ManRen	Pos	0.02283	0	0.02283	8.565	Log	0.0155	0	0.01549	7.6365
WomOwn	Pos	0.06467	0.04077	0.0239	5.2	Log	0.0604	0.04	0.02045	4.4616
WomRen	Pos	0.10926	0.05433	0.05493	5.4765	Log	0.0968	0.0529	0.04389	4.2963
NFOwn	Pos	0.20272	0.19088	0.01184	3.9489	Log	0.1793	0.1747	0.0046	3.179
NFRen	Neg	0.33781	0.33912	-0.0013	2.6516	ArcSin	0.355	0.346	0.009	3.5128

Green shows the variable used. Only two variables were not normalized.

Each explanatory variable set represents housing or homeowner characteristics. The explanatory variable datasets are as follows: log of percent of the total number of people twenty-four years or above according to education level obtained, log or arcsin of percent of the total home value, log or arcsin of the percent for the total house-hold income, log or arcsin of percent of occupied homes, log or arcsin of the percentage of the total type of housing (attached and number of attached units), and the log or arcsin of the percentage of the total head of households that showed family, marital status, and ownership status. The individual explanatory variables and the data sets are presented in Table 1.

Housing characteristics may portray capable guardianship for any passerby. Two variable sets use housing attributes to show the perception of a capable guardian: home value and housing type.

The first explanatory variable set is based on the ‘value of home’. The value of the home is separated into three different levels: under $199,999, $200,000 - $999,999, and above $1,000,000. The reason for the wide range for the middle variable rests on the high median value of a home in DC: $535,000. Most likely, breaking this variable into two would not produce a different result.

The second housing characteristic uses the type of home. Detached homes are chosen because there is a physical space which increases the risk of being observed in a private area. Below in Table 6 are the raw numbers prior to being converted to ratios. The total number of housing units used in the survey is in the left column. The percentage is obtained by dividing the number of each type of unit by the total number of units. Furthermore, multiple variables are joined (Table 7). For example, instead of separating ‘1 unit attached,’ ‘2 units attached,’ and ‘three or four units attached,’ the different ranges are placed into a single variable ‘1 – 4 units attached’ to provide a larger sample size for the alternative variable.

Table 6: An example of count data for an explanatory variable

Total housing units							
	1-unit, detached	1-unit, attached	2 units	3 or 4 units	5 to 9 units	10 to 19 units	20 or more units
2,828	220	1,233	40	209	186	115	825

Table 7: Variable after grouping and as a ratio

Total housing units				
	1-unit, detached	1-4 unit, attached	5-9 units	10+ units
2,828	0.0779	0.5241	0.0658	0.3324

The third housing characteristic uses information about the homeowner. This is composed of marital status and ownership. There are eight possibilities between married, not married, man or woman head of household without spouse, nonfamily, and own or rent. The Census Bureau's survey did not consider families outside of the traditional makeup, such as same sex couples, so a larger range of potential error within this variable may exist, depending upon how these families chose to participate in the survey.

A fourth housing characteristic uses occupancy and vacancy percentages. This variable is the most basic of those being tested. If merely being occupied measures as significant, then capability is undermined.

The fifth variable set uses the household income as the measurement. There are four different brackets: up to $39,999; $40,000 - $74,999; $75,000 - $199,999, and above $200,000.

The sixth variable shows the highest education level obtained: no high school diploma, high school diploma, some college, bachelor's degree and higher.

3.5 Exploratory Regression and Ordinary Least Squares (OLS)

Two phases of exploratory regression are used to find explanatory variables for further analysis in Ordinary Least Squares (OLS). Phase one identifies variables with a negative relationship to violent crime. All twenty-five explanatory variables are entered into exploratory regression, and those with a negative linear relationship to violent crime are moved for further analysis. Phase two required three characteristics for the explanatory variable to be analyzed further in OLS: When belonging to a model (with the other explanatory variables) in exploratory regression, the variable must 1) show a negative linear relationship to violent crime; 2) measure as significant ($p<0.05$); and 3) show a variance inflation factor (VIF) score of less than seven. The VIF score measures the amount of collinearity between variables. After meeting these three criteria, the explanatory variables are analyzed using OLS.

Esri's tool Ordinary Least Squares (OLS) is used to show whether or not each explanatory variable and the cumulative model is significant (results were considered significant when $p<0.05$). There are six subdivisions or assessments.

1) The coefficient ought to be negative to ensure a negative linear relationship exists between violent crime (the dependent variable) and the explanatory variable.

2) Esri's OLS tool checks for co-linearity and computes an index called the Variance Inflation Factor (VIF). Essentially this test makes sure two variables are not redundant by providing a score based on co-linearity and the guidance is not to keep variables with scores above seven (Esri 2016). For example, variance may be high between income and home value because, for the most part, to be able to afford an expensive home, a person would most likely be in the higher income bracket. However, this may not end up being the case. There may be many high-income learners who do not buy or live inside expensive homes. In any case, variance needs to be tested to make sure a variable can stand on its own.

3) To assess the statistical significance of the resultant model, the Koenker (BP) statistic is run. The Koenker (BP) statistic to assesses stationarity between the dependent variable and the

explanatory variable. Stationarity exists when relationships between the variables contain consistency. This is also known as homoscedasticity. For example, if high-income earners are consistently present in high percentages in violent crime cold spots all of the time, then the relationship is stationary. However, if high-income earners are present in high percentages in both cold and some hot spots, then the relationship is said to be heteroscedastic, and the Koenker (BP) test shows this as statistically significant ($p<0.05$). To be determined as significant, the results from the Koenker (BP) statistic needs to be greater than the chi-squared calculation based on the number of degrees of freedom (dependent on sample size). When the results are statistically significant, the robust probability along with the Joint Wald Statistic (described below) is used to show the model's significance. However, if the Koenker (BP) statistic is not significant, then the probability along with the Joint F-Statistic (also described below) shows the overall model's significance. Either the Joint F-Statistic and/or the Joint Wald Statistic shows whether the model is or is not significant (ArcGIS 2012).

4) The Joint Wald and Joint F-Statistic are used to measure the overall performance of the model by setting the null hypothesis for the explanatory variable at ninety-five percent. The Joint Wald statistic measures as being significant when the result is greater than the chi-squared calculation. The Joint F-Statistic uses a t-statistic and requires normal distribution of the data. In other words, when variables show heteroscedasticity, then this statistic is not used. So, when the Koenker (BP) statistic is significant, then the Joint Wald statistic is used and vice versa for when the Koenker (BP) statistic is not significant (ArcGIS 2012).

5) The Jarque-Bera statistic checks whether the residuals are normally distributed and for model bias (Esri 2016). The residuals ought to be randomly distributed. If not, then, most likely a key variable is missing or there is strong heteroscedasticity between the dependent and

explanatory variables. The Jarque-Bera statistic shows to be probable when the results are greater than the chi-squared calculation based on two degrees of freedom.

6) R-squared indicates a percentage of the response variable variation between the dependent variable (denoted by y) and one or more independent variables (denoted by x) in a linear regression model. The R-squared values are given as a decimal between zero and one to show the strength of the relationship between the explanatory variables and the dependent variable, with a value closer to one showing a stronger relationship exists. However, the R-squared results are not the main focus for this study.

The adjusted R-squared score may measure the strength of the relationship, but high R-squared scores are not expected nor is there a requirement to show that the model shows a significant relationship exists between violent crime and the proxy model. The purpose is to test whether a proxy model suggests the perceived presence of a capable guardian deters crime, but not to what extent. Any number to show a benchmark would be arbitrary, so none is put forth.

3.6 Geographically Weighted Regression (GWR)

Geographically Weighted Regression (GWR) works well when data is nonstationary because it uses a local form of linear regression. Instead of fitting a model to an entire study region (global), GWR estimates coefficient values for every chosen point, giving most of the weight to the points that are closest to the center; the results provide information about the relationship between the dependent variable and one or more independent variables concerning geographical differences (Fotheringham, Brunsdon and Charlton 2002). Although different bandwidth, or the number of neighbors, can be chosen by the researcher, this analysis uses Esri’s built in function to optimize bandwidth based on AIC. The primary use for GWR is to attempt to understand the strengths and weaknesses of the relationships for each explanatory variable to violent crime.

Prior to running GWR, a model is built and run in OLS. GWR may help a researcher understand the variables and relationships better, but GWR cannot decipher collinear relationships, so this must be observed using OLS. According to Esri, the VIF should not be above seven. If the VIF is greater than seven, then GWR does not yield results that can be used. GWR analysis enables the researcher to understand nonstationary data along with areas where the model performs well and where it does not. This analysis provides local R-squared values from 0.0 to 1.0. High values show high performance.

If the model run in OLS passes most tests, but the Koenker (BP) is significant, suggesting the data is non-stationary, then the model is a fitting candidate for GWR analysis. A GWR analysis may show where the locally weighted regression coefficients move away from their global values, which may provide possible reasons for non-stationarity within an explanatory variable and provide better oversight in future projects with a similar scope. For example, the census data may need to be filtered differently, or a different aggregation may need to be applied to improve the model's performance.

Chapter 4: Results

This study uses housing characteristics as proxies for a capable guardian and tests whether a correlation exists between capable guardianship and the deterrence of violent crime. The following chapter provides the results. In sections 4.1 and 4.2 the results of a hot spot analysis and the relationships amongst violent crime data are examined. In the subsequent sections of this chapter, regression analysis is used to explore the housing and homeowner explanatory variables, including a proxy model that measures as being significant, suggesting that the presence of a capable guardian deters violent crime. Finally, Geographically Weighted Regression (GWR) is used to map the coefficients and provide a visual display of the relationship between the explanatory variables and violent crime in relation to each of the other variables.

The overall results show violent crimes to be clustered at a ninety-nine percent confidence level. Also, at least three hot spots and one cold spot is visually identified in the results of the optimized hot spot analysis. The strong inter-relationships of the four violent crimes shown in a hot spot analysis and measured during regression analysis (whereby each violent crime is a dependent variable and the other violent crimes are the explanatory variables) justifies combining these into one variable for use as a dependent variable in the final regression analysis.

4.1 Hot Spot Analysis of the Violent Crimes

Hot spot analysis shows each of the four violent crimes (homicide, assault with a dangerous weapon, robbery, and sexual assault) are clustered at a ninety-nine percent confidence level (Figures 5 – 8). Although the hot spots and cold spots may expand and contract depending upon the violent crime, enough overlap is visible to suggest a strong relationship exists between the violent crimes when aggregated at the census block group level. Figures 5 - 8 depict at least three hot spots and one cold spot for all of the violent crimes independently run. Overall, three hot spots and one cold spot is identified within Washington, DC for all of the violent crimes combined (Figure 9). A cold spot is identified near the white house for homicide and assault with a dangerous weapon, but this cold spot is no longer identified for robbery or sexual assault. Also, Central DC shows crimes to be along the main roads in many places, but the numbers of crimes compared to the hot spots identified simply keep these areas from being clustered. Reference numbers are placed near the center of the masses to show the same areas are both plagued (hot spots) and vacant (cold spot) of violent crime, regardless of which violent crime is input.

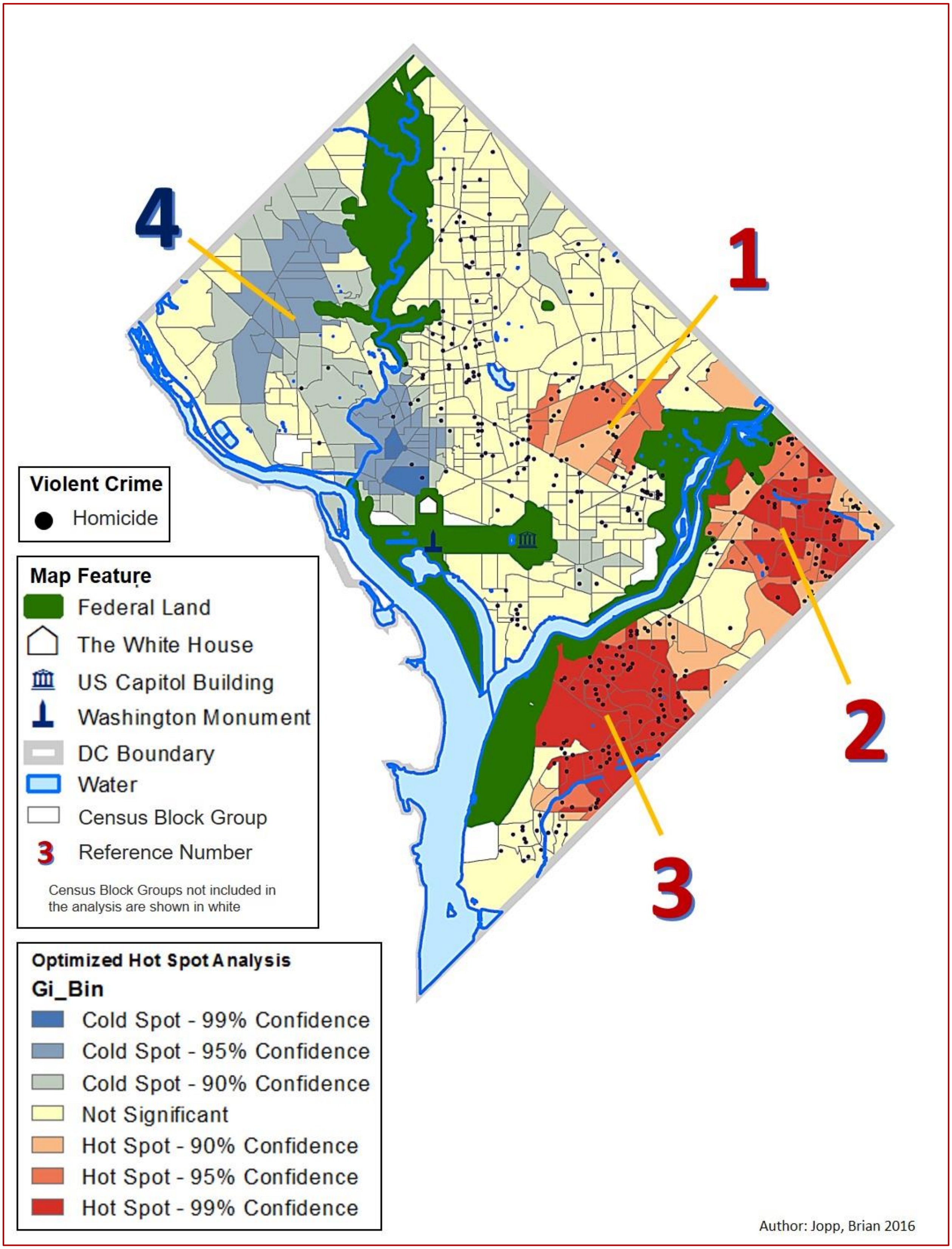

Figure 5: Homicide hot spots by census block group

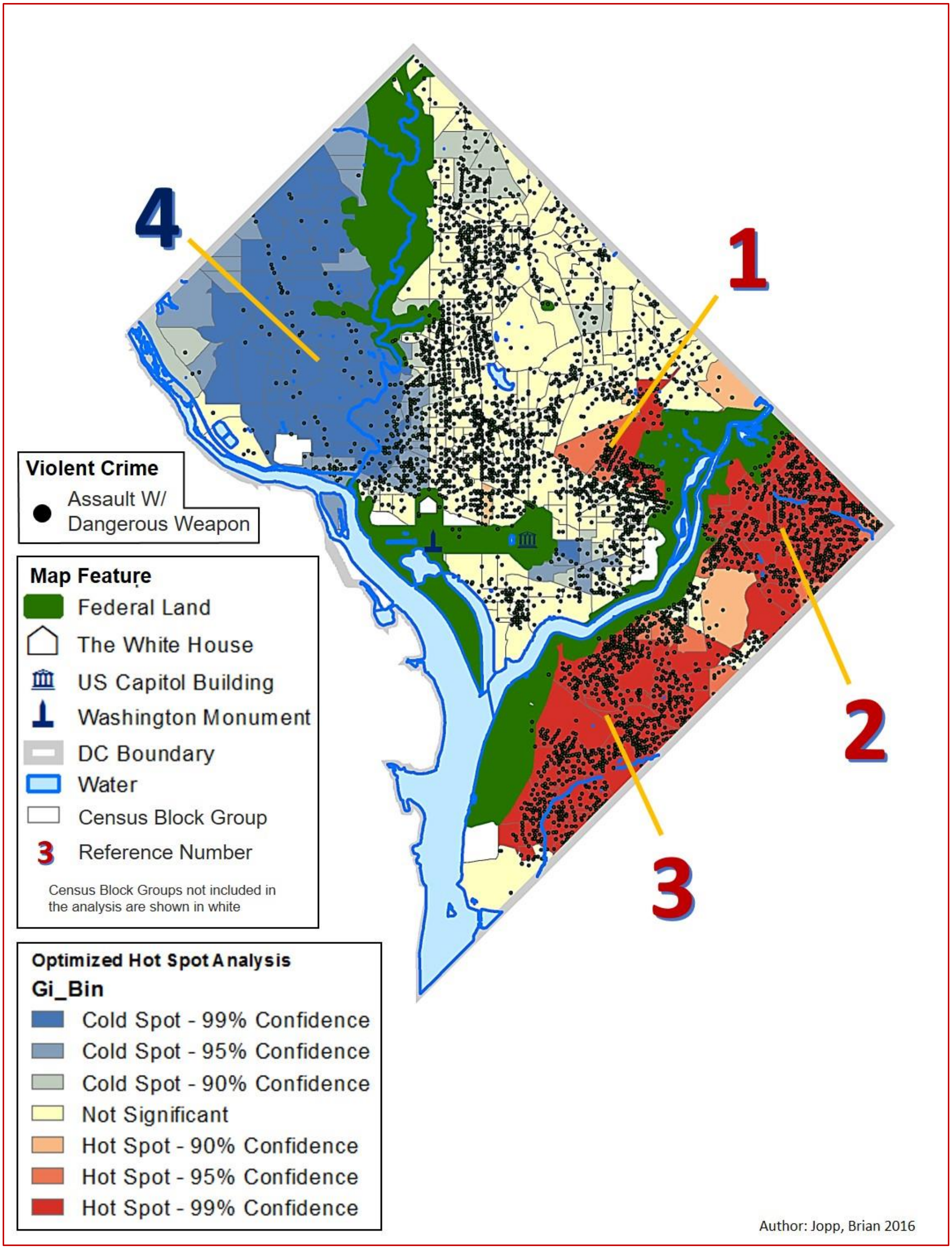

Figure 6: Assault with a dangerous weapon hot spots by census block group

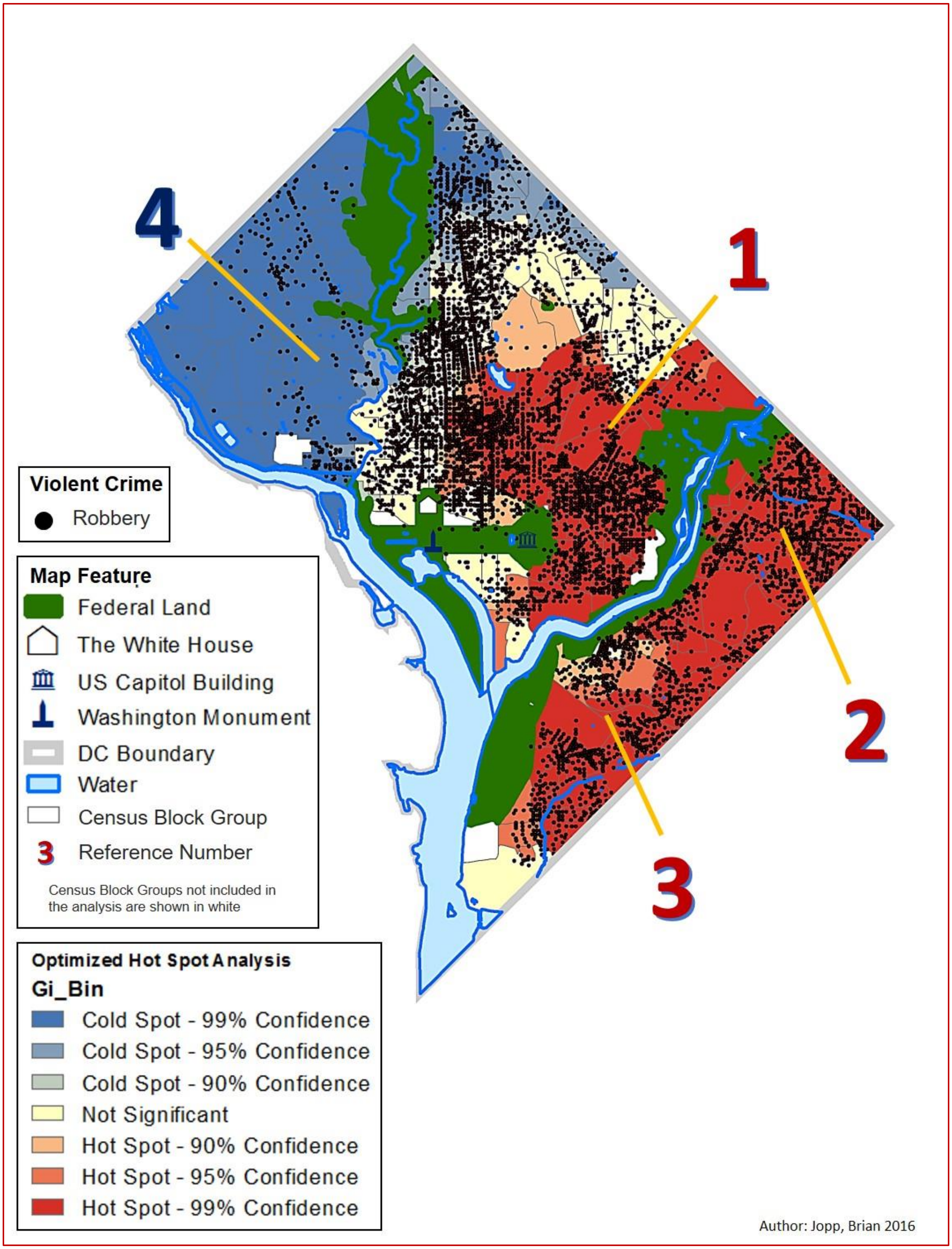

Figure 7: Robbery hot spots by census block group

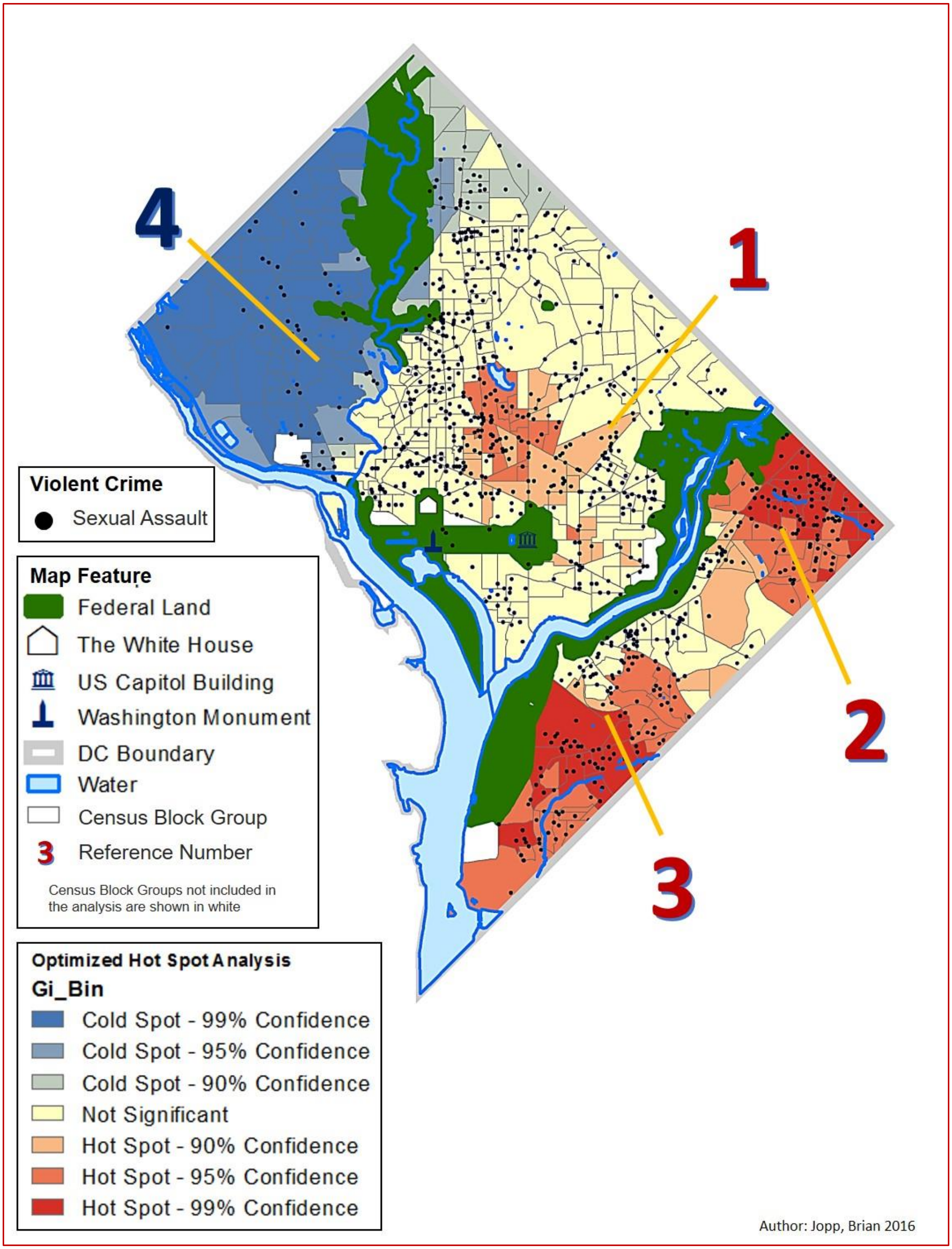

Figure 8: Sexual assault hot spots by census block group

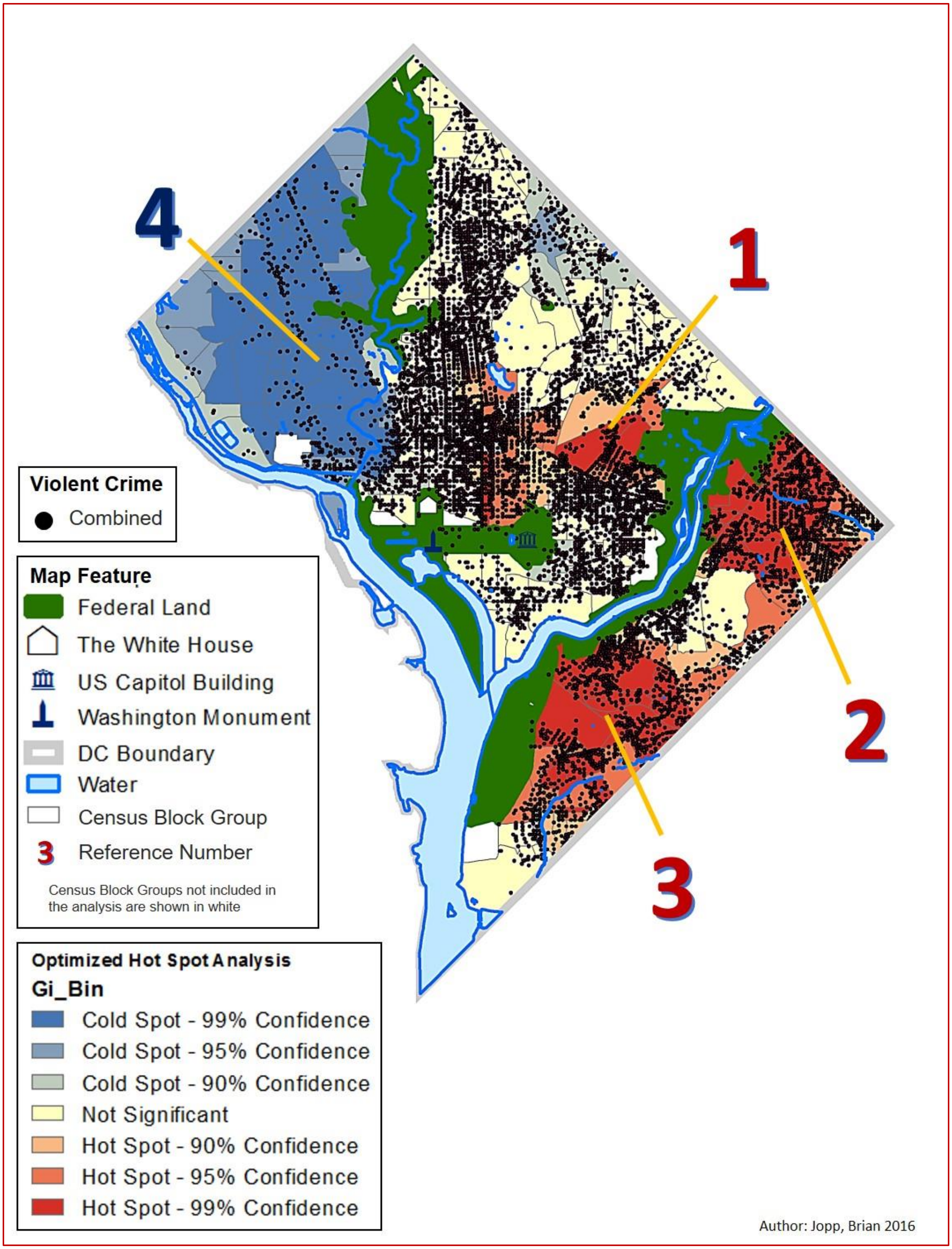

Figure 9: Combined violent crime hot spots by census block group

The results show the violent crimes are clustered. To better understand these relationships, the next section explains the results of exploratory regression.

4.2 Exploratory Regression and Ordinary Least Squares of the Violent Crimes

Exploratory regression shows a 100% positive relationship exists between all four violent crimes (homicide, assault with a dangerous weapon, robbery, and sexual assault). However, homicide only shows to be 100% significant in relation to assault with a dangerous weapon, whereas homicide is only fifty percent significant to robbery and sexual assault, as shown in Table 8 below.

Table 8: Exploratory regression results for four violent crimes

Dep Var	Exp Var	Significant %	Negative	Positive	AdjR2
Homicide	D Weapon	100	0	100	0.38
Homicide	Robbery	50	0	100	0.28
Homicide	S Assault	50	0	100	0.22
D Weapon	Homicide	100	0	100	0.38
D Weapon	Robbery	100	0	100	0.69
D Weapon	S Assault	100	0	100	0.45
Robbery	Homicide	50	0	100	0.28
Robbery	D Weapon	100	0	100	0.69
Robbery	S Assault	100	0	100	0.37
S Assault	Homicide	50	0	100	0.22
S Assault	D Weapon	100	0	100	0.45
S Assault	Robbery	100	0	100	0.37

Table 9: OLS results for violent crime models

Dep Var	Exp Var	Robust_Pr	VIF	AdjR2
Hom				0.39
	D Weapon	0.000000	3.721837	
D Weapon				0.76
	Homicide	0.000000	1.455205	
	Robbery	0.000000	1.820588	
	S Assault	0.000000	1.678653	
Robbery				0.70
	D Weapon	0.000000	1.814231	
	S Assault	0.004949	1.814231	
S Assault				0.46
	D Weapon	0.000000	3.224401	
	Robbery	0.003531	3.224401	

Ordinary Least Squares defines this relationship further, as shown in Table 9 above. When assault with a dangerous weapon is input as the dependent variable, then all violent crimes are significant (Table 8) and the adjusted R-squared score measures at 0.76 (Table 9). When using Ordinary Least Squares, only assault with a dangerous weapon measures as being significant to homicide. Whereas, robbery, sexual assault, and assault with a dangerous weapon shows a significant relationship to one another. The variability remains low in all models.

These results combined with the prior hot spot analysis provides enough rationale to combine all of the violent crimes into one dependent variable to be used in the main study testing whether the presence of a cable guardian deters crime. The dependent variable consists of 26,003 violent crimes occurring from 2012 through 2015: homicide (n = 440), assault with a dangerous weapon (n = 9,603), robbery (n = 14,824), and sexual assault (n = 1,136). The dependent variable is aggregated by census block group, count data is changed to percentages, and since it is not normally distributed, the variables are normalized using the log function.

4.3 Exploratory Regression and Identifying Explanatory Variables for OLS

To test whether the presence or perceived presence of a capable guardian shows a negative correlation to violent crime, six variable sets are chosen: 1) education attained for people who were twenty-four years of age and above; 2) home values; 3) income per household; 4) occupancy; 5) type of housing; 6) and type of household (family) along with ownership. Two phases of exploratory regression are used to isolate significant explanatory variables with negative linear relationships to violent crime.

In the first phase of exploratory regression, twenty-five explanatory variables belonging to one of the six data sets are tested to identify the variables with a negative linear relationship to violent crime. The settings within Esri's exploratory regression tool are set to allow for seven possible variables in the model. In Table 10, five explanatory variables shows a significant negative linear relationship to violent crime: the log of percent of the total number of homes with values over $1,000,000 (LV1M); the log of the percent of the total number of homes owned by married couples (LMAROWN); log of percent of the total number of the population who are twenty-four years of age and over and earned a bachelor's degree or higher (LBACH); the log of percent of the total number of homes that are detached (LDET); and the log of percent of the total number of households with income over $200,000 per year (although not as strong as the other four explanatory variables).

For the second phase of exploratory regression, the five identified explanatory variables with a negative linear relationship are entered into Esri's exploratory regression analysis tool. Due to lack of significance, income over two hundred thousand dollars is dropped, as shown in

Table 11 by the lack of an asterisk. After the second phase of exploratory regression, a model consisting of four explanatory variables remain (Table 11). I refer to this as the guardian model.

Table 10: Exploratory regression phase one - negative linear relationships to violent crime

Variable	Significant Percentage	Negative Percentage	Positive Percentage
Home Values Over $1,000,000	99.90	100.00	0.00
Married and Owns Home	99.00	99.92	0.08
No High School Diploma	98.44	0.00	100.00
Bachelor's Degree or Higher	95.77	98.54	1.46
Detached Home	92.73	99.90	0.10
Attached Home 4 and Under	85.95	0.20	99.80
Home Value $300,000 - $999,000	84.56	2.92	97.08
Income Over $200,000	82.86	95.66	4.34
High School Diploma	81.32	1.07	98.93
Woman No Spouse, Rents Home	71.11	3.64	96.36
Income $75,000 - $199,999	69.48	20.82	79.18
Some College	61.64	21.79	78.21
Attached Homes 10 and Above	55.12	78.73	21.27
Income Under $40,000	53.89	14.83	85.17
Non family, Rents Home	52.24	25.27	74.73
Home Value $199,999 and Under	49.68	1.41	98.59
Woman No Spouse, Owns Home	37.19	38.64	61.36
Man No Spouse, Rents Home	34.42	2.04	97.96
Income $40,000 - $74,999	29.43	61.45	38.55
Non family, Owns Home	20.78	26.76	73.24
Attached Homes 5 -9 Units	19.48	6.01	93.99
Vacant	17.70	14.94	85.06
Occupied	14.43	57.36	42.64
Married, Rents Home	14.12	29.96	70.04
Man No Spouse, Owns Home	7.16	31.09	68.91

Table 11: Exploratory regression phase two - four standing significant explanatory variables

Highest Adjusted R-Squared Results

AdjR2	AICc	JB	K(BP)	VIF	SA	Model			
0.46	1155.10	0.00	0.00	2.28	0.00	-LBACH***	-LV1M***	-LDET***	-LMAROWN***
0.46	1162.22	0.00	0.00	3.87	0.00	-LBACH***	-LV1M***	-LI200	-LDET***
0.45	1166.02	0.00	0.00	4.52	0.00	-LBACH***	-LV1M***	+LI200	-LMAROWN***

4.4 Results of Ordinary Least Squares Model and Details

The guardian model (highlighted in Table 11 above) consists of a dependent variable encompassing all four violent crimes (homicide, assault with a dangerous weapon, robbery, and sexual assault), and four explanatory variables (LBACH, LV1M, LDET, and LMAROWN). When the guardian model is plugged into Ordinary Least Squares (OLS), all four variables show a negative linear relationship (shown by the negative symbol in the Coefficient column), all measure as being statistically significant (shown by the Robust Probability column), and register as maintaining a low variability in relation to each of the other explanatory variables, being under three (shown in the VIF column), as shown in Table 12 below.

Table 12: Summary of OLS results for the guardian model

Variable	Coefficient	StdError	t-statistic	Probability	Robust SE	Robust t	Robust Pr	VIF
Intercept	5.002586	0.10337	48.394911	0.000000*	0.084808	58.987219	0.000000*	-----------
LBACH	-2.306739	0.253663	-9.093728	0.000000*	0.277158	-8.322816	0.000000*	1.605068
LV1M	-1.723681	0.386703	-4.457371	0.000013*	0.406172	-4.243717	0.000031*	1.613482
LDET	-0.966088	0.291938	-3.30922	0.001026*	0.27787	-3.476767	0.000572*	1.718721
LMAROWN	-1.37913	0.504748	-2732313	0.006540*	0.447836	-3.079541	0.002214*	2.282136

Since the Koenker (BP) statistic is significant, the Joint Wald statistic is used to measure the overall performance of the guardian model: the results show the model to be statistically significant, $p<0.05$* (Table 13). Using the robust probabilities and the Joint Wald statistic, the model suggests capable guardianship or the perception of capable guardianship deters crime.

However, the Jarque-Bera Statistic measures as being statistically significant ($p<0.000000$*), and the residuals are spatially clustered at a 99% confidence level, which suggests a key variable is not included in the model. The adjusted R-squared score is 0.46, which supports the model is missing a key variable (Table 13).

Table 13: OLS diagnostics for the guardian model

Input Features	Capable G.	Dependent Variable	Violent Crime
Number of Observations	446	Akaike's Information Criterion (AICc)	1155.096018
Multiple R-Squared	0.46964	Adjusted R-Squared	0.464829
Joint F-Statistic	97.6277539	Prob (>F), (4,441) degrees of freedom	0.000000*
Joint Wald Statistic	467.317281	Prob (>chi-squared), (4) degrees of freedom	0.000000*
Koenker (BP) Statistic	25.527869	Prob (>chi-squared), (4) degrees of freedom	0.000039*
Jarque-Bera Statistic	111.879492	Prob (>chi-squared), (2) degrees of freedom	0.000000*

Below are the histograms and scatterplots of the OLS (Figure 10) that show the relationship between each explanatory variable and the dependent variable (violent crime). The scatterplots for LV1M and LDET do not show to be as strongly correlated to violent crime as LMAROWN and LBACH, which explains the potential heteroscedasticity and missing variables signified by the Jarques-Bera statistic and clustered residuals. The positively skewed variables LV1M and LDET suggest a different transformation or a more fitting aggregation may yield better results.

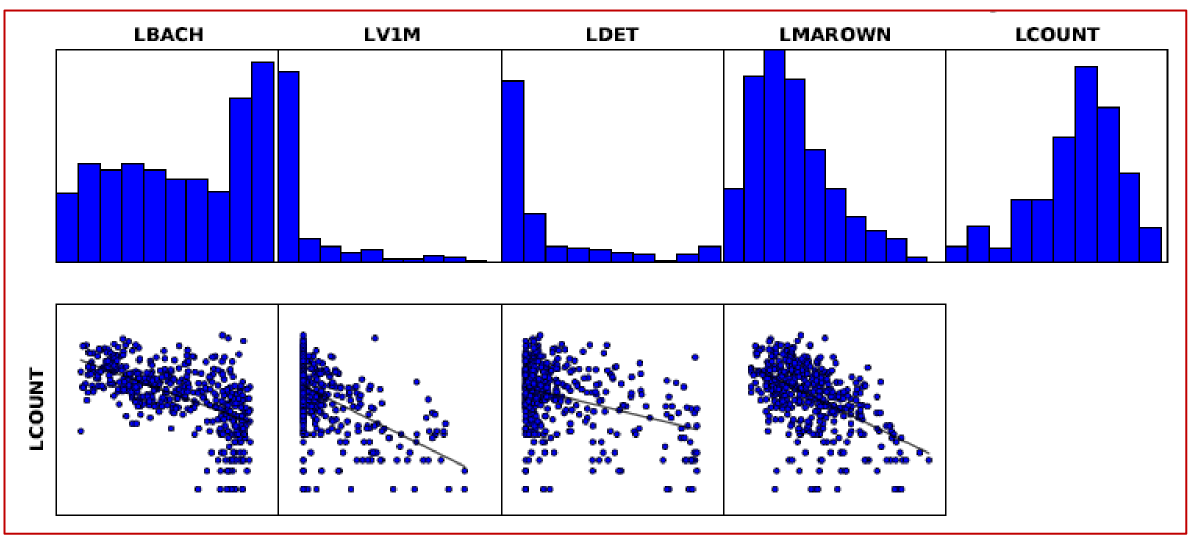

Figure 10: Variable distributions and relationships in the guardian model

The histogram of the standardized residuals did resemble a Gaussian curve, showing the model is probably not biased (Figure 11).

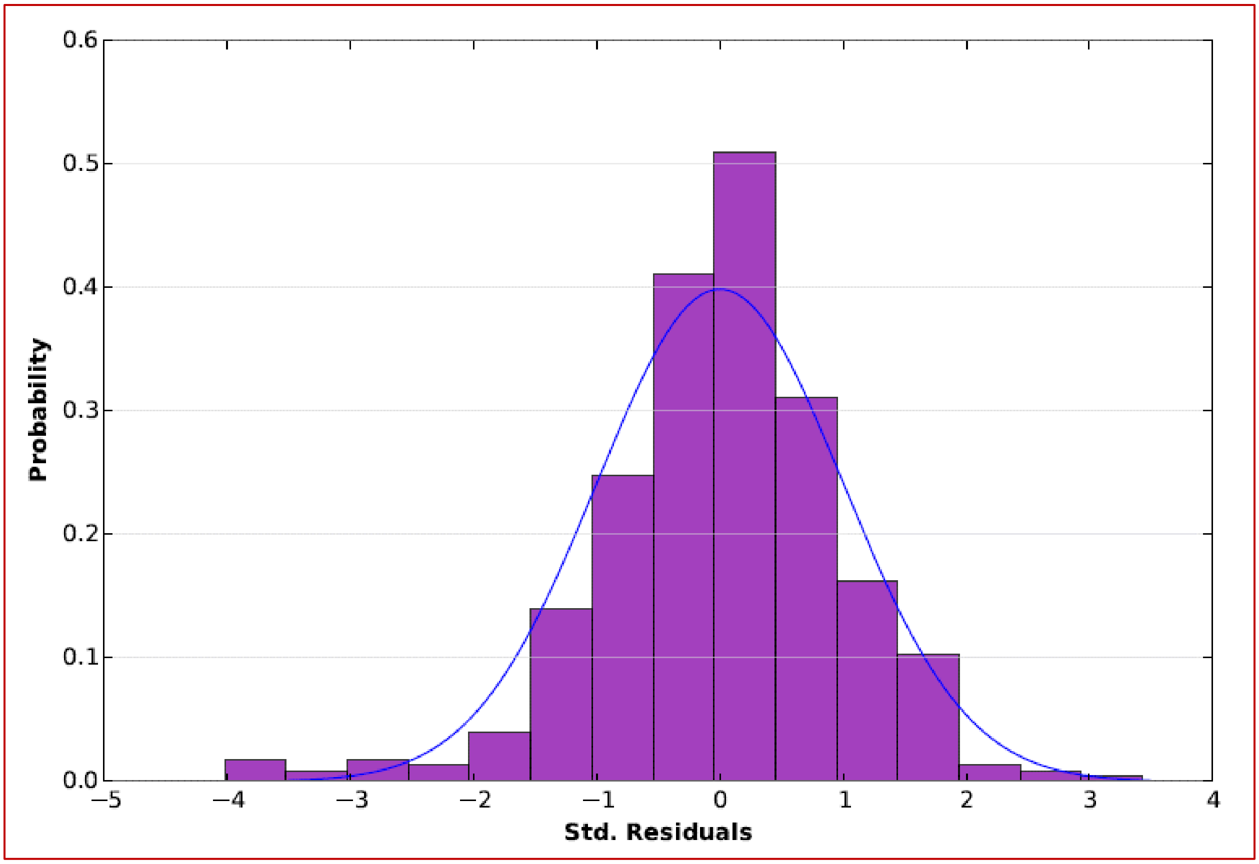

Figure 11: Histogram of standardized residuals for the guardian model

Finally, the standard residuals are compared to the predicted plot; an archetype result does not show any pattern. On the next page, Figure 12 does not show an obvious pattern, so the model appears to work. The positive as well as negative standard residuals are relatively even as can be observed through the identified colors above and below the X axis.

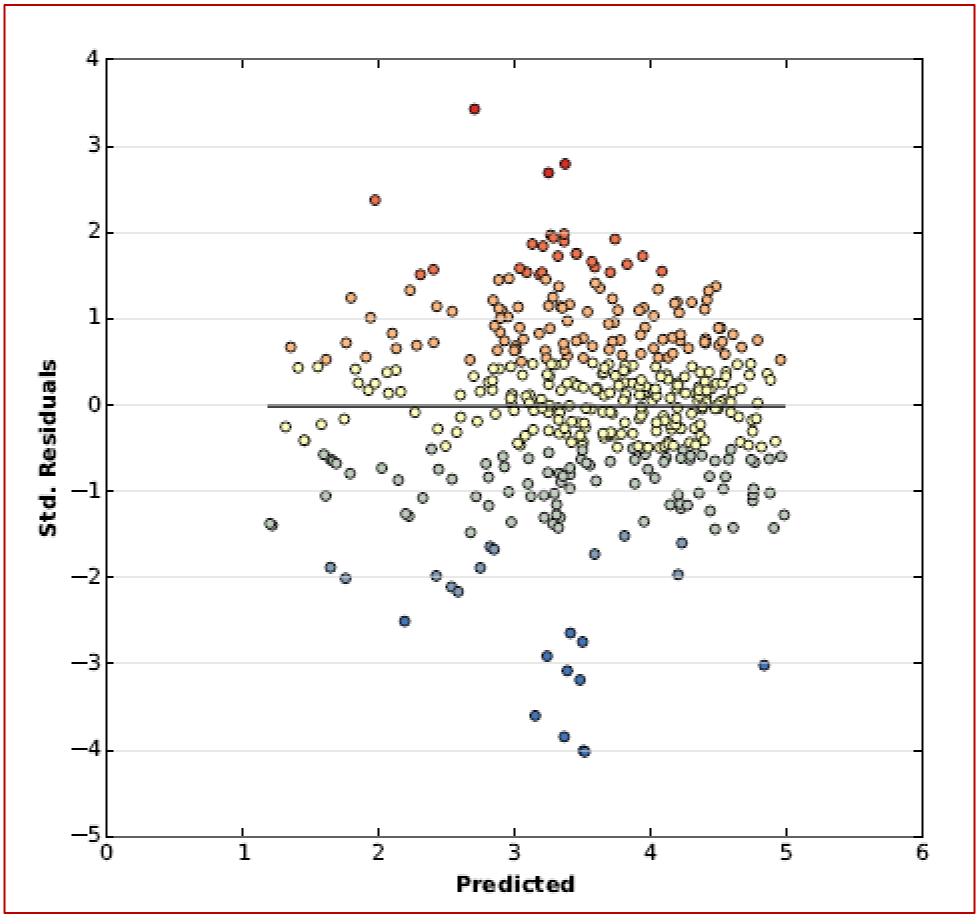

Figure 12: Residual vs predicted plot guardian model

4.5 Geographically Weighted Regression

Geographically Weighted Regression's adjusted R-squared value of 0.57 is an improvement from the Ordinary Least Squares (OLS) R-squared value of 0.46. The dependent variable is violent crime. The explanatory variables are LV1M, LBACH, LDET, and LMAROWN; for reference the distribution of these values is shown in Figure 13 with the values categorized into quintiles. Within all of the explanatory variables, non-stationarity exists, which provides the main reason GWR shows better results in comparison to OLS. The model coefficients are mapped and can be found in Figure 14. Reference numbers are placed in the graphic to match with areas of interest identified in a prior hot spot analysis.

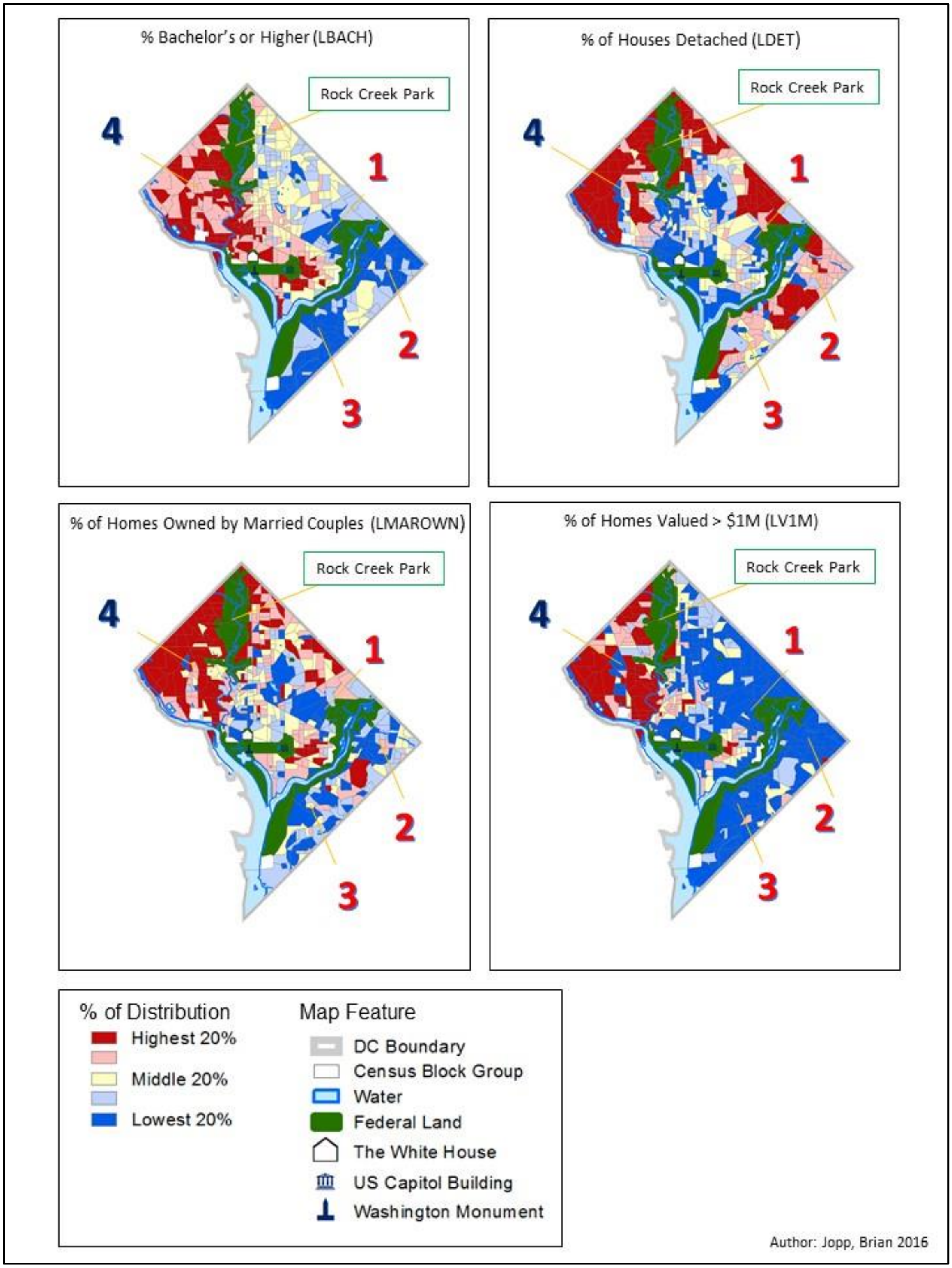

Figure 13: Distribution of the values of the explanatory variables in the guardianship model with values classed into quintiles

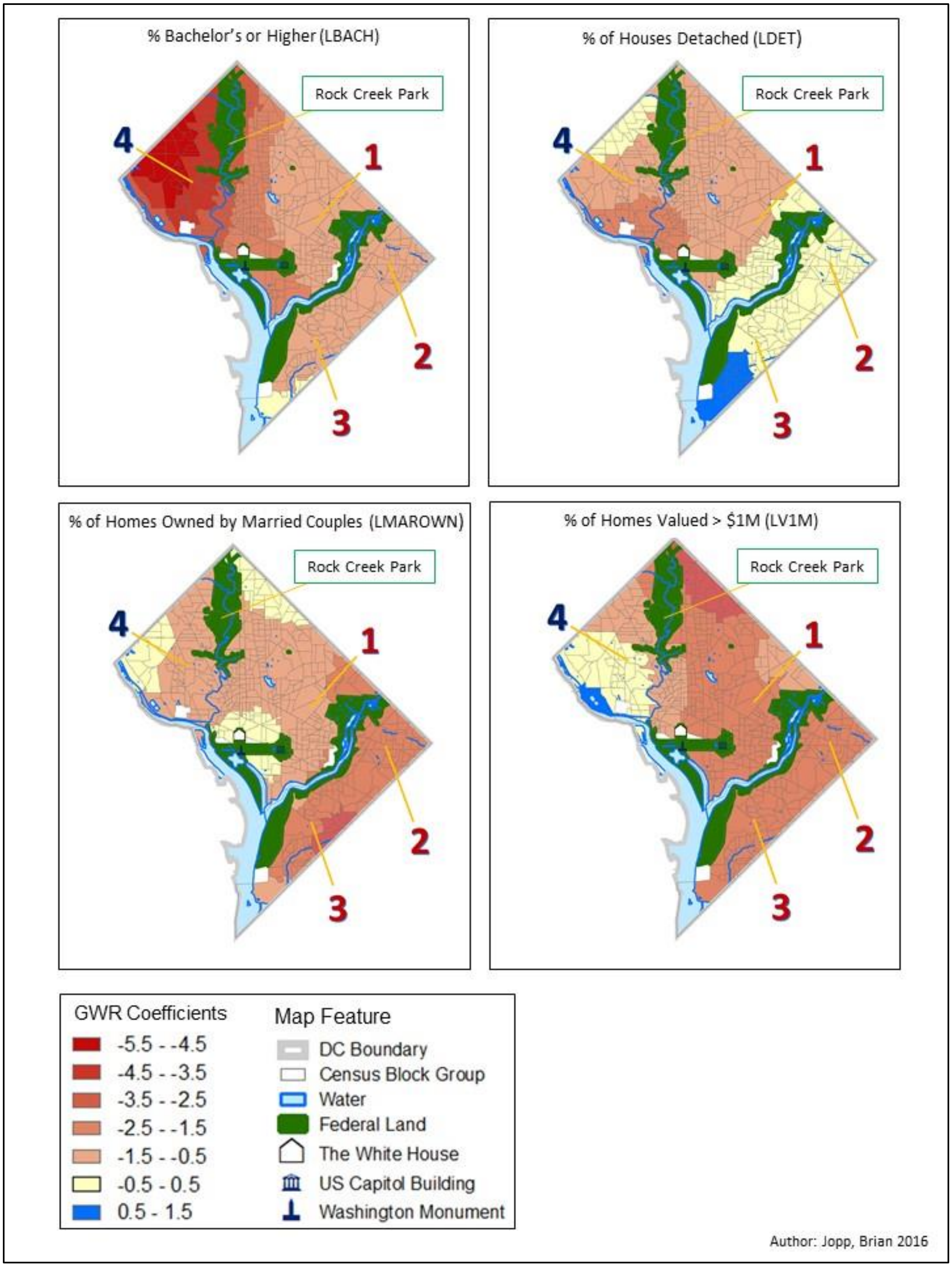

Figure 14: GWR guardianship model coefficients

Figure 14 shows all four of the model coefficients on a single page for comparison. The area of interest for this study is in northwestern DC, west of Rock Creek Park, near reference number 4. In this area, the coefficients for LBACH ranged down to -5.02, which is the strongest negative coefficient of any explanatory variable, suggesting LBACH is the most important factor in the model. Furthermore, the distribution of the variables match well with the area of interest (Figure 13), meaning numerically as well as spatially LBACH matches with the area of interest.

LDET's coefficient, when mapped in Figure 14, shows a moderately strong negative value in the area of interest near reference number 4. The coefficient's minimum is -2.45, which suggests the variable is important but to a lesser degree than LBACH. However, the distribution (Figure 13) for LDET shows a widely dispersed number of census tracts in the top quintile throughout DC. Furthermore, at least six census block groups within the area of interest are in the lowest quintile, weakening the strength of the variable.

According to Figure 14, the coefficients for LMAROWN and LV1M show no particular relationship to the area of interest. For LMAROWN, many census block groups with the lowest coefficients are located outside of the area of interest. According to the distribution (Figure 13) of LV1M, many of the census block groups in the area of interest are in the lowest percentile, bringing down the significance of the variable. Furthermore, there are many census block groups with high percentages of homes over $1 million located outside the area of interest. Census block groups where LMAROWN shows the strongest negative coefficients are located outside of the area of interest (Figure 14), and the percentage of homes in this category varies throughout the study area (Figure 13).

Chapter 5: Conclusions and Discussion

In the routine activity theory there are three components that must come together in time and space for a crime to exist: 1) a likely offender; 2) a suitable target; and 3) the absence of a capable guardian (Cohen and Felson 1979). This thesis examines homeowner and housing characteristics in relation to spatial cold spots of violent crimes. Variables with a negative correlation with violent crimes are fit into an Ordinary Least Squares (OLS) and a Geographically Weighted Regression (GWR) analyses. The final variables (LBACH, LV1M, LDET, and LMAROWN) are chosen as the "guardian model." The results in OLS suggests a negative correlation exists between housing and homeowner characteristics and violent crime. Thus, capable guardianship may be correlated with violent crime cold spots. However, when the coefficients are mapped in GWR, only LBACH shows a strong influence within the area of interest, west of Rock Creek Park near reference number four (Figure 14).

In the following three sections, a discussion and some conclusions are drawn concerning the relationship between housing and homeowner characteristics to violent crimes. The first section focus is on the relationships of the violent crimes using hot spot analysis and regression analysis. The second section discusses the explanatory variables and the results of the model fit into OLS and GWR. In the final section of this chapter, future work is discussed concerning the application of quantitative spatial analysis in accord with violent crime.

5.1 Violent Crimes Using Hot Spot and Regression Analysis

A hot spot analysis of four violent crimes (homicide, assault with a dangerous weapon, robbery, and sexual assault), from 2012 through 2015 for Washington, DC shows a divided city. For the most part, the only cold spot by time and place throughout the study area is limited to the north and west of Rock Creek Park. At the same time, three hot spots are identified throughout the time period in all four violent crimes. The cold and hot spots for all four violent crimes remain clustered in many of the same areas throughout the whole study period. Therefore, all four violent crimes are grouped into one dependent variable.

The four violent crimes, when being used as both dependent and explanatory variables, measures at 100% as being positively related in an exploratory regression analysis. However, further analysis shows homicide as a weak connection to robbery and sexual assault, but homicide does show a significant relationship to assault with a deadly weapon. Robbery may end up as a homicide, but generally speaking, the motivation for robbery is to gain material possessions. Whereas, assault with a deadly weapon seems be more connected to murder, with respect to motivation. However, the low number of homicides (n = 450) in comparison to the other violent crimes, assault with a dangerous weapon (n = 9,603), robbery (n = 14,824), and sexual assault (n = 1,136), may skew the results.

When using exploratory regression, the strongest relationship exists between robbery and assault with a dangerous weapon (adjusted r-square = 0.69). Furthermore, the strongest model in OLS is measured when assault with a dangerous weapon stands as the dependent variable and the other three violent crimes are plugged in as explanatory variables. The adjusted R-squared score is 0.76, suggesting strong correlations exist between assault with a dangerous weapon and the other three violent crimes.

5.2 The Model Developed to Measure Capable Guardianship

A successful model uses the sum of all of the violent crimes from 2012 – 2015 as the dependent variable and four housing and homeowner characteristics as the explanatory variables to suggest capable guardianship deters violent crime. Six data sets consisting of housing and homeowner characteristics are used to identify variables negatively correlated with violent crime.

A model made up of four variables (LBACH, LDET, LV1M, and LMAROWN) suggests capable guardianship is correlated to violent crime.

LBACH is the strongest explanatory variable in the model, as the coefficients ranged as low as -5.02 in the census block groups within the area of interest. Furthermore, the distribution map (Figure 13) shows the area of interest had a high percentage of people who are at least twenty-four years old and who earned a bachelor's degree or above. For the most part, the areas with higher crime rates do not have a high percentage of people with a bachelor's degree or higher. There could be many reasons why a relationship between attainment of a bachelor's degree or higher and low crime exists. For example, a person with a bachelor's degree or higher may be more inclined to report crime, may be more resourceful in regards to using security systems along with other preventative measures, and may have stronger ties to influential people in the community. All of these reasons may lead to the deterrence of violent crime. However, another possible reason for areas with low crime rates and high percentages of people with a bachelor's degree or higher may be that people who gain higher education choose not to live in high crime areas.

Regardless of the reason, occupancy alone cannot explain areas with low crime rates. Occupancy was discarded after the first phase of exploratory regression. If occupancy alone made it into the final model, then capability would not be important. Occupancy ranked 22^{nd} out of the 25 variables tested with a measured significance at fourteen percent and with a fifty-seven percent negative linear relationship in models tested during exploratory regression analysis. This shows guardians need to be both present and capable. If occupancy alone determined deterrence, then this model is not valuable, in regards to the capable guardian theory. However, given this variable measures exceedingly low, highlights capability as a possible factor to deter violent crime.

Although the results and analysis of LBACH and LDET in OLS and GWR show strong and moderate relationships exist between education level and detached homes, in regards to low crime rates in northwestern DC, the other two variables are not as useful when concentrating on northwestern DC. In OLS, the adjusted R-squared score reaches only 0.46, the Jarque-Bera test is significant, and the residuals are clustered at the 99% level. In short, the model is missing at least one key variable. To resolve these problems, a finer aggregation or an aggregation based on housing type may greatly improve the results. For example, the boundaries for census block groups did not necessarily consider housing characteristics such as detached homes or rented versus owned.

GWR works well when data is nonstationary because it fits a model for every area and estimates coefficient values for each chosen point within each area. GWR gives weights to the points according to proximity, and the results show information concerning the relationships between the independent and dependent variables (Fotheringham, Brunsdon and Charlton 2002). GWR's adjusted R-squared score shows an increase from 0.46 in OLS to 0.57, which suggests the variables are nonstationary.

The purpose and intention of this study is not to show a direct cause/effect relationship exists between these selected variables and violent crime. In fact, the routine activity theory uses three components that must come together in time and space. In the original theory, Felson et al use the absence of a capable guardian. The active parts of the theory consist of a likely offender and a suitable victim coming together in time and space. A capable guardian may act as a deterrent to keep likely offenders from entering certain areas, and therefore, would be a difficult model to increase to a significant adjusted R-squared value. Yet, a hot spot analysis shows a divided DC (Figure 9), whereby a large cold spot is identified north and west of Rock Creek Park. In addition, when the coefficients are mapped (Figure 14), LBACH is the only variable where census block groups with the lowest coefficients match up with the area of interest. Does education level obtained (in this case bachelor's degree or higher) provide the best explanation for where the most capable guardians reside? More research in other cities could be conducted to find out if this is a universal trait: do communities with high percentages of people who have obtained a bachelor's degree or higher have lower violent crime rates?

5.3 Future Work in the Application of Spatial Analysis of Violent Crime

Quantitative spatial analysis may enable the researcher to better understand violent crime cold spots, the presence or perceived presence of a capable guardian (the third component of the routine activity theory), and housing characteristics. In an analysis of the four violent crimes (homicide, robbery, assault with a dangerous weapon, and sexual assault), a regression model with an adjusted R-squared value of 0.76 shows strong correlations exist when assault with a dangerous weapon is the dependent variable and homicide, robbery and sexual assault are entered as the independent variables. When all four violent crimes are combined and entered as the dependent variable, and four housing characteristics are entered as the independent variables (LV1M, LBACH, LDET, and LMAROWN) the adjusted R-squared value is 0.57 in a Geographically Weighted Regression (GWR) analysis. Both of these models suggest quantitative spatial analysis may be a tool for researchers to use when studying violent crimes. A future study to expand upon this analysis could be to take a closer look at housing and homeowner characteristics and violent crime cold or hot spots to search for other variables (the capable guardian model is missing at least one key variable).

Perhaps the housing and homeowner characteristics explored in this study may be a link for understanding how criminals decide to take on new areas to commit crimes. Do they decide to explore areas synonymous with less capable guardians present? Likewise, do they tend to avoid areas where capable guardians seem to live? If so, then it may be possible to predict where existing hot spots of crime will expand or not expand. Do criminals move into the areas with the least resistance, potentially measured by housing and homeowner characteristics? This study produced housing and homeowner variables that may provide insight into violent crime for researchers.

The study of physical housing characteristics may lead to a better understanding of criminal's decisions to expand into new areas. Although this study concentrates on violent crime cold spots, a researcher may want to reverse the analysis and study the hot spots. Do housing characteristics enable and deter where the criminals expand their area of operations? Studying the hot spots may better explain the explanatory variables when mapped in GWR. If so, then a larger-scale and more focused study concerning hot spots may show new areas where criminals are more likely to move into before committing violent crimes. These decisions may be based on the criminal's judgment whether capable guardianship exists or does not exist, and physical housing characteristics may factor into this judgment.

However, the variable with the strongest spatial relationship to areas with low levels of violent crime are not physical housing characteristics. In the guardianship model, LBACH measures as the strongest negative coefficient to violent crime in OLS and GWR within the area of interest. This study highlights education as an area for more quantitative spatial research to better understand violent crime and capable guardianship. Conducting this same analysis in other cities in the United States may be useful to understand if these results are part of the culture inherent to DC or universal to all major cities in the United States. If higher education is negatively correlated to violent crime on a universal level, then this adds yet another reason for society and decision-makers to invest into enabling individuals to earn a higher education.

Works Cited

Anderson, Tammy L. 2014. *Understanding Deviance: Connecting Classical and Contemporary Perspectives.* New York: Routledge Taylor and Francis group.

Bolstad, Paul. 2012. *GIS Fundamentals.* White Bear Lake: Eider Press.

Bowser, Muriel. 2016. *Metropolitan Police Department.* May 20. Accessed May 21, 2016. http://mpdc.dc.gov/page/district-crime-data-glance.

Bowser, Muriel. 2013. *Metropolitan Police Department Annual Report 2013.* Washington: Metropolitan Police Department.

Bowsner, Muriel, and Jeffrey S DeWittt. 2015. *District of Columbia Economic and Revenue Trends: January 2015.* Washinton: DC Office of Revenue Analysis.

Bozeman, J Michael. 2014. *The Common Language of Homicide and Suicide: Evidence of the Value of Durkheim's Typologies.* El Paso: LFB Scholarly Publishing LLC.

Browning, Christopher R. 2002. "The Span of Collective Efficacy: Extending Social Disorganization Theory to Partner Violence." *Journal of Marriage and Family* 833 - 850.

Cahill, Meagan, and Gordon Mulligan. 2007. "Using Geographically Weighted Regression to Explore Local Crime Patterns." *Social Science Computer Review* 174-193.

Chiricos, T G. 1987. "Rates of Crime and Unemployment: An Analysis of Aggregate Research Evidence." *Social Problems* 187-212.

Cohen, Lawrence E, and Marcus Felson. 1979. "Social Change and Crime Rate Trends: A Routine Activity Approach." *American Sociological Review* 588-608.

Cozens, Paul, and Terence Love. 2015. "A Review and Current Status of Crime Prevention Through Environmental Design (CPTED)." *Journal of Planning Literature* 393-412.

Delisi, Matt, Anna Kosloski, and Molly Sween. 2010. *Murder by Numbers: Monetary Costs Imposed by Sample Homicide Offenders.* Ames: Iowa State University.

Departmet, Los Angeles Police. 2016. *Crime Mapping and COMPSTAT.* Accessed June 29, 2016. http://www.lapdonline.org/crime_mapping_and_compstat.

DeWitt, Jeffrey S. 2015. *2015 DC Facts.* Government Release, Washington, District of Columbia: Office of the Chief Financial Officer.

Esri. 2016. *ArcGIS for Desktop.* Accessed June 5, 2016. http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/optimized-hot-spot-analysis.htm.

Esri. 2016. *Interpreting OLS results.* Accessed June 20, 2016. https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-statistics-toolbox/interpreting-ols-results.htm.

Esri. 2012. *Tutorial: Regression Analysis in ArcGIS.* May 9. Accessed June 5, 2016. http://www.arcgis.com/home/item.html?id=71a65d35688a4502b123cbdfc99afdee.

Felson, Marcus. 2006. "Routine Activity Theory." *Seminar: Crime's Everyday Routines.* Loughborugh University: Midlands Centre for Criminology and the Criminal Justice, February 2.

Fotheringham, Stewart A, Chris Brunsdon, Martin Charlton. 2002. " *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*.

Garofalo, James, Leslie Siegel, and John Laub. 1987. "School-Related Victimizations Among Adolescents: An Analysis of National Crime Survey Narratives." *Journal of Quantitative Criminology* 321-338.

Heuer, Jr, Richard J. 1999. *Psychology of Intelligence Analysis.* Center for the Study of Intelligence, Central Intelligence Agency.

Hipp, John R. 2010. "A Dynamic View of Neighborhoods: The Reciprocal Relationship Between Crime and Neighborhood Structural Characteristics." *Social Problems* 205-230.

Hollis, Meghan E, Marcus Felson, and Brandon C Welsh. 2013. "The Capable Guardian in Routine Activities Theory: A Theoretical and Conceptual Reappraisal." *Crime Prevention and Community Safety* 65-79.

Jones, Roderick W, and William Alex Pridemore. 2012. "The Foreclosure Crisis and Crime: Is Housing-Mortgage Stress Associated with Violent and Property Crime in US Metropolitan Areas." *Social Science Quarterly* 671-691.

Klevens, Joanne, Luis Fernando Duque, and Clemencia Ramirez. 2002. "The Victim-Perpetrator Overlap and Routine Activities." *Journal of Interpersonal Violence* 206-216.

Lauritsen, Janet L, and Kenna F Davis Quinet. 1995. "Repeat Victimization Among Adolescents and Young Adults." *Journal of Quantitative Criminology* 143-166.

Marx, Karl, and Frederick Engels. 1848. *The Communist Manifesto.* London: Communist League.

McCollister, Kathryn E, Michael T French, and Hai Fang. 2010. "The Cost of Crime to Society: New Crime-Specific Estimates for Policy and Program Evaluation." *Drug Alcohol Depend* (Drug and Alcohol Depend) 98-109.

Morenoff, Jeffrey D, Robert J Sampson, and Stephen W Raudenbush. 2001. "Neighborhood Inequality, Collective Efficacy, and the Dynamics of Urban Violence." *Criminology* 517-560.

Nicolaidis, Christiina, Mary Ann Curry, and Yvonne Ulrich. 2003. "Could We Have Known? A Qualitative Analysis of Data from Women Who Survived an Attempted Homicide by an Intimate Partner." *Journal of General Internal Medicine* 788-794.

Nofziger, Stacey, and Don Kurtz. 2005. "Violent Lives: A Lifestyle Model Linking Exposure to Juveniile Offending." *Journal of Research in Crime and Delinquincy* 3-26.

Perry, Walter L, Brian McInnis, and Carter C Price. 2013. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations.* Santa Monica: RAND.

Pratt, Travis C, and Christopher T Lowenkamp. 2002. "Conflict Theory, Economic Conditions, and Homicide." *Homicide Studies* 61 - 83.

Regoeczi, Wendy C, and John P Jarvis. 2011. "Beyond the Social Production of Homicide Rates: Extending Social Disorganization Theory to Explain Homicide Case Outcomes." *Justice Quarterly* 983-1014.

Reynald, Danielle M. 2015. "Environmental Design and Crime Events." *Journal of Contemporary Criminal Justice* 71-89.

Sampson, Robert J, and Stephen W Raudenbush. 2004. "Seeing Disorder: Neighborhood Stigma and the Social Construction of "Broken Windows". *Social Psychology Quarterly* 319 - 342.

Sampson, Robert J, and W Byron Groves. 1989. "Community Structure and Crime: Testing Social-Disorganization Theory." *American Journal of Sociology* 774-802.

Sampson, Robert J, Stephen W Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 918 - 924.

Scheider, Matthew C, Deborah L Spence, and Jessica Mansourian. 2012. *The Relationship Between Economic Conditions, Policing, and Crime Trends.* Washington: US Department of Justice.

Sherman, Lawrence W, Patrick R Gartin, and Michael E Buerger. 1989. "Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place." *Criminology* 27-55.

Siegel, Larry J. 2000. "Conflict Theory." *National Criminal Justice Reference Service* 254-284.

Spano, Richard, and Joshua D Freilich. 2009. "An Assessment of the Empirical Validity and Conceptualization of Individual Level Multivariate Studies of Lifestyle/routine Activities Theory Published from 1995 to 2005." *Journal of Criminal Justice* 305-314.

St. Jean, Peter K.B. 2007. *Pockets of Crime: Broken Windows, Collective Efficacy, and the Criminal Point of View.* Chicago: The University Chicago Press.

Taylor, Ralph B, and Adele V Harrell. 1996. *Physical Environment and Crime.* Research Report, Rockville: National Institute of Justice.

Tewskbury, Richard, and Elizabeth Ehrhardt Mustaine. 2003. "College Students' Lifestyles and Self-protective Behaviors: Further Considerations of the Guardianship Concept in Routine Activity Theory." *Crimininal Justice and Behavior* 302-327.

Tobler, W. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 234-240.

Tseloni, Adnromachi. 2000. "Personal Criminal Victimization in the United States: Fixed and Random Effects of Individual and Household Characteristics." *Journal of Quantitative Criminology* 415-442.

United States Department of Labor. 2016. *Burea of Labor Statistics.* Accessed March 27, 2016. http://www.bls.gov/eag/eag.dc.htm.

US Department of Justice–Federal Bureau of Investigation. 2012. *Crime in the United States, 2011*. September. Accessed September 11, 2016. https://ucr.fbi.gov/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/offenses-known-to-law-enforcement/expanded/expanded-homicide-data.

Vazsonyi, Alexander T, Lara M Belliston, and Dick Hessing. 2002. "Routine Activities and Deviant Behaviors: American, Dutch, Hungarian, and Swiss Youth." *Journal of Quantitative Criminology* 397-422.

Vilalta, Carlos, and Robert Muggah. 2014. "Violent Disorder in Ciudad Jaurez:: A Spatial Analysis of Homicide." *Trends in Organized Crime.*

Walker, Jefferey T. 2009. *Social Disorganization Theory in 21st Centruy Criminology: A Reference Handbook.* Thousand Oaks: SAGE Publications, Inc.

Wilson, John P, and Stewart A Fotheringham. 2008. *The Handbook of Geographic Information Science.* Malden: Blackwell Publishing.