Understanding Colombian Violence Through Geographic Information Systems and Statistical Approaches

Brandon Fowler
Western Kentucky University, brandon.cufflink@gmail.com

Follow this and additional works at: http://digitalcommons.wku.edu/theses
Part of the Geographic Information Sciences Commons, Latin American Studies Commons, and the Other Social and Behavioral Sciences Commons

Recommended Citation

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.
UNDERSTANDING COLOMBIAN VIOLENCE THROUGH GEOGRAPHIC INFORMATION SYSTEMS AND STATISTICAL APPROACHES

A Thesis
Presented to
The Faculty of the Department of Geography and Geology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Brandon Fowler

May 2013
UNDERSTANDING COLOMBIAN VIOLENCE THROUGH GEOGRAPHIC INFORMATION SYSTEMS AND STATISTICAL APPROACHES

Date Recommended December 6, 2012

Jun Yan, Director of Thesis

Margaret Gaspshover

David Keeling

Kevin Cary

Dean, Graduate Studies and Research Date

1-15-13
ACKNOWLEDGEMENTS

I want to extend my appreciation and gratitude to those who have continually given their patience and support throughout this project. First and foremost, I would like to acknowledge the Department of Geography and Geology at Western Kentucky University for their knowledge, experience, and passion for academic excellence. I want to thank my advisor, Dr. Jun Yan, for his patience, expertise, and relentless encouragement throughout the entirety of this thesis project. I am truly grateful for his optimism and commitment to ensure my success. I would also like to acknowledge Mr. Kevin Cary for providing some of the most educational and rewarding courses of my college career. While arduous and incredibly time-consuming, they instilled within me the diligence and academic discipline necessary to succeed in my future endeavors.

In addition, I want to extend my gratitude to Dr. Margaret Gripshover for graciously serving as a committee member for this thesis project. I would like to extend my gratitude to Dr. David Keeling for also serving as a committee member and for including me in the development of a GIS training program for the Bowman Expedition to Colombia. This partnership between Colombia and the American Geographical Society connected me with the GEOSCIRE research group and their efforts to improve the GIS capabilities of Colombia. It was this incredible opportunity that granted me a unique perspective of Colombian violence from the very people who live there. I want to thank all of them, especially David Aponte, Joan Mauricio Palacios, and Sebastián Díaz Angel for their knowledge, insight, and inspiration for this project.
Last but not least, I want to thank my parents, friends, and family for their continuous support and understanding throughout the course of this thesis project and my college career. None of my academic accomplishments would have been possible without these amazing people in my life.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xii</td>
</tr>
<tr>
<td>Chapter 1: Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 2: Literature Review</td>
<td>5</td>
</tr>
<tr>
<td>2.1 Violence and Crime Theories</td>
<td>5</td>
</tr>
<tr>
<td>2.2 Historical Perspective on a Contemporary Crisis in Colombia</td>
<td>10</td>
</tr>
<tr>
<td>2.2.1 A Summary of La Violencia</td>
<td>10</td>
</tr>
<tr>
<td>2.2.2 Insurgent Organizations</td>
<td>14</td>
</tr>
<tr>
<td>2.3 Cross-National Homicide Studies</td>
<td>18</td>
</tr>
<tr>
<td>2.4 Geographic Isolation and Crime</td>
<td>21</td>
</tr>
<tr>
<td>2.5 Studies of Colombian Violence</td>
<td>22</td>
</tr>
<tr>
<td>2.5.1 State Presence: The Illicit Drug Trade</td>
<td>24</td>
</tr>
<tr>
<td>2.5.2 Economic Conditions: Poverty and Inequality</td>
<td>25</td>
</tr>
<tr>
<td>2.5.3 Political Stability: Guerrilla Organizations</td>
<td>26</td>
</tr>
<tr>
<td>2.5.4 Social Strife: Displacement</td>
<td>28</td>
</tr>
<tr>
<td>2.6 Geographic Information Systems and Crime Studies</td>
<td>29</td>
</tr>
<tr>
<td>2.6.1 Crime Hotspot Detection and Spatial Autocorrelation</td>
<td>29</td>
</tr>
<tr>
<td>2.6.2 Regression Analysis and Geographic Factors of Crime</td>
<td>31</td>
</tr>
<tr>
<td>Chapter 3: Study Area, Data, and Methods</td>
<td>34</td>
</tr>
<tr>
<td>3.1 Overview of the Study Area</td>
<td>34</td>
</tr>
<tr>
<td>3.2 Data and Data Processing</td>
<td>44</td>
</tr>
</tbody>
</table>
3.2.1 Homicide Rates ................................................................. 45
3.2.2 Coca Crops ................................................................. 47
3.2.3 Political Participation ...................................................... 49
3.2.4 Displaced Persons ........................................................ 51
3.2.5 Poverty Rate ................................................................. 53
3.2.6 Standard of Living Index .............................................. 55
3.2.7 Terrain Ruggedness Index ............................................. 57
3.2.8 Armed Actions ............................................................... 67

3.3 Spatial Statistical Approaches ............................................ 72
3.3.1 Global Moran’s I ......................................................... 72
3.3.2 Anselin Local Moran’s I ................................................. 72
3.3.3 Regression Analysis and Spatial Data .......................... 73
3.3.4 Geographically Weighted Regression (GWR) ................ 76

3.4 GIS Implementation .......................................................... 78
3.4.1 Global Moran’s I ......................................................... 78
3.4.2 Local Moran’s I ............................................................ 81
3.4.3 Potential Regression Variables ..................................... 82
3.4.4 Transforming Regression Variables ............................. 85
3.4.5 Selecting Regression Variables .................................... 88
3.4.6 Ordinary Least Squares Regression ............................. 92
3.4.7 Geographically Weighted Regression ........................... 95

Chapter 4: Analysis and Results ............................................. 102
4.1 Scatter Plots ................................................................. 102
LIST OF FIGURES

Figure 2.1 International homicide rates per 100,000 people, 2004 ........................................ 7
Figure 2.2 ELN evolution sketch by the ELN commander Felipe Torres ......................... 17
Figure 3.1 Natural regions of Colombia ............................................................................ 36
Figure 3.2 Colombia Population Density and Major Cities, 2005 ..................................... 40
Figure 3.3 Administrative units of Colombia in 2005 .................................................. 41
Figure 3.4 Regional map of Colombia study area ......................................................... 43
Figure 3.5 Homicides per 100,000 people ........................................................................ 46
Figure 3.6 Coca cultivation density, 2005 ...................................................................... 48
Figure 3.7 Participation rate for the presidential election, 2006 ...................................... 50
Figure 3.8 Displaced persons per 100,000 people, 2005 ................................................. 52
Figure 3.9 Percentage of the population in poverty status, 2005 ..................................... 54
Figure 3.10 Colombia standard of living index, 2005 .................................................. 56
Figure 3.11 Example of how to obtain TRI values .......................................................... 58
Figure 3.12 Calculating the sum of the DEM values in a 3x3 neighborhood .................... 60
Figure 3.13 Calculating the squared value of the DEM ................................................... 61
Figure 3.14 Calculating the sum the squared DEM values in a 3x3 neighborhood .......... 62
Figure 3.15 Calculating the $R^2$ (squared TRI) value ................................................... 63
Figure 3.16 Calculating the square root to obtain the $r$ (TRI) values .............................. 64
Figure 3.17 Side-by-side depiction of original DEM and derived TRI rasters ................. 65
Figure 3.18 Terrain Ruggedness Index ........................................................................... 66
Figure 3.19 Armed actions initiated by the FARC per 100,000 people, 2005 ................. 69
Figure 3.20 Armed actions initiated by the ELN per 100,000 people, 2005 .................... 70
Figure 3.21 Armed actions initiated by the Public Force per 100,000 people, 2005 ...... 71
Figure 3.22 Spatial Autocorrelation (Global Moran’s I) tool................................. 79
Figure 3.23 Example of spatial autocorrelation report......................................... 80
Figure 3.24 Cluster and Outlier Analysis (Anselin Local Moran’s I) tool ............... 82
Figure 3.25 Scatterplot matrix for variables .......................................................... 86
Figure 3.26 Scatterplot matrix for transformed variables....................................... 88
Figure 3.27 Exploratory Regression tool ............................................................... 89
Figure 3.28 Ordinary Least Squares tool ............................................................... 93
Figure 3.29 OLS regression model illustration....................................................... 94
Figure 3.30 OLS residual example map................................................................. 95
Figure 3.31 Geographically Weighted Regression tool ........................................... 96
Figure 4.1 Scatter Plot – Coca cultivation density (2005)........................................ 102
Figure 4.2 Scatter Plot – Presidential Election Participation Rate (2006)............... 103
Figure 4.3 Scatter Plot – Displaced persons rate (2005) ......................................... 103
Figure 4.4 Scatter Plot – Standard of Living index (2005) .................................... 104
Figure 4.5 Scatter Plot – Terrain Ruggedness Index ............................................ 104
Figure 4.6 Scatter Plot – Rate of armed actions initiated by the FARC (2005) ....... 105
Figure 4.7 Scatter Plot – Rate of armed actions initiated by the Public Force (2005) . 105
Figure 4.8 Report generated by Spatial Autocorrelation (Moran’s I) tool.............. 107
Figure 4.9 Local Moran’s Index – Cluster / Outlier Type..................................... 109
Figure 4.10 OLS – Standard Residuals ................................................................. 111
Figure 4.11 OLS – Histogram of OLS Residuals....................................................... 114
Figure 4.12 Global Moran’s Index of OLS residuals ............................................. 115
Figure 4.13  Local Moran's Index of OLS residuals ............................................................... 116

Figure 4.14  Global Moran’s Index of GWR residuals .......................................................... 118

Figure 4.15  GWR – Local $R^2$ values ............................................................................... 120

Figure 4.16  GWR – Coca Density Coefficient ................................................................. 122

Figure 4.17  GWR – Presidential Election Participation Coefficient ................................. 123

Figure 4.18  GWR – Displaced Persons Coefficient ......................................................... 125

Figure 4.19  GWR – Standard of Living Index Coefficient .............................................. 126

Figure 4.20  GWR – Terrain Ruggedness Index Coefficient ........................................... 128

Figure 4.21  GWR – FARC Armed Action Coefficient .................................................... 129

Figure 4.22  GWR – Public Force Armed Actions Coefficient ......................................... 130

Figure 5.1  Distribution of Afro-Colombian and Indigenous populations ......................... 136

Figure 5.2  Territorial extent of the FARC, ELN, and paramilitaries, 2005 .................... 139
<table>
<thead>
<tr>
<th>Table 2.1</th>
<th>Summary of cross-national homicide studies</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.1</td>
<td>Categorical view of potential regression variables for this study</td>
<td>44</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>Categorical view of regression variables used in this study</td>
<td>92</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>Example of supplemental table of statistics from GWR tool</td>
<td>98</td>
</tr>
<tr>
<td>Table 3.4</td>
<td>Example of output feature class attribute table from GWR tool</td>
<td>101</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>OLS – Coefficients and Standard Errors</td>
<td>110</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>OLS – Diagnostic statistics</td>
<td>112</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>GWR – Supplemental diagnostic statistics</td>
<td>118</td>
</tr>
</tbody>
</table>
In 2002, Colombia had the highest homicide rate of any Latin American country (Berkman, 2007). The origins of this violence, however, are complex and difficult to identify. It would be sensible to argue that it cannot be explained by any one particular factor, but rather an assortment of many factors that wholly represent the social, economic, and political conditions of Colombia. By better understanding the origins of Colombian violence, policy makers can more effectively address and alleviate this prolonged issue. This study examines the geographic nature of municipal homicide rates for Colombia in 2005. The purpose of this study is to determine whether there are any discernible patterns in the geographic distribution of homicide rates across Colombia at the municipal level. It also aims to determine what combination of statistically significant predictors, if any, generates acceptable regression models for predicting the distribution of homicide rates. Spatial autocorrelation methods, particularly Global and Local Moran’s I statistics, were used to identify the clusters of high-value homicide rates. Regression models, specifically OLS and GWR, were utilized to examine the relationships between homicide rates and an assortment of geographic factors, including Coca Cultivation Density, Presidential Election Participation Rate, Displaced Persons Rate, Standard of Living Index, Terrain Ruggedness Index, FARC Armed Actions Rate, and Public Force Armed Actions Rate.
The results of this study indicate that clusters of high-value homicide rates were indeed located in the northern, southern, western, and central regions of Colombia. Among the aforementioned geographic factors, *Coca Cultivation Density*, *Displaced Persons Rate*, *Standard of Living Index*, *Terrain Ruggedness Index*, *FARC Armed Actions Rate*, and *Public Force Armed Actions Rate* all exhibited positive correlations. The variable exhibiting a negative correlation was the *Presidential Election Participation Rate*. 
CHAPTER 1: INTRODUCTION

The public perception of safety in many Latin American countries is less than desirable. Civil violence, guerilla warfare, and drug-trafficking are among many of the underlying tribulations that plague the region. Conventional wisdom recognizes that these geopolitical issues are anything but isolated and that their effects, on a global scale, are incredibly extensive. Among these countries, Colombia is infamous for its endless struggle to re-establish government control in the areas threatened by opposing resistance groups and unrestrained violence. A majority of its urban populations claim that violence is the single most cited problem facing the country (Moser and McIlwaine, 2000).

As of 2002, Colombia had the highest homicide rate of any Latin American country (Berkman, 2007). It had reached a level of violence typically only seen in those areas devastated by war. While the country experienced a dramatic decrease in these rates over the next several years, Colombia reported a staggering 18,111 homicides in 2005, which is approximately 42 homicides per 100,000 people (UNODC, 2011). Violence of this magnitude should be alleviated, but for that to occur, the violence and its spatial distribution must first be understood.

In an effort to better understand the origins of violence numerous studies have been conducted at the cross-national level to reveal the underlying influences of homicide rates around the world. The studies done by Messner and Rosenfeld (1997), LaFree and Kick (1986), Hansmann and Quigley (1982), and Fajnzylber et al. (2002) are just a few examples. Most of these studies have identified a variety of factors that express either positive or negative relationships with homicide rates, including factors related to
socioeconomic development, ethnic heterogeneity, income inequality, and other demographic characteristics of the population. These findings, however, may prove unreliable at more detailed geographic levels since regional conditions may be less pronounced at more aggregated scales, such as at the national level.

Additionally, several studies have focused primarily on Colombian violence and the underlying influences that are unique and specific to the country. Modern-day Colombia is often viewed as a simple product of the drug trade, a view, however simplistic, that is promoted by the U.S. media and informs U.S. government policy and support (Bergquist et al., 1992). Some scholars also draw this direct connection between drugs and violence within the country and believe that with the absence of drugs, violence will inexorably dissipate (Holmes et al., 2006). Others, however, believe that the principal components of Colombian violence are derived from an assortment of socioeconomic factors (Sánchez and Nuñez, 2001), social strife (Holmes et al., 2006), and political instability (Moser, 2000). Although these scholars identify different origins for the violence that continually ails Colombia, their collective research provides valuable insights into the workings of an incredibly intricate and delicate country.

These studies suggest that there may not be a single underlying factor for Colombian violence and that there are potentially many factors that are affecting the homicide rates and violence throughout the country. As a result, one would need to incorporate these factors into a systematic approach capable of better expressing the relationships among them. One such system, known as Geographic Information System (GIS), is well-suited and quite capable of performing this task. A GIS is able to facilitate data-driven decision-making by leveraging multiple sources of data that users can
analyze in a greater depth (Stoe et al., 2003, p. 12). It can utilize a database consisting of observations on spatially distributed features, activities, or events, which are definable in geographic space as points, lines, polygons, and raster datasets. Users of GIS can manipulate data about these points, lines, and areas to retrieve data for ad hoc queries and analyses (Duecker, 1979). It is therefore feasible to utilize a GIS to better understand Colombian homicide rates by incorporating the datasets of the most likely and significant underlying factors.

While some scholars have observed and analyzed Colombian violence at the departmental level, Holmes et al. (2006) for instance, the academic community appears to lack an in-depth analysis at a more localized level, such as the municipal level. This study intends to fill that void by analyzing Colombian violence at the municipal level, aiming to interpret and understand the relationships that exist between homicide rates and the underlying factors that may contribute to it. Fortunately, recent advances in spatial statistics, particularly local spatial statistical methods, have made this type of investigation possible. These local methods are capable of generating models that more accurately depict the dynamic relationships among spatial data at more localized scales. As a result, researchers are no longer restricted to global methods that do not account for the spatially varying nature of many geographic phenomena. In summary, the primary objectives of this thesis are the following:

- Objective I: To determine whether there are any discernible patterns in the geographic distribution of homicide rates across Colombia at the municipal
level. If founded, then the patterns would suggest that homicide rates are dependent upon certain underlying factors, varying across geographic space.

- Objective II: To determine what combination of statistically significant predictors, if any, generates acceptable regression models for predicting the distribution of homicide rates throughout Colombia.

To achieve these two objectives, various spatial statistical methods were adopted within a GIS context. Specifically for Objective I, the Global Moran’s I was utilized to determine if the distribution of homicide rates at the municipal level were spatially autocorrelated while Local Moran’s I was used to locate municipal clusters of high homicide rates, so-called crime hot spots. To address Objective II, two statistical regression models were developed in order to identify a set of statistically significant predictors at the national level and the municipal level respectively, a global Ordinary Least Square (OLS) model and a local spatial regression model based on geographically weighted regression (GWR).
CHAPTER 2: LITERATURE REVIEW

This chapter briefly discusses the various areas of relevance to this research. The first section defines violence and briefly examines its prevalence throughout Colombia and the rest of the world. It also presents several criminological theories in an effort to identify the social influences of criminal activities in general. Section 2.2 explains the historical, cultural, and socioeconomic nature of Colombia as a nation. Understanding these distinct, yet ever connected, aspects allows for a more holistic comprehension of modern-day Colombia and the violence that continually afflicts it. Section 2.3 reviews an assortment of cross-national studies of homicide rates and the independent variables that influence them. Section 2.4 then discusses how the geographic isolation of a region can affect the levels of criminal activity within it, which may provide valuable insight into the nature of the more disconnected and isolated areas of Colombia. Section 2.5 summarizes the previous studies of Colombian violence. These studies tend to focus on the illicit drug trade, poverty, income inequality, state presence, and displacement throughout Colombia. The last section finally reviews several crime studies that have incorporated GIS as a means of analyzing global and local patterns of crime.

2.1 Violence and Crime Theories

The term violence is somewhat nebulous but, nevertheless, it is firmly defined by the World Health Organization (WHO) as:
“The intentional use of physical force or power, threatened or actual, against oneself, another person, or against a group or community, that either results in or has a high likelihood of resulting in injury, death, psychological harm, maldevelopment or deprivation.” (WHO, 2002, p. 5)

With near certainty, one can assume that violence has always been a part of the human experience. It has plagued the human race in its various forms and continues to be the worldwide leading causes of death for people 15-44 years of age (WHO, 2002). For Colombia, a country engrossed with violence for nearly its entire history, there has never been a more pressing need than the outright abolition of political and cultural bitterness among its people.

The Geneva Declaration on Armed Violence and Development (2008, p. 5), a diplomatic initiative aimed at addressing the interrelations between armed violence and development, estimated that there were approximately 490,000 intentional homicides throughout the world in 2004. That translates to a global rate of approximately 7.6 intentional homicides per 100,000 inhabitants. As Figure 2.1 illustrates, several countries in southern Africa, Central America, and South America – including Colombia, El Salvador, Guatemala, Jamaica, South Africa, and Venezuela – reported among the highest recorded rates of violent death in the world. Between the years 2001-2007, Colombia alone endured over 153,000 homicides with an average homicide rate of 45 deaths per 100,000 inhabitants (Revista Criminalidad, 2008).
As an aspect of violence, the topic of homicide continually interests researchers as well as the public but unfortunately its underlying dynamics, motives, and very nature are often complex, confusing, and, at times, simply contradictory. Messner and Rosenfeld (1999, p. 27) reviewed key research on the structural sources of lethal violence. They found that the social-structure that influences homicide can be sub-divided into two basic types: (a) strain influences – structural forces that push or pressure persons into violence, and (b) control influences – structural weaknesses or ruptures that free or release people to engage in violence. They also pointed out that these acts of homicide are often observed as social events that are distributed in patterned ways as a result of major structural correlates – such as class, gender, race, age, etc.

In order to explain these patterns, several criminal theories have been developed to identify the social influences of criminal activities. Generally, these theories
incorporate one of the two aforementioned influence types to interpret the areas with relatively high crime rates. The following discussion briefly presents four major theories of crime, namely Anomie, General Strain, Social Disorganization, and Routine Activity. These crime theories have all been well-reviewed and summarized by Cullen and Agnew (2003).

1) **Anomie**: In 1938, Robert Merton introduced the paradigm of social structure and anomie that became the foundation for scores of research and theoretical discussions (Featherstone and Deflem, 2003, p. 471). Essentially, his theory of anomie refers to the circumstances where individuals are unable to achieve social success through traditionally acceptable means, thus creating higher crime rates (Merton, 1938).

2) **General Strain Theory**: The original strain theory had its roots in Merton’s anomie theory, which dealt with the social structures that might influence an individual to commit a crime. Later improved upon by Robert Agnew (1992), the general strain theory accounts for not only an individual’s inability to achieve success, but also the loss of positively valued stimuli and the introduction of negatively valued stimuli. This theory suggests that criminal activities are more likely to occur when negatively valued stimuli are introduced that result in anger and frustration (Agnew, 1992).

3) **Social Disorganization Theory**: The social disorganization theory, unlike the two aforementioned theories, focuses on communities as a whole rather than individuals. Sampson and Groves (1989, p. 777) clarified that social disorganization, in the most general sense, occurs when a community fails to maintain social controls and is incapable of enforcing the common values of its people. They also noted that this theory was originally pioneered by Shaw and McKay (1942) with their classic work
explained that the arguments made by Shaw and McKay characterized social
disorganization as being the result of three structural factors: low economic status,
ethnic heterogeneity, and residential mobility. According to this theory, these factors
are considered the catalysts for community disruption and criminal activities. Among
the major criminological theories, social disorganization, focusing on the interplays of
crime and place, appears to offer the best representation of Colombian violence from
geographic perspective. For instance, the political instability and social strife may be
the direct result of the economic status and the level of residential mobility
throughout the country.

4) **Routine Activity**: The routine activity theory focuses on the circumstances of space
and time rather than individuals. Developed by Cohen and Felson (1979), this theory
outlined an approach for analyzing and explaining crime rate trends as a byproduct of
changes in labor force participation and the number of single-adult households (1979,
p. 588). They claimed that direct-contact predatory violations, illegal acts that
intentionally take or damage the person or the property of another person, require the
convergence in space and time of motivated offenders, suitable targets, and the
absence of guardians capable of preventing the violation. Therefore, if any one of
these elements is missing, then it is unlikely that a successful direct-contact predatory
violation will occur (1979, p. 589). With respect to Colombia, there are likely
numerous instances where all three of these elements are absent, especially where
authority and control transition between state and guerrilla territories.
2.2  Historical Perspective on a Contemporary Crisis in Colombia

2.2.1  A Summary of La Violencia

Colombia has endured disparity, oppression, and wars for centuries. It is these series of events that laid the foundation for the 20th century era known as *La Violencia*, an epidemic of irregular clashes between liberals and conservatives where some 200,000 Colombians would lose their lives (Bailey, 1967, p. 562). Much of the conflict began between 1899 and 1902, when Colombia experienced its most intense period of violence known as “The War of the Thousand Days”. During this time, there were an estimated 100,000 deaths among a total population of 4 million. Shortly thereafter, Panamá separated and claimed independence as a consequence of the war. In the following years, Colombia maintained a relatively peaceful existence. Economic reconstruction ensued, minority party representation was guaranteed, coffee production and prices brought unprecedented growth, and bipartisan consensus was prevalent throughout the political regimes (Bergquist et al., 1992).

In the election of 1930, the conservative party lost its majority power to the liberal party, which initiated a period of social reform, economic development, and growing tensions between the parties. Over the next several years, tensions continued to grow until the 1946 presidential election. Jorge Gaitán split the liberal party when he ran against the official candidate. As a result, conservative moderate Mariano Ospina Pérez won the election by majority vote (Bergquist et al., 1992).

After Ospina Pérez won the presidency, conservatives replaced and persecuted liberals in many locales (Bergquist et al., 1992). On April 9, 1948, following the assassination of Gaitán, a civil war had erupted once again. The city of Bogotá was being
decimated as hordes of people attacked the presidential palace, demolished hundreds of
government and religious structures, and destroyed countless homes and stores. Soon
thereafter, President Pérez claimed that an agreement had been reached that would form a
bipartisan government to represent both liberal and conservative interests. This bipartisan
unity agreement fell apart in less than a year due to overwhelming partisanship and
violence that had once again consumed the political arena. The liberal insurrection
eventually resulted in their abstention from the 1950 presidential election, an attempt to
delegitimize the new president. Consequently, this brought to power the regime of
ultraconservative Laureano Gómez, who attempted to quell liberal insurgency through
various means of terror (Safford and Palacios, 2002).

In 1950, the strategy of the liberal resistance changed. They tried to organize
themselves as guerrillas and the character of the violence shifted from one of archaic
family vengeance to that of a more generalized civil war. Despite that, neither political
party publically endorsed the violence. To the detriment of their cause, however, the
existence of the guerrillas enabled the government to justify an indefinite prolonging of a
state of siege and a two-year recess of Congress. As Pope Pius XII aligned the Church to
be anti-communist, the conservatives won the support of many bishops, intensifying the
hostility between liberal and conservatives once again (Safford and Palacios, 2002).

A military coup d’état was accomplished by General Rojas Pinilla in 1953. After
assuming his de facto presidency, he found a country devastated by terror and a state of
generalized insurgency. The nation was on the verge of a social and political crisis and
the new government began to focus on three essential tasks: ending terror and anarchy,
demobilizing the guerrillas, and reconstructing the areas affected by violence.
The first task of Rojas Pinilla, to end the terror and anarchy, was essentially fulfilled by the immediate impact from the recent coup and the subsequent political shift. Terror was no longer a government strategy and the people began to speak of reconciliation instead of war. Rojas created a new atmosphere of tranquility and thousands of uprooted peasants made their way back home. The second task of Rojas Pinilla, to disarm and demobilize the guerrillas, was undertaken by offering them government aid and an unconditional amnesty. Once the offer was deemed genuine and sincere, thousands complied and there was relative peace for several months after the coup. The third task of Rojas Pinilla, reconstructing the areas affected by violence, was challenging to say the least. There were promises of agrarian reform and capitalist development. Financial resources were invested into two well-defined sectors. The first went towards certain types of agricultural investments, such as farm machinery and fertilizers, along with setting up irrigation zones. The second went towards financing infrastructure projects, such as hydroelectric plants, railways, the Eldorado International Airport in Bogotá, and the national television center. However, it became evident that the regime was not seeking to provide aid to those who had been ruined by the violence but rather to reward and stimulate those who were neither ruined nor even adversely affected by it (Bergquist et al., 1992).

Ultimately, the support for Rojas Pinilla began to collapse when the Colombian Intelligence Service used newly released killers to assassinate individuals who had been amnestied, demobilized guerrillas who had not turned over their arms, and defenseless peasants. Suddenly, guerrillas who had yet to lay down their arms were convinced that they needed to hold onto them and those who had surrendered their weapons had
regretted doing so. The atmosphere of tranquility was deemed an illusion and violence ensued once again. This presented an opportunity for the two traditional parties to regain their roles within the government. In 1957, Rojas Pinilla announced his intention to stay in power for an additional four years, but this attempt was thwarted by economic interest groups and the Church (Bergquist et al., 1992).

Following the departure of Rojas Pinilla, a liberal-conservative coalition government was instituted from 1958 to 1974 known as the National Front. This unique agreement to share power between the two parties was brought about through a combination of alternating the presidency every four years and the mandatory equal distribution of public positions between the three branches of government. Political forces outside of this bipartisan process were excluded (Bergquist et al., 1992).

With the advent of the National Front and its attempt to create a synthesis between the liberal republic and the neoconservative order, the traditional political violence had ended. What followed was the reemergence of violence from a criminal economic enterprise. Between 1954 and 1964, there was a period of mafia-type violence that grew out of the earlier years of partisan slaughter. This era of conflict differed in that it was no longer driven by political motives. Having shed much of the political forces, organized homicide had simply become a means of economic gain. Many of the armed bands that were linked to the political violence of the preceding years had established their own operations and their own rules (Safford and Palacios, 2002).

Initially, coffee that had been stolen from farms whose owners had fled could be sold by its managers and tenants without the involvement of the armed bands. Many landowners, for fear of being wiped out by their enemies, sought out the opposing gangs
to make various pacts. Farm owners became unwittingly intertwined with and locked into shady business deals. The conservative gangs were dedicated to dealings with stolen coffee while the liberal gangs were focused on the rustling of cattle. These *mafiosos*, however, did not stop there. They proceeded to purchase the farms from which they were stealing by subjecting the owners to extortion and forcing them to flee. This process became known as “conservatizing” the land, or later “re-liberalizing” it (Safford and Palacios, 2002).

The National Front became aware of the rising influence of the gangs within various localities and recognized this predatory economic violence and reasserted its power by presenting the armed bands and local political bosses with two choices: demobilization or confrontation with the state. Having become isolated from the political parties that they had supported, most of the bands demobilized and eventually disappeared (Safford and Palacios, 2002).

2.2.2 Insurgent Organizations

With the last of the armed bands having been wiped out by the military, a new period of violence was emerging. The period from 1961 to 1989 became known as the time of “armed conflict”. This was an era of insurrectional struggle by guerrilla groups whose aim was revolutionary transformation of the state. Today, the two major surviving organizations from this period are the *Fuerzas Armadas Revolucionarias de Colombia* (Revolutionary Armed Forces of Colombia – FARC) and the *Ejército de Liberación Nacional* (National Liberation Army – ELN) (Safford and Palacios, 2002). While extensive efforts have been put forth to eliminate these two guerrilla organizations, their
territorial expansion and the ensuing conflicts have become a perpetual military and political issue for decades.

Officially formed and founded in 1966, the FARC has been historically deemed as agrarian-communist. Its origins evolved from the defense organizations that peasants established in an effort to defend their parcels from the encroachments of larger landowners and the state. Initially, its primary focus was on peasants and their desire for land and various reforms that would create a society of small property owners (Richani, 2002). During its infancy, the organization was more defensive than offensive. It had no desire to engage in sabotage or terrorism and was devoid of any plans to act against local or national authorities. Its primary role was to simply protect peasant communities that would forego cherished rivalries with other peasant communities that were also protected by similar armed forces (Safford and Palacios, 2002).

The FARC and other self-defense organizations later became the target of military offensives after several right-wing politicians decried their existence. These attacks were ultimately counterproductive in that they transformed the relatively docile self-defense organizations into more mobile and offensive revolutionary guerrillas. The FARC eventually evolved into an organization with its own political and military doctrine. Consequently, with the arrival of narcotic trafficking into the regions of guerrilla influence, the FARC also became firmly linked to the cultivation of plants used in narcotic productions. Along with various economic conditions that encouraged the cultivation of coca, marijuana, and poppies, the narcotic industry moved its operations to zones that were predominately protected and influenced by the FARC (Safford and Palacios, 2002). As a result of its political, military, and economic doctrines, the FARC
have become the largest, best-trained and equipped, and most effective insurgent group in Colombia (DEA, 2006).

The ELN, considered Colombia’s second largest insurgent group (DEA, 2006), was founded in 1963 by a group of Colombian students, Catholics, and left-wing intellectuals who underwent military and political training in Cuba (Seib and Janbek, 2011). They claimed to protect the oppressed and the poor by defending them against the wealthier classes of Colombia, opposing the process of economic privatization, and impeding U.S. influences. Hoping to emulate Cuba’s Marxist revolution, the objective of the ELN was to overthrow the government and replace Colombia’s capitalist economy with a socialist system (Linden, 2004). Moreover, these founding members tended to follow the revolutionary doctrines of Che Guevara by moving to the countryside and establishing their focus among the peasantry.

The sketch in Figure 2.2, drawn by the ELN commander Felipe Torres, illustrates the three evolutionary phases of the ELN organization and how it intended to introduce and spread its ideology amongst the people. In the first phase (1965-1980), the ELN concentrated its efforts within the high mountains to further establish and build its network. The second phase (1980-1985) took place when the ELN started descending toward the small population centers. The third and final phase (1990-present) began when the ELN reached the medium-sized population centers along critical highways (Richani, 2002). The theory during this evolution was that their exposure to the peasantry, and experience of life among them, would transform the urban leaders into something truly revolutionary (Safford and Palacios, 2002).
Often resorting to terrorist acts of violence, the ELN seemed to give much more importance to guerrilla military tactics rather than political strategies (Safford and Palacios, 2002). Having established its operations primarily along Colombia’s northeastern border with Venezuela and in central and northwestern Colombia, the ELN employed traditional terrorist tactics to disrupt Colombia’s economic infrastructure and foreign economic presence. As opposed to the FARC, who would directly confront Colombian security forces, the ELN focused its terrorism and extortion efforts on oil drilling and mining companies, and on the families of kidnapped victims, primarily foreigners (DEA, 2002).

In reaction to these emerging guerrilla organizations in the 1960s, Colombia legalized civil defense groups to privately confront the guerrillas, protect land ownership, and support the army in counterinsurgency efforts. These paramilitary groups were often formed by landowners and drug traffickers and were essentially a means of offering
protection to the areas of the country with little or no state presence. While their initial aim was to provide protection for areas stricken by violence, their existence ironically intensified the levels of violence. These paramilitary groups were therefore declared illegal in 1989. (Holmes et al., 2006, p. 168).

2.3 Cross-National Homicide Studies

Over the last several decades, numerous cross-national studies have been conducted to identify the independent variables that most strongly correlate with homicide rates (Table 2.1). Their collective research, while sometimes conflicting, provides valuable insight for classifying the relationships of these independent variables as typically positive, negative, mixed, or insignificant. Some of the most frequently utilized independent variables include socioeconomic development, ethnic heterogeneity, income inequality, and population growth (LaFree, 1999, pp. 127-132).

In Messner and Rosenfeld’s (1997, p. 1405) cross-national analysis, their results indicate that socioeconomic development is negatively correlated with homicide rates. They noted that this relationship is consistent with the previous theories that urbanization and industrialization play a significant role in the decline of violent crimes. This negative relationship was also confirmed by LaFree and Kick (1986) in another study that aimed to identify the developmental, distributional, and demographic variables associated with crime.

Ethnic heterogeneity is another independent variable that is widely incorporated into cross-national studies as well. Hansmann and Quigley (1982, p. 217), for example, found that there is a consistently significant and positive relationship between homicide
rates and ethnic heterogeneity. The general consensus of Avison and Loring (1986) also found that homicide rates are influenced by income inequality and that the degree of influence is often related to the intensity of ethnic heterogeneity.

In several other cross-national studies, models have included income inequality as an independent variable for predicting and explaining homicide rates. Krahn et al. (1986) found that income inequality expresses a positive correlation with homicide rates, especially among countries that are considered to be more democratic in nature. Fajnzylber et al. (2002, pp. 25-26) found that, in general, income inequality has a positive and significant effect on crime rates, whether it be robbery or homicide. Their study reveals that greater inequality is associated with higher rates of lethal violence and that those rates decrease as economic growth improves. They concluded that faster poverty reduction leads to an overall decline in national crime rates. Another study conducted by Kelly (2000, p. 537) reached a similar conclusion, suggesting that violent crime was largely impacted by income inequality, even after controlling for various factors including the effects of poverty, race, and family composition.

LaFree (1999, p. 142) discovered that population growth is a common independent variable that is included among many cross-national studies. Multivariate analyses conducted by Krahn et al. (1986) and LaFree and Kick (1986), for example, revealed that increases in population growth are consistently associated with higher rates of homicide. LaFree also noted that it is one of the few independent variables that is absent of contradicting evidence. This and many other independent variables have been incorporated into numerous cross-national studies (Table 2.1) that have been collected and summarized by LaFree (1999, pp. 127-132).
<table>
<thead>
<tr>
<th>Study</th>
<th>Data Source</th>
<th>Sample Size</th>
<th>Type Analysis</th>
<th>Independent Variables</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>McDonald</td>
<td>Interpol</td>
<td>40</td>
<td>OLS Regression</td>
<td>GNP, per capita</td>
<td>-</td>
</tr>
<tr>
<td>1976</td>
<td></td>
<td></td>
<td></td>
<td>Governmental stability +</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intercountry inequality +</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Linguistic heterogeneity +</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population increases +</td>
<td></td>
</tr>
<tr>
<td>Messner</td>
<td>Interpol</td>
<td>39</td>
<td>OLS Regression</td>
<td>Income inequality</td>
<td>+</td>
</tr>
<tr>
<td>1980</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hagemann &amp; Quigley</td>
<td>U.N.</td>
<td>38-40</td>
<td>OLS Regression</td>
<td>Ethnic heterogeneity</td>
<td>+</td>
</tr>
<tr>
<td>1982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Income inequality</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Linguistic heterogeneity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population (log)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Youth</td>
<td>+</td>
</tr>
<tr>
<td>Messner</td>
<td>Interpol</td>
<td>39</td>
<td>OLS Regression</td>
<td>Income inequality</td>
<td>-</td>
</tr>
<tr>
<td>1982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population increase</td>
<td>+</td>
</tr>
<tr>
<td>Coalin &amp; Simpson</td>
<td>WHO</td>
<td>52</td>
<td>OLS Regression</td>
<td>Infant mortality rate (log)</td>
<td>+</td>
</tr>
<tr>
<td>1985</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Percentage males aged 15-29 (log)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Persons per square mile (log)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population (log)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Percentage urban (log)</td>
<td>-</td>
</tr>
<tr>
<td>Kick &amp; LaFree</td>
<td></td>
<td>40</td>
<td>OLS Regression</td>
<td>Development index</td>
<td>-</td>
</tr>
<tr>
<td>1983</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Income inequality</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Persons per household</td>
<td>+</td>
</tr>
<tr>
<td>Atkinson &amp; Loring</td>
<td>WHO</td>
<td>32-27</td>
<td>OLS Regression</td>
<td>Ethnic heterogeneity (a.h.)</td>
<td>+</td>
</tr>
<tr>
<td>1986</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Income inequality (a.h.)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Labor force participation, males</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(i.i) x (a.h.)</td>
<td>+</td>
</tr>
<tr>
<td>Krivok, Barrengiel,</td>
<td>Interpol &amp;</td>
<td>65</td>
<td>OLS Regression</td>
<td>Democracy index</td>
<td>+</td>
</tr>
<tr>
<td>&amp; Garrett</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td></td>
<td></td>
<td></td>
<td>Income inequality</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population growth rate</td>
<td>+</td>
</tr>
<tr>
<td>LaFree &amp; Kick</td>
<td></td>
<td>47</td>
<td>OLS Regression</td>
<td>Economic development index</td>
<td>-</td>
</tr>
<tr>
<td>1986</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Income inequality</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population growth</td>
<td>+</td>
</tr>
<tr>
<td>Messner</td>
<td>Interpol /</td>
<td>52-32</td>
<td>OLS Regression</td>
<td>Economic discrimination</td>
<td>+</td>
</tr>
<tr>
<td>1989</td>
<td>WHO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Income inequality</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population less than 15 years old</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Percentage urban</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2.1: Summary of cross-national homicide studies. *Note: R represents a positive (+) or negative correlation (-) correlation with homicide rates. Source: LaFree (1999)*
<table>
<thead>
<tr>
<th>Study</th>
<th>Data Source</th>
<th>Sample Size</th>
<th>Type Analysis</th>
<th>Independent Variables</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sieber</td>
<td>Interpol</td>
<td>44</td>
<td>Bivariate</td>
<td>Infant mortality rate (log)</td>
<td>+</td>
</tr>
<tr>
<td>1990</td>
<td></td>
<td></td>
<td></td>
<td>Newspapers/1,000 population</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population change</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population/hospital beds</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population/physicians</td>
<td>+</td>
</tr>
<tr>
<td>Bennett</td>
<td>Interpol</td>
<td>38</td>
<td>Time series</td>
<td>Juvenile proportion</td>
<td>-</td>
</tr>
<tr>
<td>1991a</td>
<td></td>
<td></td>
<td></td>
<td>GDP manufacturing / agricultural</td>
<td>+</td>
</tr>
<tr>
<td>Bennett</td>
<td>Interpol</td>
<td>43</td>
<td>Time series</td>
<td>Educational inequality</td>
<td>-</td>
</tr>
<tr>
<td>1991b</td>
<td></td>
<td></td>
<td></td>
<td>Percentage urban</td>
<td>+</td>
</tr>
<tr>
<td>Nespoulia</td>
<td>Interpol /</td>
<td>64-106</td>
<td>OLS Regression</td>
<td>Economic discrimination</td>
<td>+</td>
</tr>
<tr>
<td>1994</td>
<td>WHO</td>
<td></td>
<td></td>
<td>GDP, per capita (log)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Income inequality</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Latin American countries</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population size (log)</td>
<td>-</td>
</tr>
<tr>
<td>Nespoulia</td>
<td>Interpol /</td>
<td>58-105</td>
<td>OLS Regression</td>
<td>GDP, per capita</td>
<td>-</td>
</tr>
<tr>
<td>1996</td>
<td>WHO</td>
<td></td>
<td></td>
<td>Income inequality</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Population size (log)</td>
<td>-</td>
</tr>
<tr>
<td>Messner &amp; Rosenfeld</td>
<td>WHO</td>
<td>38-45</td>
<td>OLS Regression</td>
<td>Decommitt,ion index</td>
<td>-</td>
</tr>
<tr>
<td>1997</td>
<td></td>
<td></td>
<td></td>
<td>Development index</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Economic discrimination index</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sex ratio (log)</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2.1 (cont’d): Summary of cross-national homicide studies. Note: *R represents a positive (+) or negative correlation (-) with homicide rates.* Source: LaFree (1999)

2.4 Geographic Isolation and Crime

There are also some studies that specifically look at geographic isolation and how it affects regional crime rates. Fafchamps and Moser (2002, p. 2), for example, found instances where crime rates were actually higher in more isolated, less populated areas. They asserted that crime within Madagascar is negatively correlated with population density and positively correlated with isolation, even after controlling for other risk factors. They found that all four categories of criminal activity – cattle theft, burglary,
homicide, and rape – were more prevalent in the areas with low population density and a lack of relative proximity to traditional towns. Additionally, they noted that the isolated areas of Madagascar appear to “nurture distrust among different groups and provide safe harbor and passage for criminals” (p. 23).

The increased isolation within Madagascar can be explained in part by a temporary state of poverty that prevented travel from one area to another, leaving many households unable to feed themselves. It was during this time that there was an increase in crop theft from the more isolated and rural areas of the country. Although other forms of crime, such as burglaries and homicides, showed a less systematic relationship with temporary poverty, this is an example of how behavioral changes can occur in response to the absence of goods, services, and opportunities (Fafchamps and Minten, 2005, p. 1).

While the circumstances of Madagascar and Colombia are very different, they may both experience instances where the geographic isolation of a region directly affects occurrence of crime. The leftist guerrilla organizations of Colombia, for instance, took it upon themselves in the mid-1960s to provide a presence in geographic areas where the state had formerly been absent. As a result, these organizations discovered that the remote isolation of the jungle and mountain terrain served as a sanctuary for their organization to develop and grow (Maddaloni, 2009, p. 25).

2.5 Studies of Colombian Violence

In general, the extent to which countries in Central and South America are affected by crime and violence varies significantly. In 2004, the United Nation Office on Drugs and Crime (UNODC) reported that homicide rates per 100,000 were 56.4 in El
Salvador, 45.5 in Colombia, 32.5 in Venezuela, 26.2 in Brazil, 3.7 in Bolivia, and 3.0 in Perú (UNODC, 2004). While much of the violence in these countries takes place in urban settings, a high incidence of rural violence commonly occurs in conflict and post-conflict countries including El Salvador, Guatemala, and Colombia (Heinemann and Verner, 2006, p. 3).

In an effort to better understand the violence in Colombia, Moser (2000, p. 2) developed a framework that distinguishes between political, economic, and social acts of violence. She provided the definition of political violence as “acts motivated by a desire, conscious or unconscious, to obtain or maintain political power” (p. 3). These acts can be manifested by armed conflict between guerrillas, paramilitaries, or any other political party. She then defined economic violence as “violent acts motivated by a desire, conscious or unconscious, for economic gain or to obtain or maintain economic power” (p. 3). These acts can be manifested through street crime, carjacking, robbery, drug trafficking, and kidnapping. Moser defined social violence as “violent acts motivated by a desire, conscious or unconscious, for social gain or to obtain or maintain social power” (p. 3). These acts can be manifested as interpersonal violence such as spouse and child abuse, sexual assault of women and children, and arguments that simply get out of control. According to Moser and McIlwaine (2000, p. 2), Colombian focus groups identified the proportions of violence between these three categories, with economic violence at 54%, social violence at 32%, and political violence at 14%. Their investigation therefore suggests that the people tend to experience economic violence far more frequently than social or political violence.
Various scholars have tended to focus their investigations around these three categories in an effort to elucidate the social, economic, and political factors that continually drive Colombian violence. While some have argued that there is a direct connection between violence and the illicit drug trade, others have focused on state presence, economic conditions, political stability, and social strife. These various approaches are briefly reviewed in the following sections.

2.5.1 State Presence: The Illicit Drug Trade

Rabasa and Chalk (2001) claimed that the current insatiability in Colombia is derived from the “interaction and resulting synergies of an underground criminal drug economy and the growth of armed challenges to the state’s authority” (p. 1). They emphasized that there are two sources of instability. The first source of instability is the increasingly pervasive influence of the drug trade and their infiltration of societal institutions. The second source of instability is the growth of guerrilla armies, paramilitaries, and privatized security forces. Consequently, they believed that the confluence of drug traffickers, guerrillas, and paramilitaries have exacerbated the loss of authoritative, economic, and social control.

Dion and Russler (2008, pp. 401-402) conducted a study that models the sub-national pattern of coca cultivation in Colombia after the implementation of Plan Colombia (2001-2005). They noticed that coca cultivation tends to occur in the areas with adequate labor and land resources but an insufficient public infrastructure and state presence. These isolated regions of the country, according to Sanchez et al. (2003, p. 376), receive little funding for infrastructure, technical assistance, education, and health
services. As a result, legal and viable economic opportunities are limited, a circumstance that highlights the importance of an actively involved and invested national government.

Moser (2000, pp. 5-11) noted that the illicit drug trade of Colombia is a critical source of economic and political violence. As one of the three most important coca cultivation countries in Latin America, Colombia has different types of drug-related violence that varies in both significance and intensity. The first type is violence that is linked to organized crime that is associated with drug production and trafficking. This includes violent attacks to prevent aerial eradication, gang-related killings, and assassinations of prominent political and judicial figures. The second type of violence is linked to the militarization of the fight against drugs, which has re-legitimized the domestic role of military forces and has blurred the distinction between the roles of military and police. The third type is considered disorganized violence, which occurs around distribution and consumption when drug addicts need money for their habit or street dealers are challenged for their profits.

2.5.2 Economic Conditions: Poverty and Inequality

A large body of contemporary research has examined the relationship between homicide and two particularly important features of economic stratification: poverty and income inequality. According to Messner and Rosenfeld (1999, p. 29), “poverty is generally understood to represent economic deprivation”. They stated that “it refers to a situation wherein persons have difficulty securing the basic necessities for a healthy life”. It can therefore sometimes be the strain that pressures people to commit acts of lethal
violence. According to their findings, poverty, however, is merely a representation of income levels and not necessarily indicative of economic stratification.

Sánchez and Nuñez (2001, p. 25) produced an investigation of 711 municipalities to identify the origins of criminal activity in Colombia. Their results indicated that a relationship exists between homicide rates and socioeconomic factors such as poverty and inequality. They found that between 3% and 13% of the differences in homicide rates between the most and least violent municipalities could be explained by socioeconomic factors. The remaining differences were explained by the presence of armed groups (paramilitaries, guerrillas), the inefficiency of the justice system, the intensity of drug trafficking, and the interaction between armed groups and drug trafficking.

A study by Londoño and Guerrero (1999, pp. 38-39) revealed that throughout the 1990’s, the level of violence in Colombia had been three times greater than the rest of Latin America. Their investigation also found that income inequality and poverty were significantly and positively correlated with homicidal violence. Even after controlling for the effect of income and the education gap, income inequality and poverty remained statistically significant. Their assertion was that the incidence of homicide was greater in low-income population areas than in middle- and high-income population areas. Additionally, they discovered that an increase of 1 point in the proportion of poor was associated with an additional 3,186 murders per year.

2.5.3 Political Stability: Guerrilla Organizations

Colombia has historically lacked a uniform authoritative presence throughout the country. A study produced by Fearon and Laitin (2003, pp. 75-76) argued that
financially, organizationally, and politically weak governments are more apt to render insurrection. In these conditions, the feasibility and attractiveness of insurgency is “due to weak local policing or inept and corrupt counterinsurgency practices”. They also noted the importance of a government presence in remote areas, claiming that “insurgents are better able to survive and prosper if the government and military that they oppose are relatively weak – badly financed, organizationally inept, corrupt, politically divided, and poorly informed about goings-on at the local level” (Fearon and Laitin, 2003, p. 80).

As noted by Waldmann (2007, p. 73), the state is “hardly present” in many rural areas of Colombia. Consequently, the support for guerrilla organizations in these regions is often based on their ability to provide basic order in the areas that do not have a significant government presence. As a result, these organizations have grown particularly strong in the areas that exhibit a vacuum of authority (Fajardo, 2003, p. 30). In order to fund their efforts, however, they have turned to a variety of sources over last several decades, most notably the illicit drug trade. Enchandia (2000, p. 121) asserted that while guerrilla organizations are often funded by the drug trade, they do not directly profit from it. In short, they are simply external actors that provide the means to process and export the finished goods. This revenue is essentially a tax for providing their services and resources. Therefore, as a source of financial security and strength for guerrilla organizations, there may be a relationship between armed groups, coca cultivation, and violence. Fortunately, Holmes et al. (2006, p. 158) conducted a department level analysis of this relationship which suggests that coca cultivation was not a major factor in explaining the differences in guerrilla violence throughout Colombia. They did, however,
find a persistent, significant, and positive relationship between guerrilla violence and coca eradication.

2.5.4 Social Strife: Displacement

The escalation of internal armed conflict and violence has been linked to the creation of a large population of internal refugees, also known as displacement (Holmes et al., 2006, p. 159). This forced displacement has covered nearly every region of Colombia and affects disproportionately vulnerable groups of the population (Ibáñez and Vélez, 2003, p. 1). Ultimately, the displacement has been the result of political conflicts between guerrilla organizations, paramilitaries, and the state. Crimes directed at the civil population, however, have been a low cost and effective strategy for illegally armed groups to expand their territories. In turn, these groups are better able to strengthen their area of control, transport weapons, and further develop their illegal activities (Ibáñez and Vélez, 2003, p. 2).

According to Holmes et al. (2006, pp. 161, 176, 179), there is a strong, positive relationship between displacement, coca production, coca eradication, and leftist guerrilla violence in Colombia. They pointed out that an influx of displaced people may have taxed an already overburdened and weakened state, further facilitating insurgent activity. Likewise, Byman et al. (2001, p. 66) claimed that insurgents often seek out and dominate refugee camps “because they are well armed and organized, while the displaced population is weak and disorganized; in addition, there may be no government or aid agency capable of imposing order”. Under these circumstances, rebel groups can demand money, provisions, or recruits from displaced populations with relative ease.
Ibáñez and Kirchhoff (2001, p. 41) conducted a study to better identify the determinants and effects of displacement in Colombia. They found that violence and the perceptions of insecurity are significant factors for motivating displacement. According to their results, land owners, members of local organization, and younger household heads face the highest risk of targeted threats, which appear to be the most important trigger of displacement.

2.6 Geographic Information Systems and Crime Studies

The applications of Geographic Information Systems (GIS) are as diverse as they are extensive. Having emerged as a ubiquitous means to analyzing geospatial data, GISs have become an all-pervasive technology used for in-car navigation, construction, weather forecasting, military planning, and other geographically oriented fields. Crime analyses, in particular, benefit from GISs in that they provide an invaluable context for spatial methods capable of recognizing and mapping patterns of crime. Law enforcement personnel can then investigate these patterns to actively reduce and/or prevent crime, whether it is theft, assault, rape, homicide, etc. For that reason, the identification of these crime hotspots is an important step in understanding the mechanisms and dynamics that contribute to their fruition (Chainey and Ratcliffe, 2005, pp. 2-4).

2.6.1 Crime Hotspot Detection and Spatial Autocorrelation

Crime, whether violent or non-violent, is not spread consistently across geographic space. It tends to be clustered in some areas while being relatively absent in others. These clustered areas are often referred to as crime hotspots by researchers and law
enforcement. Though no common definition of hotspot exists, the general understanding is that a hotspot is an area with an identifiable boundary that has a “greater than average number of criminal or disorder events, or an area where people have a higher than average risk of victimization”. This also suggests the presence of crime cool spots – “places or areas with less than the average amount of crime or disorder” (Eck et al., 2005, p. 2). There are generally two categories of spatial statistical methods that can be useful for identifying crime hot spots: 1) methods of spatial point pattern analysis (SPPA) for examining the distribution of the actual crime occurrences, often in point locations, across geographic space; and 2) methods for detecting the spatial autocorrelation effects on crime rates. The latter category is more relevant to this study as point location data was unavailable for Colombian homicides.

According to Griffith (2003, p. 3), spatial autocorrelation is “the correlation among values of a single variable strictly attributable to their relatively close locational positions on a two-dimension surface”. This phenomenon therefore violates the independent observation assumption of classical statistics. He then noted that positive spatial autocorrelation is an indication of geographical areas that tend to express similar values. These areas are comprised of high values located near high values, medium values located near medium values, and low values located near low values (Griffith, 2003, p. 5). Demographic and socio-economic variables are often prime examples of variables exhibiting positive spatial autocorrelation. When these values are positive, they represent the spatial clustering of high or low values for a particular area. In contrast, when these values are negative, they represent spatial dispersion of high values surrounded by low values and vice versa (Griffith, 2003, p. 5).
Messner et al. (1999) conducted a study to examine the distribution of homicides in the 78 counties around the St. Louis metropolitan area between 1984 and 1988 and also between 1988 and 1993. They utilized the global Moran’s I statistic to provide “an indication of the extent to which the spatial pattern of the whole data set is compatible with a null hypothesis of randomness” (p. 439). In addition, their study also incorporated the local version of Moran’s I to categorize the patterns of spatial association into several descriptive categories. They reported a positive and statistically significant global Moran’s I, confirming that the homicide rates of the counties were indeed clustered into hot spots and cool spots throughout the study area. The local Moran’s I then indicated that the high homicide rate counties were clustered around the St. Louis urban core and that the low homicide rate counties were clustered in the territories that were far removed from the core. Messner et al. (1999) concluded that these exploratory data methods are extremely useful for investigating spatial autocorrelation and enhancing our understanding of the phenomenon of homicide.

2.6.2 Regression Analysis and Geographic Factors of Crime

Regression analysis has been widely used by geographers to identify geographic factors of many geographic phenomena, including crime studies. Paul and Patricia Brantingham (1981, p. 8) described a crime as having four dimensions: legal, offender, target, and place. Most notably, they explained that the dimension of place is a “discrete location in time and space at which the other three dimensions intersect and a criminal event occurs”. It is this particular aspect of crime that has attracted and established the

Recent technological and analytical advances have made GIS a crucial component of contemporary spatial studies. Researchers are now able to process large volumes of data in a shorter period of time and with greater precision. Analytical processes that once took an inordinate amount of time are now quickly and easily executed by computers within minutes or even seconds (Lee and Wong, 2001, p. vii). As a result, there has been a growing interest in GIS as a means for helping explain the occurrence of criminal activities (Murray et al., 2001, p. 309).

In a study conducted by Graif and Sampson (2009), they employed a “GIS analytical framework” (p. 246) to examine the relationship between heterogeneity and homicide rates of Chicago neighborhoods. Their approach incorporated Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR) models with several independent variables including socioeconomic disadvantage, residential stability, population density, percentage foreign born, language diversity, and spatial dependence (pp. 248-250). According to their results, the GWR model offered a better fit and explained more of the data variance than the OLS model, suggesting that the GWR model was able to account for localized spatial variability (p. 257).

Troy et al. (2012) also utilized a GIS by examining the relationship between indicators of crime and measures of tree cover in Baltimore City and Baltimore County, Maryland. One of the goals of their study was “to understand how the relationship between trees and crime might vary across space in a way that cannot adequately be controlled for in a linear regression model” (p. 263). Their OLS regression models
incorporated an assortment of independent variables including land cover percentages, population density, median household income, and other local demographics (p. 264). None of the OLS regression models, however, were able to show the spatially non-constant relationship that may have existed between the independent variables and the dependent variable. For this reason, they generated a GWR model that identified several parameters expressing this spatially varying relationship. By mapping these relationships, they were better able to reveal patterns of crime that were the result of dissimilar circumstances (p. 266).

Cahill and Mulligan (2007) studied crime patterns with a GIS as well. Their research, while admittedly not unique, examined the covariates of violent crime in Portland, Oregon. Their study emphasized the possibility of spatial variation in crime measures and their covariates by presenting a local analysis of crime using GWR and comparing the results to a global OLS model (p. 175). Ultimately, their results demonstrated the effectiveness of GWR in the context of providing insight into the spatial patterns of data. They found that the local GWR model was better suited than the global OLS model for exploring the varying relationships between independent and dependent variables. Cahill and Mulligan note that a local analysis is therefore more appropriate for developing crime intervention methods that are varied across space as well (p. 192).
CHAPTER 3: STUDY AREA, DATA, AND METHODS

This chapter discusses the geographic settings, methods, and datasets used in this study. The first section presents an overview of Colombian and then discusses the process of how the study units were selected. It also depicts the physical nature of the country and the distribution of its general population. Section 3.2 then describes the datasets used in this study, including the data sources and the data processing steps. After that, Section 3.3 introduces the main spatial statistical methods utilized in this study, including spatial autocorrelative methods to identify spatial clusters (crime hotspots) of homicide rates and the regression methods for identifying a set of significant predictors of homicide rates. The last section then details how these methods were implemented in GIS for analyzing Colombian homicide rates in 2005.

3.1 Overview of the Study Area

Colombia is located in the north-west region of South America. It is bordered by Panamá and the Caribbean Sea to the north, by Venezuela and Brazil to the east, by Ecuador and Perú to the south, and by the Pacific Ocean to the west. Its geographic center is located at 4°0' N, 72°0'W (CIA, 2011). With its rugged topography and physical structure, Colombia is traditionally divided into six distinct natural regions (Figure 3.1). These regions are comprised of: 1) the Andina (Andean) region, covering the three principal mountain ranges; 2) the Llanura Del Caribe (Caribbean) region, covering the area adjacent to the Caribbean Sea; 3) the Pacífico (Pacific) region, lying between the border of Panamá and Ecuador, bounded by the Pacific Ocean and the Cordillera
Occidental; 4) the Orinoquía region, a vast area east of the Cordillera Oriental; 5) the Amazonia (Amazon) region, located in Southern Colombia, bordered by the Guaviare River to its North; and 6) the Insular region, areas outside of the continental territories of Colombia that are located in both the Caribbean Sea and the Pacific Ocean (Encyclopedia Britannica, 2010).
Figure 3.1: Natural regions of Colombia, Source: IGAC (2002)
**REGIONES NATURALES**

<table>
<thead>
<tr>
<th>Región Andina</th>
<th></th>
<th>Región del Pacífico</th>
<th></th>
<th>Región de la Orinoquia</th>
<th></th>
<th>Región de la Amazonia</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>Nudo de los Pastos</td>
<td>18</td>
<td>Sector norte de las estribaciones occidentales Cordillera Occidental</td>
<td>36</td>
<td>Piedemonte Llanero</td>
<td>43</td>
<td>Piedemonte Amazoniaco</td>
</tr>
<tr>
<td>17</td>
<td>Sector nororiental de la Cordillera Occidental</td>
<td>17</td>
<td>Sector sur de las estribaciones occidentales de la Cordillera Occidental</td>
<td>37</td>
<td>Llanuras de destornillo del Piedemonte</td>
<td>44</td>
<td>Llanuras altas y desérticas del río Caquetá</td>
</tr>
<tr>
<td>16</td>
<td>Fosa del Ota</td>
<td>16</td>
<td>Altiplano de Popayán</td>
<td>38</td>
<td>Llanuras del río Meta</td>
<td>45</td>
<td>Llanuras de los ríos Guaviare - Ñindia</td>
</tr>
<tr>
<td>15</td>
<td>Valle del Cauca</td>
<td>15</td>
<td>Valle del Cauca</td>
<td>39</td>
<td>Llanuras del río Orinoco</td>
<td>46</td>
<td>Confluencia de los ríos Putumayo y Caquetá</td>
</tr>
<tr>
<td>14</td>
<td>Cañón del Cauca</td>
<td>14</td>
<td>Cañón del Cauca</td>
<td>40</td>
<td>Llanuras de los ríos Guaviare - Ñindia</td>
<td>47</td>
<td>Penínsulas al sur de Puerto Ñindia</td>
</tr>
<tr>
<td>13</td>
<td>Medio de los Ríos de las Montañas</td>
<td>13</td>
<td>Medio de los Ríos de las Montañas</td>
<td>48</td>
<td>Llanuras entre los ríos Ñindia - Yari</td>
<td>49</td>
<td>Amazonía meridional</td>
</tr>
<tr>
<td>12</td>
<td>Cordillera Central meridional</td>
<td>12</td>
<td>Cordillera Central meridional</td>
<td>50</td>
<td>Llanuras de los ríos Ñindia - Putumayo</td>
<td>51</td>
<td>Confluencia de los ríos Apaporis - Caquetá</td>
</tr>
<tr>
<td>11</td>
<td>Medio del Tairona</td>
<td>11</td>
<td>Medio del Tairona</td>
<td>52</td>
<td>Cordillera de los Andes</td>
<td>53</td>
<td>Semillas y montes Ñindia</td>
</tr>
<tr>
<td>10</td>
<td>Medio de los Ríos de las Montañas</td>
<td>10</td>
<td>Medio de los Ríos de las Montañas</td>
<td>54</td>
<td>Llanuras de destornillo - confluye Guaviare - Ñindia en el río Ñindia</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Límite de región**

Figure 3.1 (cont’d): Natural regions of Colombia, Source: IGAC (2002)
A majority of the country’s land area is made up of lowland plains that consist of either tropical grasses or Amazonian forest. These plains are appropriately called *tierra caliente*, or “hot lands”. As the elevation increases, however, the average temperature declines and the surrounding environments change from hot river valleys to snow-topped mountain peaks. Amongst the mountain ranges, one can find a string of basins and plateaus, at an elevation between 1,500 - 3,000 meters above mean sea level (AMSL) that boast moderate temperatures and often some of the best soils and living conditions in the country. As a result, these middle-elevation areas have held Colombia’s densest concentration of human inhabitants for centuries (Bushnell, 1993, p. 1).

Out of all the geographic features in Colombia, no one has so molded the history of Colombia as the Andes Mountains. These mountains are separated into three principal ranges: 1) the Cordillera Occidental, between the Pacific Ocean and the valley of the Cauca River; 2) the Cordillera Central, between the Cauca and the Magdalena River; and 3) the Cordillera Oriental, which branches off toward Venezuela. These three ranges provide the basic structure of Colombia (Bushnell, 1993, p. 1). And they have created a natural fragmentation that has greatly influenced the development of distinct regional economies and cultural characteristics. For a majority of its history, Colombia’s population has been relatively sparse and scattered throughout the country as disconnected communities. The relatively few accounts of travel during the 18\textsuperscript{th} century, and the more abundant accounts in the 19\textsuperscript{th} century, have made it clear that historically large stretches of territory have been lightly populated, or almost devoid of inhabitants entirely. This dispersion and discontinuity of the people, along with the natural
fragmentation of the land, has tended to deter and delay the development of economic integration and growth in Colombia (Safford and Palacios, 2002, p. 1).

According to the general census conducted by the Departamento Administrativo Nacional de Estadística (National Administrative Department of Statistics – DANE) (2005), Colombia had an adjusted population estimate of 42,888,592 people in 2005. The top three major cities were Bogotá, Medellín, and Cali. The map in Figure 3.2 shows the location of several major cities along with the density and distribution of the Colombian population. It reveals that approximately 75% of Colombia’s population is concentrated in the Andean region, which represents only 25% of the country’s total area. In contrast, the Amazon-Orinoquía regions represent approximately 50% of the country’s total area but only accounts for less than 1% of the population. While likely the product of acute population disparity, the mountains, including the areas surrounded by mountains, are where armed conflict has been felt most intensely (Ospina, 2006).

The administrative boundaries of the Colombia are structured into several levels. There is, of course, the country as a whole at the national level, then 32 administrative units at the department level, and 1,120 administrative units at the municipal level. These administrative units are shown in Figure 3.3.
Figure 3.2: Colombia Population Density and Major Cities, 2005.
Data Source: DANE (2005), SIGOT (2012)
Figure 3.3: Administrative units of Colombia in 2005.
Data Source: SIGOT (2012)
The map in Figure 3.4 depicts the regions of Colombia that are both included and excluded from the study area. Of the 1,120 municipalities, 243 were excluded from this study, leaving a total of 877 among the study area. These 243 municipalities were not included because they either physically disconnected from the mainland (e.g. islands) or reported as having a homicide rate of zero in 2005. While it may be possible that some of these areas did not experience homicide, it is doubtful that 233 out of the 1,120 municipalities can actually make this claim. The more likely scenario is that the homicide rate dataset classifies unavailable records as zero instead of a null value. Therefore, all municipalities reporting zero homicides were removed from the study area.
Figure 3.4: Regional map of Colombia study area. 
3.2 Data and Data Processing

Based on the unique nature of Colombia and the variables adopted by previous studies, nine independent variables (Table 3.1) were initially selected for this study. It should also be noted that these variables were selected also because of their availability at the municipal level from various data sources. They, excluding homicide rate, are divided into five categories, representing social strife, economic condition, state presence, political stability, and physical geography of Colombia.

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violence</td>
<td>Homicide Rate (D)</td>
<td>SIGOT (2012)</td>
</tr>
<tr>
<td>Social Strife</td>
<td>Displacement Rate (I)</td>
<td>SIGOT (2012)</td>
</tr>
<tr>
<td>Economic Condition</td>
<td>Poverty Rate (I)</td>
<td>SIGOT (2012)</td>
</tr>
<tr>
<td></td>
<td>Standard of Living Index (I)</td>
<td>SIGOT (2012)</td>
</tr>
<tr>
<td>State Presence</td>
<td>Coca Cultivation Density (I)</td>
<td>SIMCI (2005)</td>
</tr>
<tr>
<td></td>
<td>Presidential Election Participation Rate (I)</td>
<td>Registraduría Nacional del Estado Civil (2006)</td>
</tr>
<tr>
<td>Political Stability</td>
<td>FARC Armed Actions Rate (I)</td>
<td>SIGOT (2012)</td>
</tr>
<tr>
<td></td>
<td>ELN Armed Actions Rate (I)</td>
<td>SIGOT (2012)</td>
</tr>
<tr>
<td></td>
<td>Public Force Armed Actions Rate (I)</td>
<td>SIGOT (2012)</td>
</tr>
<tr>
<td>Physical Geography</td>
<td>Terrain Ruggedness Index (I)</td>
<td>CGIAR-CSI (2008)</td>
</tr>
</tbody>
</table>

Table 3.1: Categorical view of potential regression variables used in this study. Note: (D) indicates dependent variable and (I) indicates independent variable.
3.2.1 *Homicide Rates*

The municipal homicide rate dataset for Colombia in 2005 was obtained from Sistema de Información Geográfica para la Planeación y el Ordenamiento Territorial Nacional (Geographic information systems for planning and land – SIGOT), which constitutes a network of institutions, technological resources and data to facilitate access and use of geospatial information in Colombia. This organization provides information for the management of development at national, regional, and local government levels. The metadata for the homicide rates dataset gives credit to the Vicepresidencia de la República as the original source. The homicide rates are defined as the ratio of total homicide cases in a municipality over the municipality's total population, measured as the number of killings per 100,000 people. Additionally, the metadata notes that the different rates between municipalities may be affected by the security and armed conflict of those areas (SIGOT, 2012).

In Figure 3.5, a choropleth map depicts the spatial distribution of homicide rates throughout the Colombian study area. Many of the areas expressing the highest homicide rates are located in the western and southeast regions. These two regions appear to be divided by municipalities expressing relatively low homicide rates. Interestingly, the two regions with the highest homicide rates are separated by the physical boundaries of the Cordillera Oriental mountain range.
Figure 3.5: Homicides per 100,000 people, 2005. Data Source: SIGOT (2012)
3.2.2 Coca Crops

The United Nations Office on Drugs and Crime (UNODC) developed a project known as the Sistema Integrado de Monitoreo de Cultivos Ilícitos (SIMCI), which employs technological resources to generate a census of coca crops, monitor their dynamics, and determine the potential production of cocaine hydrochloride in the country. According to their official website, the SIMCI project has measured the extent of coca cultivation in the country since 1999 by using satellite imagery technology, supported by aerial photographs and field verification (UNODC, 2012). For this study, tabular data was acquired from the SIMCI project which provided the area (hectares) of coca cultivation for each Colombian municipality in 2005. Coca crops were detected, measured, and recorded in 190 out of the 1,120 total municipalities.

In Figure 3.6, a choropleth map depicts the spatial distribution of coca cultivation throughout the Colombian municipalities. The distribution is clearly non-random and suggests that there are regions of Colombia that are more conducive for cultivating coca than others. After further examination, it appears that the areas of high-density coca cultivation tend to be located outside of urban areas and adjacent to the principal mountain ranges. While interesting, this is not surprising since many of these areas are likely located at the fringe of state authority and control.
Figure 3.6: Coca cultivation density, 2005. Data Source: SIMCI (2005), SIGOT (2012)
3.2.3 Political Participation

The Registraduría Nacional del Estado Civil (2012) is a Colombian organization whose mission is to ensure the transparency and reliability of the electoral process at various levels of government: local, congressional, presidential. During the 2006 presidential election of Colombia, they consolidated data on the actual and potential electoral voters for each municipality throughout the country. These values were then converted to a percentage that reflected voter turnout rates for their respective municipalities. Soon thereafter, they provided their election data to the general public via their website. Unfortunately, the data was unavailable in a tabular format for this study and therefore had to be manually entered and formatted for all 1,120 municipalities.

In Figure 3.7, a choropleth map depicts the spatial distribution of the participation rates for the Colombian presidential election in 2006. The distribution clearly shows that municipalities in the northwest region of Colombia express some of the highest participation rates in the country. There also tends to be higher participation rates in and around urban areas. Nevertheless, the Casanare department, which extends easterly from the Cordillera Oriental, is a rural region of Colombia that also expresses an apparent investment in the political processes of the country.
Figure 3.7: Participation rate for the presidential election, 2006. Data Source: Registraduría Nacional del Estado Civil (2006), SIGOT (2012)
3.2.4 *Displaced Persons*

People that were driven out of their municipality as a result of persecution, violence, or economic forces are considered displaced persons. The dataset for this variable includes the count of individuals within each municipality that have registered with the state as displaced territorial entities. For this study, the SIGOT (2012) project provided the number of displaced persons from Colombian municipalities in 2005. These values were then divided by the total population of their respective municipality and multiplied by 100,000 to generate a forced displacement rate. The metadata for this dataset provided credit to the Departamento para la Prosperidad Social as the original source. Figure 3.8 shows the spatial distribution of the displaced persons rate throughout Colombian, which appears to suggest that the displacement is most often a rural phenomenon. Regions that do not express high rates of displacement tend to be urban areas that are among the principal mountain ranges. Low rates of displacement are also found in the more remote areas of Colombia’s southern Amazon region.
Figure 3.8: Displaced persons per 100,000 people, 2005. Data Source: SIGOT (2012)
3.2.5 Poverty Rate

The Departamento Administrativo Nacional de Estadística (DANE) is responsible for the planning, collecting, processing, analyzing, and the dissemination of official statistics in Colombia (DANE, 2012). In 2005, DANE conducted a national census of its population to quantify the economic and social development of Colombia. They developed an unsatisfied basic needs (NBI) index to represent the unmet basic needs for each municipality. The index identifies people dissatisfied with any (one or more) of the needs identified as basic and simple indicators of unmet basic needs include inadequate housing, housing with critical overcrowding, homes with inadequate services, homes with high economic dependence, and households with school-age children not attending school. To measure the magnitude of poverty in relation to population, DANE decided that people living in households with NBI were in the same condition of their respective housing (SIGOT, 2012).

Figure 3.9 shows the spatial distribution of the poverty rate throughout the Colombian municipalities. Once again, the distribution appears to suggest that high poverty rates are a rural phenomenon. Areas in and around the major cities and urban areas of the principal mountain ranges express some of the lowest poverty rates in the country. The highest rates of poverty are located along the Pacific coast, the northern region, and a majority of the eastern and southern regions. It is apparent that much of the wealth is contained within the more densely populated urban areas of Colombia.
Figure 3.9: Percentage of the population in poverty status, 2005.
Data Source: SIGOT (2012)
3.2.6 Standard of Living Index

Provided by the SIGOT (2012) project for municipalities in 2005, the standard of living index is composed of variables related to education and human capital, access and quality of services, and household size and composition, and housing quality. It allows for an approximation of the quality of life for households. The standard of living index consists of four factors: 1) individual human capital (education); 2) household structure and demographics (children under six years and overcrowding); 3) collective physical capital (basic sanitation and cooking fuel); and 4) individual physical capital (flooring material and walls of houses). Figure 3.10 depicts the spatial distribution of the standard of living index throughout the Colombian municipalities. It is not surprising that the municipalities with the highest standards of living are also more urbanized and more highly populated than other parts of Colombia. The municipalities with the lowest standards of living appear to be in the western region along the Pacific coast and in the eastern region beyond the Cordillera Oriental. Essentially, these two regions are considerably more rural than the other areas of the country and are home to the Afro-Colombian and indigenous populations respectively.
Figure 3.10: Colombia standard of living index, 2005. Data Source: SIGOT (2012)
3.2.7 **Terrain Ruggedness Index**

The ruggedness of the Colombian landscape can be obtained by calculating the terrain ruggedness index (TRI). This index provides an objective quantitative measure of topographic heterogeneity that is derived from a digital elevation model (DEM) (Riley et al., 1999). A DEM was acquired from the Consultative Group on International Agricultural Research (CGIAR). This organization is a global partnership dedicated to reducing poverty and hunger, improving human health and nutrition, and enhancing ecosystem resilience through advanced agricultural research. Within this organization, the Consortium for Spatial Information (CGIAR-CSI) promotes and practices the application of spatial science to most effectively achieve their goals. For this study, they provided Shuttle Radar Topography Mission (SRTM) 90m DEM data, originally produced by NASA, which has been processed to fill data voids.

The topographic ruggedness index (TRI) is a measurement developed by Riley et al. (1999) to express the elevation difference between the adjacent cells of a DEM. Essentially it calculates the difference in elevation values between a center cell and its eight surrounding cells. These values are then squared to ensure that they are positive and then added together. The TRI is then derived by calculating the square root of this summation. In other words, this index represents the average change in elevation surrounding any point of a DEM. The formula can be expressed as $\left[ \sum (x_i - x_{0,0})^2 \right]^{1/2}$ where $x_i$ is the elevation of each cell adjacent to cell (0,0) within the 3x3 neighborhood. Figure 3.11 provides an example of how to obtain the TRI value for a particular cell within a hypothetical digital elevation grid.
Figure 3.11: Hypothetical elevation data where (a) is a peak type topography, (b) are the squared differences between the center cell, and (c) is the calculated TRI value.

While the mathematical equation for the TRI measure is relatively straightforward, the technique to calculate this index with ArcGIS Desktop 10 is considerably more complex. By employing some basic algebra, the equation can be transformed to better suit the available tools within ArcToolbox. To elaborate, let $x$ represent the center cell, let $x_i$ represent the values of the eight neighboring squares, and let $r$ represent the topographic ruggedness index. Since $r = \left[ \sum (x_i - x_{0,0})^2 \right]^{1/2}$, it can also be expressed as $r^2 = \sum (x_i - x_{0,0})^2$.

In addition, two other variables also need to be calculated using the Focal Statistics tool within ArcToolbox. Let $s$ be the first variable, represented as the sum of the values in the neighborhood. This can be expressed as $s = \sum x_i + x$. Let $t$ represent the second variable, which is equal to the sum of the squared values in the neighborhood. This can be expressed as $t = \sum x_i^2 + x^2$.

Having defined all of the variables, the following algebraic operations and substitutions can be then performed to calculate the TRI. The expansion of $r^2$ becomes:
\[ r^2 = \Sigma [x_i^2 + x^2 - (2x \cdot x_i)] \]
\[ r^2 = \Sigma (x_i^2) + 8x^2 - (2x \cdot \Sigma x_i) \]
\[ r^2 = [\Sigma (x_i^2) + x^2] + 7x^2 - \{[2x \cdot (\Sigma x_i + x - x)]\} \]
\[ r^2 = t + 7x^2 - [2x \cdot (\Sigma x_i + x)] + 2x^2 \]
\[ r^2 = t + 9x^2 - (2x \cdot s) \]

The final expansion of \( r^2 \) allows for an easily executed workflow of raster operations with ArcGIS Desktop 10. As seen in Figure 3.12, the variable \( s \) can calculated with the Focal Statistics tool within the Neighborhood toolset of the Spatial Analyst Tools toolbox. This tool requires an input raster (original DEM) and an output raster (DEM_S). The neighborhood was set to Rectangle with a Height and Width of 3 units and the Statistics type to SUM. After executing the operation, the newly generated DEM_S raster assumes the role of the \( s \) variable in the expanded \( r^2 \) equation.
The next operation is to calculate the squared values of $x_i$ and $x$. Using the Raster Calculator tool (Figure 3.13) within the Map Algebra toolset of the Spatial Analyst Tools toolbox, these squared values can be calculated by simply squaring the DEM raster. The algebra expression is thus $DEM * DEM$, which generates the output raster $DEM_{SQ}$. 

![Image of Focal Statistics tool with 3x3 neighborhood settings and SUM statistic.]
Having calculated the squared values of $x_t$ and $x$, the raster equivalent of the $t$ variable can then be calculated by performing another focal sum. The Focal Statistics tool was employed once again (Figure 3.14) with the input raster was set to DEM_SQ and the output raster was named DEM_T. The neighborhood was set to Rectangle with a Height and Width of 3 units and the Statistics type was set to SUM. After executing the operation, the newly generated DEM_T raster assumes its role as the $t$ variable in the expanded $r^2$ equation.
Based on the aforementioned equation, \( r^2 = \sum \left( \frac{t}{g^2} \right)^2 \), \( r^2 \) could finally be computed by using the raster datasets that are generated from the Focal Statistics and Raster Calculator tools. By substituting the variables with their respective raster datasets, the equation can be expanded once more as the following:

\[
  r^2 = DEM_T + (9 \times DEM_{SQ}) - (2 \times DEM \times DEM_S)
\]
As seen in Figure 3.15, this equation was entered into the Raster Calculator tool to generate another raster dataset named DEM_R2. The output from this operation represents TRI² and thus requires an additional operation to compute its square root. The final DEM_TRI raster dataset can be calculated again by the Raster Calculator (Figure 3.16) with the equation, \texttt{SquareRoot("DEM\_R2")}.
Figure 3.16: Calculating the square root to obtain the $r$ (TRI) values.
Source: ESRI (2011)

The resulting DEM_TRI raster dataset characterizes the terrain ruggedness index for every cell within the original digital elevation model. It should be noted that the original DEM and the subsequent TRI raster datasets were not clipped to the political boundaries of Colombia before the analysis. This would have created an edge effect, which arises where an artificial boundary is imposed on a study area. Had the original DEM been clipped to the political boundaries of Colombia, the resulting TRI values along the border would have been incorrect since the required neighborhood cells would be absent from the calculation. Only after the analysis were either of these datasets clipped to the national boundaries. In Figure 3.17, there is a side-by-side depiction of the
original DEM raster and the derived TRI raster for Colombia. Notice that several plateaus in the mountainous regions exhibit very high elevation values and very low TRI values. This reflects the notion that elevation does not necessarily reflect terrain ruggedness; such is the case with many of the urban areas of the principal mountain ranges.

Figure 3.17: (a) Original DEM raster (b) Computed TRI raster derived from DEM raster. Data Source: CGIAR-CSI (2008)

Figure 3.18 depicts the average terrain ruggedness index (TRI) throughout the Colombian municipalities. As expected, this map indicates that the most rugged areas of the country are among the principal mountain ranges. The TRI in the outlying rural areas is relatively low as the changes in elevation are typically minimal. Additionally, high-elevation plateaus exhibit low TRI values, depicting a less-rugged terrain surrounded by highly-rugged mountain ranges.
Figure 3.18: Terrain Ruggedness Index.
Data Source: CGIAR-CSI (2008), SIGOT (2012)
3.2.8 Armed Actions

The SIGOT (2012) project also provided the municipal level data for the number of armed conflicts initiated by different armed groups in 2005. These armed groups included the FARC, the ELN, and the Public Force, which is comprised exclusively of armed military forces and the National Police (Constitución Política De Colombia, 2011). For each municipality, the number of armed actions initiated by these groups were divided by their respective total population and multiplied by 100,000. This generated a municipal rate to represent the number of initiated armed actions per 100,000 people for each of the armed groups.

Figure 3.19 shows the spatial distribution of the armed actions initiated by the FARC throughout the Colombian municipalities. The FARC accounts for 649 of the armed actions in 2005, which is approximately 30% of the total initiated armed actions that occurred in Colombia that year. Most of these armed actions tend to occur in the southwest and central regions. Those cases in the central region are located to the east of the Cordillera Oriental mountain range. There are also several armed actions that occur along the coasts of the north-west and north-east regions as well.

Figure 3.20 depicts the spatial distribution of the armed actions initiated by the ELN. There are only 44 armed actions initiated by the ELN in 2005, which is approximately 2% of the total initiated armed actions that occurred in Colombia that year. The armed actions of the ELN are spread somewhat erratically from the Pacific coast to the Atlantic coast across the north-central region. While relatively few in number, the ELN continually aim to disrupt Colombia’s economic infrastructure by sabotaging and bombing strategic assets, including oil pipelines and electrical systems (DEA, 2002).
Figure 3.21 shows the spatial distribution of the armed actions initiated by the Public Force throughout the Colombian municipalities. The Public Force accounts for 1,483 of the initiated armed actions in 2005, which is approximately 68% of the total initiated armed actions in Colombia that year. Most of the armed actions appear to occur in the southern and eastern regions of the country. Regions that do not experience a high number of armed actions by the Public Force tend to be in and around major urban areas.
Figure 3.19: Armed actions initiated by the FARC per 100,000 people, 2005.
Data Source: SIGOT (2012)
Figure 3.20: Armed actions initiated by the ELN per 100,000 people, 2005.
Data Source: SIGOT (2012)
Figure 3.21: Armed actions initiated by the Public Force per 100,000 people, 2005. 
Data Source: SIGOT (2012)
3.3 Spatial Statistical Approaches

Based on the findings from many existing studies of violence in Colombia, it has been well established that violence does not occur in random areas of the country. While certain factors continually attribute to the occurrence of violent crimes, some areas are vastly more affected than others. The areas of Colombia with high homicide rates usually have similar rates in their adjacent municipalities and this clustering creates a discernable geographic pattern. The following sections provide a detailed discussion of spatial statistics categories and methods that are adopted into this study.

3.3.1 Global Moran’s I

To measure the spatial autocorrelation of homicide rates throughout Colombia, this study incorporates Global Moran’s I (GMI). It is a global measure of the correlation among neighboring observations in a pattern (Boots and Getis, 1988). Given a set of features and an associated attribute, it evaluates whether the pattern expressed is clustered, dispersed, or random. When values for neighboring features are either both larger than the mean or both smaller than the mean, the cross-product would be positive. When one value is smaller than the mean and the other is larger than the mean, the cross-product would be negative. This cross-product is represented by the Moran’s Index, bounded from -1.0 to 1.0, to assess the overall pattern and trend of data (ESRI, 2012a).

3.3.2 Anselin Local Moran’s I

In addition to Global Moran’s I, this study also utilized a local cluster and outlier analysis known as Anselin Local Moran’s I (LMI). Given a set of weighted features, an
LMI analysis can identify the locations of spatial clusters and outliers of features with attribute values similar in magnitude (ESRI, 2012b). Unlike the GMI, which only provides an indication of patterns over an entire study area, the LMI is a local indicator of spatial association (LISA) that measures the degree of spatial dependence between a particular location and its neighbors, where neighborhood is defined according to some measure of proximity (Anselin, 1995). The output from this analysis is broken down into four categorical quadrants that correspond to the Moran scatterplot. Areas with a positive spatial autocorrelation are classified by both high-high (HH) and low-low (LL) combinations. The HH clusters are labeled as hotspots and the LL clusters are labeled as coldspots. The other two classifications are high-low (HL) and low-high (LH), which are indicative of patterns with negative spatial autocorrelation. These negatively correlated areas contain outliers where a high value is within a low value neighborhood and vice versa (Messner et al., 1999, p. 439).

3.3.3 Regression Analysis and Spatial Data

Regression analysis has become one of the most broadly used statistical methods for analyzing multifactor data. Its appeal stems from its simple method for investigating relationships among variables. The general approach for a regression analysis is to acquire data, generate a model, and then evaluate its fit (Chatterjee and Hadi, 2006, p. xiii). According to Charlton and Fotheringham (2009a, p. 1), a regression “encompasses a wide range of methods for modeling the relationship between a dependent variable and a set of one or more independent variables”. In its simplest form, a linear regression model can be expressed in the following equation:
\[ y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad \text{for } i=1 \ldots n \]

where \( y_i \) is the response variable, measured at location \( i \), \( x_i \) is the independent variable, \( \varepsilon_i \) is the error term, and \( \beta_0 \) and \( \beta_1 \) are parameters which are to be estimated such that the value \( \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \) (sum of squared errors) is minimized over the \( n \) observations in the dataset. The \( \hat{y}_i \), is the predicted or fitted value for the \( i^{th} \) observation (Charlton and Fotheringham, 2009a, p. 1). A linear regression model can also be developed for the relationship between a dependent variable and a set of independent variables (ESRI, 2012d), represented in the following form:

\[ y_i = \beta_0 + \beta_1 x_i + \beta_2 x_{2i} + \ldots + \beta_m x_{mi} + \varepsilon_i \quad \text{for } i=1 \ldots n \]

Usually, the unknown parameters, mainly the coefficients \( \beta_0, \beta_1, \ldots \beta_m \), in a linear regression model can be estimated via a method called Ordinary Least Squares (OLS), which minimizes the sum of all squared differences between the observed and predicted values of a model. Be that as it may, they require some underlying assumptions. The first is that the observations should be independent of one another. With respect to spatial data, this is not always the case (Charlton and Fotheringham, 2009a, p. 3). As noted by Tobler (1970, p. 236) in his infamous 1st Law of Geography – “everything is related to everything else, but near things are more related than distant things”. It is therefore reasonable to expect that the variables and residuals of a regression model might exhibit spatial dependence where nearby locations have similar values (Charlton and
Fotheringham, 2009a, p. 3). Typically, a regression model is applied to the entire study area, and as a result it is often referred as global regression model.

When applying a global regression model, the second assumption is the relationships being modeled are consistent throughout the study area. This assumption often fails when spatially varying data is included in the model (Charlton and Fotheringham, 2009a, p. 3). To compensate for this, the model would need to account for the localized spatial variations of the data by weighting observation values of nearer locations more heavily than distant ones.

In addition to these assumptions, several issues can impact the performance of a regression model. For example, multicollinarity amongst the independent variables has the potential to affect the overall performance of a regression model. This occurs when there are strong correlations among explanatory variables themselves that create an unnecessary redundancy within the model. Multicollinarity leads to an over-counting type of bias which results in the model becoming unstable and sometimes unreliable. Another issue that can plague regression models is spatially autocorrelated residuals. When there is a spatial clustering of the under/over predictions from the model, it introduces an over-counting type of bias that can also render the model unreliable. Lastly, a common issue is the omission of explanatory variables, referred to as misspecification. This occurs when key explanatory variables are missing from a regression model. As a result, variable coefficients and their associated p-values cannot be trusted (ESRI, 2012c).
3.3.4 Geographically Weighted Regression (GWR)

In an effort to overcome some of the issues that surround the modeling of spatial data, the Geographically Weighted Regression (GWR) model was a local regression model developed by Brunsdon et al in 1996. This model addresses the notion that coefficients do not remain fixed over space, referred to as spatial non-stationarity (Brunsdon et al., 1998). The underlying idea of GWR is that “parameters may be estimated anywhere in the study area given a dependent variable and a set of one or more independent variables which have been measured at places whose location is known” (Charlton and Fotheringham, 2009a, p. 5). Additionally, it accounts for the 1st Law of Geography by weighting nearer observations more than those that are further away. The equation for a typical GWR model is represented in the following form:

\[ y_i(u) = \beta_{0i}(u) + \beta_{1i}(u) x_{1i} + \beta_{2i}(u) x_{2i} + \ldots + \beta_{mi}(u) x_{mi} \]

The notation \( \beta_{0i}(u) \) indicates that this particular parameter describes a relationship around location \( u \) and is specific to that location. A prediction for the dependent variable can be made if measurements for the independent variables are also available at the location \( u \) (Charlton and Fotheringham, 2009a, p. 5). Essentially, the GWR model allows one to estimate variable coefficients at local levels across a global study area. Hence GWR is a member of so-called local regression models. The estimator for this model utilizes a weighted least squares method that is conditioned on the location \( u \) relative to the other observations in the dataset and thus changes for each location (Charlton and Fotheringham, 2009a, p. 5). The estimator takes the following form:
\[ \hat{\beta}(u) = (X^T W(u) X^{-1} X^T W(u) y \]

Within this equation, \( W(u) \) is a square matrix of weights relative to the position of \( u \) in the study area, \( X^T W(u) X \) is the geographically weighted variance-covariance matrix, and \( y \) is the vector of the values of the dependent variable. Charlton and Fotheringham (2009a, p. 6) further explained that the \( W(u) \) matrix contains the geographical weights in its leading diagonal and 0 in its off-diagonal elements. These weights are computed from a weighting scheme that is also known as a *kernel* which typically has a Gaussian shape. The following is a representation of the square matrix of weights near location \( u \):

\[
\begin{bmatrix}
  w_1(u) & 0 & 0 & 0 \\
  0 & w_2(u) & 0 & 0 \\
  0 & 0 & \ldots & 0 \\
  0 & 0 & 0 & w_n(u)
\end{bmatrix}
\]
3.4 GIS Implementation

3.4.1 Global Moran’s I

Before engaging in a comprehensive spatial analysis, the phenomenon being studied needs to be evaluated for a measure of spatial autocorrelation, known as Global Moran’s I. The Spatial Autocorrelation (Global Moran’s I) tool (Figure 3.22), within the Spatial Statistics Tools toolbox of ArcGIS Desktop 10, measures levels of spatial autocorrelation based on both feature locations and feature values simultaneously. It requires an Input Feature Class and an Input Field to be selected. For this study, the municipal level feature class and the rate of homicide were chosen respectively. Thereafter, several parameters were adjusted to reflect the inherent relationships and desired measure of distances among the features being analyzed.

The first parameter is the Conceptualization of Spatial Relationships, which determines how spatial relationships among features are conceptualized. For this parameter, the Inverse Distance method was selected because it more appropriate for modeling processes where the closer two features are in space, the more likely they are to influence each other (ESRI, 2012a).

The second parameter is the Distance Method, which specifies how distances are calculated from each feature to neighboring features. For this study, a Euclidean method is selected to measure the straight-line distance between polygon centroids since travel is not necessarily restricted to a street network (ESRI, 2012a).

The third parameter is the Standardization, which is recommended whenever the distribution of features is potentially biased due to sampling design or an imposed aggregation scheme. The options for this parameter are either None or Row. The row
standardization divides each weight by the sum of the weights of all neighboring features. This method is often used with fixed distance neighborhoods and almost always used for neighborhoods based on polygon contiguity, i.e., polygons that share a boundary or corner. It is intended to mitigate bias due to features having different numbers of neighbors. Row standardization scales all weights so that they are between 0 and 1, which creates a relative, rather than absolute, weighting scheme. Since the polygon features in this study represent administrative boundaries, the Row standardization was selected (ESRI, 2012a).

![Spatial Autocorrelation (Morans I) tool](image)

Figure 3.22: Spatial Autocorrelation (Global Moran’s I) tool. Source: ESRI (2011)

The Spatial Autocorrelation (Global Moran’s I) tool is an inferential statistic where the results of the analysis are interpreted within the context of the null hypothesis.
For the Global Moran’s I statistic, the null hypothesis states that the attribute being analyzed is randomly distributed throughout the study area. As a point of caution, researchers should be aware that this tool may result in a Type I error where the null hypothesis is rejected when it is in fact true (ESRI, 2012a).

After executing this tool, it generates a report containing a Moran’s Index, a z-score, and a p-value that dictates whether to accept or reject this null hypothesis. An example of this report is shown in Figure 3.23. According to ESRI (2012a), the null hypothesis should be rejected if the p-value is statistically significant and the z-score is positive. This indicates that the spatial distributions of high and low values in a dataset are more spatially clustered than would be expected if the underlying spatial processes were random.

![Image of Moran’s I statistic report](image_url)

Figure 3.23: Example of spatial autocorrelation report. Source: ESRI (2011)
3.4.2 *Local Moran’s I*

After evaluating and confirming the presence of global clusters for a selected feature class and input field, the locations of those clusters can be determined by using the Local Moran’s I (LMI). Given a set of weighted features, this statistical measure can identify and classify the spatial clusters of features with attribute values that are similar in magnitude. The Cluster and Outlier Analysis (Anselin Local Moran’s I) tool (Figure 3.24) within the Spatial Statistics Tools toolbox of ArcGIS Desktop 10, identifies the location of statistically significant hot spots, cold spots, and spatial outliers using the Anselin Local Moran’s I statistic. This tool also requires an input feature class, an input field, and several parameters that account for the spatial relationships, measure of distance, and standardization among the features. It then generates an output feature class with the following attributes for each feature in the input feature class: Local Moran’s I index, z-score-p-value, and cluster/outlier type (COType) (ESRI, 2012b).

Once again, the z-scores and p-values are measures of statistical significance which determine whether or not to reject the null hypothesis, feature by feature. A feature with a high positive z-score indicates that the surrounding features have similar values, which are either high or low values. The COType field for these features is therefore expressed as HH for a statically significant (0.05 level) cluster of high values and LL for a statistically significant (0.05 level) cluster of low values. Those features with a low negative z-score indicate that the feature was a statistically significant (0.05 level) spatial outlier. The COType field for these features is therefore expressed as HL for high value features surrounded by low value features and LH for low value features surrounded by high value features (ESRI, 2012b).
3.4.3 Potential Regression Variables

This study utilizes both Ordinary Least Squares (OLS) regression and Geographically Weighted Regression (GWR) models to help identify and interpret the causal relationships among homicide rates and an assortment of independent variables at the municipal level for Colombia. These independent variables can be divided into several categories that include social strife, economic conditions, state presence, political stability, and physical geography (listed in Table 3.1). The following describes how they represent their respective categories:
1) **Social Strife**: The causal relationship between homicide rates and the social nature of Colombia were characterized by the rate of displacement within the municipalities. Areas that experience higher rates of displacement are more indicative of the social strife between opposing political and military ideologies. The independent variable **Displaced Persons Rate** was selected to represent this category by providing a percentage of displaced people within the municipalities of Colombia in 2005.

2) **Economic Conditions**: The causal relationship between homicide rates and the economic nature of Colombia were characterized by the poverty rate and the standard of living index of the municipalities. Areas that express high rates of poverty and high standards of living may indicate certain levels of economic disparity that could contribute to municipal violence. The independent variables **Poverty Rate** and **Standard of Living Index** were selected to represent this category by providing a percentage of people with limited income along with the overall quality and accessibility of adequate housing and services for the municipalities of Colombia in 2005.

3) **State Presence**: The causal relationship between homicide rates and the state presence in Colombia was characterized by the political participation rates and the density of coca cultivation within the municipalities. While it is difficult to quantify the state presence and its authority within a given area, these two independent variables were chosen as suitable proxy indicators for the state presence in Colombia. The independent variable **Presidential Election Participation Rate** was one of two independent variables selected to represent this category by providing the voter turnout rate of the municipalities in 2006. Areas with a relatively high voter turnout
generally reflect the legitimacy and implicit endorsement of the current system of the government. Those areas with a relatively low voter turnout may indicate voter suppression, dissatisfaction, or indifference of the people. The other independent variable, *Coca Cultivation Density*, might also identify areas where the state has little or no presence. Those areas where coca cultivation tends to thrive may indicate a lack of authority and enforcement of government policies.

4) **Political Stability**: The causal relationship between homicide rates and the political stability of Colombia were characterized by the rate of armed actions among the guerrillas and public forces. Areas that experience a high rate of armed actions indicate the presence of opposing political and military forces that may directly correlate with the levels of violence among conflicted communities. The independent variables *FARC Armed Actions Rate, ELN Armed Actions Rate*, and *Public Force Armed Actions Rate* were selected to represent this category by providing a municipal rate for the number of armed actions recorded in Colombia for 2005.

5) **Physical Geography**: The causal relationship between homicide rates and the physical geography of Colombia was characterized by the topographic ruggedness index (TRI). Areas that are characterized as being more rugged may indicate where guerrilla groups have found sanctuary for their organizations. The terrain could also be a significant indicator for zones of transition that separate opposing political and military ideologies. The independent variable *Terrain Ruggedness Index* was selected to represent this category by providing an average TRI for the municipalities of Colombia.
3.4.4 *Transforming Regression Variables*

All dependent and independent variables should be examined prior to executing any linear regression analysis to ensure that they are normally distributed to meet with the basic assumption of linear regression and otherwise a linear regression model will perform poorly (ESRI, 2012c). The scatterplot matrix in Figure 3.25 includes the histograms that show the distribution for each variable. The scatterplots also illustrate the initial relationships among the variables. The values for several of these variables appear to be positively skewed and lack a normal distribution, specifically *Homicide Rate*, *Coca Cultivation Density*, *Displaced Persons Rate*, *Poverty Rate*, *FARC Armed Actions Rate*, *ELN Armed Actions Rate*, and *Public Force Armed Actions Rate*. 
ESRI (2012c) notes that these skewed variable distributions can often be remedied by transforming them with a variety of methods. This process can potentially improve model performance and eliminate bias. Transformation methods were thus applied to several variables in an effort to improve their distributions. As a result, log transformations were found to be the best method for transforming *Homicide Rate*, *Displaced Persons Rate*, *Poverty Rate*, *FARC Armed Actions Rate*, *ELN Armed Actions Rate*, and *Public Force Armed Actions Rate* into more normal distributions. For the independent variables that contained values equal to zero, the transformation $\log(x+1)$
was applied as log(0) is equal to infinity. Additionally, the square root transformation was found to be the best method for transforming Coca Cultivation Density into a more normal distribution.

After completing the transformations, a new scatterplot matrix was generated (Figure 3.26) and, while not perfect, it reveals a far better distribution of values for the variables. The variables Coca Cultivation Density, FARC Armed Actions Rate, ELN Armed Actions Rate, and Public Force Armed Actions Rate, however, contain a high number of values that are equal to zero and continue to lack a normal distribution. Be that as it may, the log and square root transformations did improve them. These particular variables would benefit from more advanced transformation techniques that are beyond the scope of this study.
3.4.5 Selecting Regression Variables

Having normalized the distributions of these variables, the next step is to properly specify an OLS model with the most important contributing factors. This procedure can be rather difficult as it requires one to evaluate model performances for the numerous combinations of potential independent variables. To achieve this, an Exploratory Regression tool (Figure 3.27) can be used within the Spatial Statistics Tools toolbox of ArcGIS 10.1. This tool automatically evaluates all of the possible combinations of
potential explanatory variables. It generates a report of the redundancy, completeness, significance, bias, and performance results for all variable combinations (ESRI, 2012e).

Figure 3.27: Exploratory Regression tool. Source: ESRI (2012f)

The output results table from the Exploratory Regression tool contains model diagnostics for each combination of potential independent variables. These diagnostics include Adjusted R-Squared (R2), corrected Akaike Information Criteria (AICc), Jarque-
Bera p-value (JB), Koenker’s studentized Breusch-Pagan p-value (BP), Variance
Inflation Factor (VIF), and Global Moran’s I p-value (SA). ESRI (2012g) recommends
using the AICc values as a way to identify the best and worst performing models. These
AICc values are determined by taking into account the divergence between the observed
and predicted values and by measuring the complexity of each model (Charlton and
Fotheringham, 2009b). Therefore, the models with the lowest AICc values are considered
to be the overall best performing models (ESRI, 2012g).

The Exploratory Regression tool was executed using Homicide Rate as the
dependent variable and Displacement Rate, Poverty Rate, Standard of Living Index, Coca
Cultivation Density, Presidential Election Participation Rate, FARC Armed Actions Rate,
ELN Armed Actions Rate, and Public Force Armed Actions Rate as the potential
independent variables to be tested. The additional criteria were set to the following:

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum Acceptable Adj R Squared:</td>
<td>0.25</td>
</tr>
<tr>
<td>Maximum Coefficient p value Cutoff:</td>
<td>0.05</td>
</tr>
<tr>
<td>Maximum VIF Value Cutoff:</td>
<td>2.50</td>
</tr>
<tr>
<td>Minimum Acceptable Jarque Bera p value:</td>
<td>0.10</td>
</tr>
<tr>
<td>Minimum Acceptable Spatial Autocorrelation p value:</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Three parameters for the Exploratory Regression tool were changed from their
default values: Minimum Acceptable Adj R Squared, Maximum VIF Value Cutoff Value
and Minimum Acceptable Spatial Autocorrelation p-value. The Minimum Acceptable
Adj R Squared value was changed from 0.50 to 0.25 as no models expressed an adjusted
R² value greater than 0.30. The VIF value measures the multicollinearity between variables within the model. Its default cutoff value was changed from 7.5 to a custom value of 2.5 to ensure that there were no redundant variables within the model. The Spatial Autocorrelation p-value indicates whether the model residuals are randomly distributed. Its default p-value was changed from 0.10 to a custom value of 0.00 as all models were reporting to be spatially autocorrelated.

With these criteria, the Exploratory Regression tool revealed two models with similarly low AICc values of 1993.87 and 1995.51. According to ESRI (2012h), AICc values that are less than 3 are inconsequential. A choice had to therefore be made about which model was better suited for this study. Both of these models included *Homicide Rate* as the dependent variable and *Displacement Rate, Coca Cultivation Density, Presidential Election Participation Rate, FARC Armed Actions Rate*, and *Public Force Armed Actions Rate* as the independent variables. The *ELN Armed Actions Rate* variable was not included in either as it was ultimately deemed insignificant. The primary difference between these two models was that one included the *Poverty Rate* variable while the other included the *Standard of Living Index variable*. These two variables were also marked as being slightly redundant in the diagnostic report, suggesting that only one of them would be necessary for this study. As a result, the choice between these two models was based on the results of the other diagnostic values.

The Jarque-Bera p-value is a measure of how normally a model’s residuals are distributed. If this value is low, it indicates that values are not normally distributed and that the model may include nonlinear relationships or data outliers. The Koenker’s studentized Breusch-Pagan p-value is a measure of how consistent the relationship is to
the dependent variable. If this value is low, it indicates that the variation in the predicted and observed values change throughout the study area (ESRI, 2012i).

The model that included the Standard of Living Index variable expressed both a larger Jarque-Bera p-value and a larger Koenker’s studentized Breusch-Pagan p-value and was thus selected as the strongest and most suitable OLS regression model for this study. The final list of regression variables is shown in Table 3.2.

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violence</td>
<td>Homicide Rate (D)</td>
</tr>
<tr>
<td>Social Strife</td>
<td>Displacement Rate (I)</td>
</tr>
<tr>
<td>Economic Condition</td>
<td>Standard of Living Index (I)</td>
</tr>
<tr>
<td>State Presence</td>
<td>Coca Cultivation Density (I)</td>
</tr>
<tr>
<td></td>
<td>Presidential Election Participation Rate (I)</td>
</tr>
<tr>
<td>Political Stability</td>
<td>FARC Armed Actions Rate (I)</td>
</tr>
<tr>
<td></td>
<td>Public Force Armed Actions Rate (I)</td>
</tr>
<tr>
<td>Physical Geography</td>
<td>Terrain Ruggedness Index (I)</td>
</tr>
</tbody>
</table>

Table 3.2: Categorical view of regression variables used in this study. 
*Note: (D) indicates dependent variable and (I) indicates independent variable.*

3.4.6 *Ordinary Least Squares Regression*

An OLS regression analysis is often the starting point for all spatial regression analyses by providing a global model for any variable or process that is trying to be understood (ESRI, 2012d). It allows one to assess the incremental improvements as
potential explanatory variables are introduced and removed from the model. It also provides the foundation for comparing the results of subsequent GWR analyses. Having prepared, transformed, and selected the most important regression variables for analysis, an OLS regression analysis was conducted with the Ordinary Least Squares tool (Figure 3.28) available within the Spatial Statistics Tools toolbox of ArcGIS Desktop 10.

Figure 3.28: Ordinary Least Squares tool. Source: ESRI (2011)
After executing the OLS tool in ArcGIS Desktop 10, it generates an output feature class and a report of statistical results. The report assesses the model's performance by providing *Multiple R-Squared* and *Adjusted R-Squared* values. These values are both derived from the regression equation to quantify model performance and fit. The value of R-squared can range from 0.0 to 1.0. If a model is to perfectly fit the dependent variable, the R-squared value would be 1.0. If, for example, there is an R-squared value of 0.49, it could be interpreted as the model explaining roughly 49% of the variation in the dependent variable. The adjusted R-squared value is always less than its corresponding multiple R-squared value because it reflects model complexity, or the number of variables, as it relates to the data. Therefore the adjusted R-squared value is a more accurate measure of model performance (ESRI, 2012d). As seen in Figure 3.29, this model can be plotted against the observed and predicted/estimated values. R-squared is then determined by how well the observed and estimated values match.

Figure 3.29: OLS regression - (a) model of observed and predicted values, (b) R-squared derived from how well estimated y-values match the observed y-values.

Source: ESRI (2012c)
In some instances, an OLS regression model can encounter a problem where there is a spatial clustering of under/over predictions which renders it unreliable. Fortunately, this can be detected by mapping the residual values from the OLS analysis as seen in Figure 3.30. Indications of spatial autocorrelation are evident when the values of features express patterns of clustering or dispersion. This also violates the assumption that the observed features are independent of one another and creates a circumstance where the standard errors are artificially low. Consequently, variables of a model may be deemed significant when in fact they are not. Spatially autocorrelated data, however, can be incorporated into an analysis that performs a local form of linear regression capable of modeling spatially varying relationships, such as GWR (ESRI, 2012d).

![Figure 3.30: OLS residual map.](image)

3.4.7 Geographically Weighted Regression

A Geographically Weighted Regression (GWR) is a member of emerging spatial regression methods that performs a local form of linear regression capable of modeling spatially varying relationships (ESRI, 2011). In this study, a GWR analysis was
performed by using the Geographically Weighted Regression tool (Figure 3.31) within the Spatial Statistics Tools toolbox of ArcGIS Desktop 10. Similar to the OLS model, it requires a dependent variable and a set of explanatory variables.

The GWR tool also required two additional parameters, namely *Kernel Type* and *Bandwidth Method*. The first parameter, *Kernel Type*, specifies if the kernel is constructed as a fixed distance or if it is allowed to vary in extent as a function of feature
density. The Kernel Type can thus be set to either Fixed or Adaptive. A Fixed Kernel Type is ideal when the locations of observed events are themselves not spatially clustered and therefore randomly distributed. Those observations that are spatially clustered, however, are better suited for an Adaptive Kernel. This type of kernel uses a spatial context that is the function of a specified number of neighbors. The spatial context is smaller in areas where feature distribution is relatively dense and larger in areas where the feature distribution is sparse. For this particular study, the Adaptive kernel type was selected.

The second parameter, Bandwidth Method, specifies how the extent of the kernel should be determined. The selectable methods are AICc, CV, and Bandwidth Parameter. The corrected Akaike Information Criterion (AICc) method determines the extent of the kernel by taking into account the divergence between the observed and predicted values, and by measuring the complexity of the model (Charlton and Fotheringham, 2009b). The Cross-Validation (CV) method determines the extent of the kernel by predicting the sum of squares for the prediction residuals of the observations (Tarpey, 2000). The Bandwidth Parameter determines the extent of the kernel by a fixed distance or a fixed number of neighbors that are specified by the user. If the Kernel Type is set to fixed, then the coefficient will be a distance which is in the same units as the coordinate system being used by the feature class. In other words, if the coordinates are in meters, this distance will also be in meters. If the Kernel Type is set to adaptive, then the extent of the kernel becomes a count of the nearest observations to be included. A GWR model can be complex in that it depends not just on the number for variables in the model, but also on the bandwidth. This interaction between the bandwidth and model complexity suggests
that the AICc method is the most preferable for an adaptive kernel (Charlton and Fotheringham, 2009b).

After executing the GWR tool in ArcGIS Desktop 10, it generates a supplemental table of statistics and an output feature class of model predictions. There are seven diagnostic results generated within the supplemental table of statistics (Table 3.3), which provides a means for further assessing the overall performance of a model.

Table 3.3: Example of supplemental table of statistics from GWR tool.
Source: ESRI (2011)
ESRI (2012j) has provided the following descriptions for each of these values:

- **Neighbors**: This represents the number of neighbors used for each local estimation and it is perhaps the most important parameter for a GWR model. It controls the degree of smoothing for the model.

- **ResidualSquares**: This value is the sum of the squared residuals in the model, which is the difference between all observed values and their respective estimated values of the model. As this measure become smaller, the closer the predicted values by the GWR model fit the observed values.

- **EffectiveNumber**: This value reflects the tradeoff between the variance of the fitted values and the bias in the coefficient estimates. As the bandwidth approaches infinity, the geographic weights for every observation approach 1, and the coefficient estimates become very close to those for a global OLS model. Conversely, as the bandwidth approaches zero, the geographic weights for every observation approach zero and the coefficient estimates exhibit large variances.

- **Sigma**: This value is the square root of the normalized residual sum of squares, where the effective degrees of freedom of the residual. This is the estimated standard deviation for the residuals. Sigma is used for AICc computations.

- **AICc**: This is a measure of model performance. The AICc value also accounts for the complexity of a model and helps provide a better fit to the observed data. It is not an absolute measure of fit however; it is useful for comparing models with different combinations of explanatory variables as long as they use the same dependent variable. When comparing the two models, the model with the lower AICc is considered to be better.
• **R2**: This R-squared value is a measure of good of fit. Its value can range between 0.0 and 1.0. Higher values are preferable and it can be interpreted as the proportion of dependent variable variance accounted for by the regression model. The denominator for the R2 computation is the sum of squared dependent variable values. Therefore, any additional explanatory values can alter the numerator and give the impression of improvement in model fit that may not be genuine.

• **R2Adjusted**: Due to the aforementioned problem described for R2, the calculations for the adjusted R-squared value normalize the numerator and denominator by their degrees of freedom. This compensates for the number of variables in a model and, as a result producing an Adjusted R2 value that is almost always smaller than the R2 value.

The other output from the GWR tool is an output feature class that includes the model predictions. An example of the attribute table for the output feature class is shown in Table 3.4. It is a combination of the original feature class attributes along with the additional statistical data that is generated by the tool. By default, this output feature class renders the standard regression residuals of the model. The over and under predictions for a well-specified regression model will appear to be randomly distributed throughout the feature class. Any clusters or dispersions are evidence that at least one key explanatory variable is missing from the model. If this occurs, the patterns of the GWR model residuals may provide insight into what variables are missing. Even if the residuals do appear to be spatially random, the Spatial Autocorrelation (Moran’s I) tool should be used on the regression residuals just to ensure that they are in fact distributed randomly.
If the results indicate either clustered or dispersed features, then the GWR model could be misspecified.

Table 3.4: Example of output feature class attribute table from GWR tool.
Source: ESRI (2011)
CHAPTER 4: ANALYSIS AND RESULTS

4.1 Scatter Plots

The scatter plots for the transformed dependent and independent variables are shown in Figures 4.1 through 4.7, revealing both positive and negative correlations. *Coca Cultivation Density*, *Displaced Persons Rate*, *Standard of Living Index*, *Terrain Ruggedness Index*, *FARC Armed Actions Rate*, and *Public Force Armed Actions Rate* all express positive correlations with *Homicide Rate*. Of these variables, *Public Force Armed Actions Rate* shows the strongest correlation with *Homicide Rate*. The only independent variable that expressed a negative correlation with *Homicide Rate* was the *Presidential Election Participation Rate*.

![Colombia Coca Cultivation Density 2005](image)

*Figure 4.1: Scatter Plot – Coca cultivation density (2005).*
*Data Source: SIMCI (2005), SIGOT (2012)*
Figure 4.2: Scatter Plot – Presidential Election Participation Rate (2006). Data Source: Registraduría Nacional del Estado Civil (2006), SIGOT (2012)

Figure 4.3: Scatter Plot – Displaced Persons rate (2005). Data Source: SIGOT (2012)
Figure 4.4: Scatter Plot – Standard of Living index (2005). Data Source: SIGOT (2012)

Figure 4.5: Scatter Plot – Terrain Ruggedness Index.
Data Source: CGIAR-CSI (2008), SIGOT (2012)
Figure 4.6: Scatter Plot – Rate of armed actions initiated by the FARC (2005). Data Source: SIGOT (2012)

Figure 4.7: Scatter Plot – Rate of armed actions initiated by the Public Force (2005). Data Source: SIGOT (2012)
4.2 Global Moran’s I

After calculating the Global Moran’s I (GMI) with the Spatial Autocorrelation (Moran’s I) tool, a report (Figure 4.8) was generated containing a Moran’s Index, a p-value, and a z-score. These values dictate whether to accept or reject the null hypothesis that the homicide rate values are randomly distributed throughout the study area. For this study, the report indicated a Moran’s Index of 0.146017, a z-score of 25.066872, and a p-value of 0.000000. Given the z-score of 25.066872, there is less than 1% likelihood that the homicide rate patterns are the result of random chance. The null hypothesis is therefore rejected because the p-value is statistically significant and the z-score is positive, indicating that the spatial distribution of high and low homicide rates in the dataset are more spatially clustered than would have been expected if the underlying spatial processes were random (ESRI, 2012a).
4.3 Local Moran’s I

The Local Moran’s Index (LMI) was calculated using the Cluster and Outlier Analysis (Local Moran’s I) tool. The results indicate that there are clusters of municipalities with similar homicide rates in the northern, southern, western, and central regions of Colombia. The map of the output feature class (Figure 4.9) clearly shows the existence of municipalities with statistically significant (0.05 level) clusters of high values as HH and statistically significant (0.05 level) clusters of low values as LL. As
shown in Figure 4.9, several statistically significant (0.05 level) outliers as HL were also observed where high value features are surrounded by low value features. Likewise, a few statistically significant (0.05 level) outliers as LH were identified where low value features are surrounded by high value features.

The map in Figure 4.9 shows that high-value homicide rate clusters are located in the Valle del Cauca, Tolima, Antioquia, Chocó, Caldas, Risaralda, and Quindio departments of the western region. This area extends from the Cordillera Occidental through the Cordillera Central mountain ranges. It contains several urban areas and stretches from two of the largest cities in Colombia; Medellín and Cali. There are also high-value homicide rate clusters in the Caquetá, Meta, Guaviare, and Putumayo departments of the southern region. This cluster extends from the eastern edge of the Cordillera Oriental and contains mostly rural areas with relatively low population density.

The map in Figure 4.9 also shows that there are low-value homicide rate clusters in the Córdoba, Sucre, Bolívar, and Magdalena departments of the northern region. This area contains several urban areas and major cities as well, however, the terrain is far less rugged as it lies beyond the reach of the principal mountain ranges. There are low-value homicide rate clusters in the Boyacá, Santander, and Cundinamarca departments of the central region as well. This area extends along the Cordillera Oriental mountain range and also contains several major cities and urban areas. The city of Bogotá, which happens to reside in this area, is one of the largest cities in Latin America and the national capital of Colombia.
Figure 4.9: Local Moran’s Index – Cluster / Outlier Type. Data Source: SIGOT (2012)
4.4 Global Model Results (OLS)

After executing the Ordinary Least Squares (OLS) tool, it generated an output feature class (Figure 4.10) of the OLS model’s standard residuals. It also generated an output table (Table 4.1) of model coefficients, standard errors, and probabilities for each explanatory variable. In addition, it generated a supplemental table (Table 4.2) with several diagnostic values to evaluate the overall performance of the global OLS model. The $R^2$ and adjusted $R^2$ values of this diagnostic statistics table indicate that the independent variables account for approximately 30% of the homicide rates throughout the study area.

### Table 4.1: OLS – Coefficients and Standard Errors.

<table>
<thead>
<tr>
<th>OBJECTID*</th>
<th>Variable</th>
<th>Coef</th>
<th>StdError</th>
<th>t Stat</th>
<th>Prob</th>
<th>Robust SE</th>
<th>Robust t</th>
<th>Robust Pr</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>1.771467</td>
<td>0.205781</td>
<td>8.608506</td>
<td>0</td>
<td>0.218621</td>
<td>0.196531</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>PRES_VOT_RT00</td>
<td>-0.01094</td>
<td>0.002893</td>
<td>-3.782273</td>
<td>0.000177</td>
<td>0.002525</td>
<td>-3.739755</td>
<td>0.000203</td>
</tr>
<tr>
<td></td>
<td>LIVING_INDEX5</td>
<td>0.023912</td>
<td>0.009556</td>
<td>2.535926</td>
<td>0</td>
<td>0.002643</td>
<td>9.646079</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TIR_MEAN</td>
<td>0.008643</td>
<td>0.000779</td>
<td>7.53181</td>
<td>0</td>
<td>0.000749</td>
<td>8.082951</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>CC_SQRT</td>
<td>0.649687</td>
<td>0.149463</td>
<td>4.432677</td>
<td>0</td>
<td>0.003427</td>
<td>4.088235</td>
<td>0.000064</td>
</tr>
<tr>
<td></td>
<td>DISPLACED_LOG_P1</td>
<td>0.041316</td>
<td>0.017276</td>
<td>2.41657</td>
<td>0.01637</td>
<td>0.001624</td>
<td>2.492833</td>
<td>0.012044</td>
</tr>
<tr>
<td></td>
<td>FFNM_AA_LOG_P1</td>
<td>0.161686</td>
<td>0.021802</td>
<td>7.424131</td>
<td>0</td>
<td>0.001536</td>
<td>7.516721</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>FARC_LOG</td>
<td>0.087054</td>
<td>0.026399</td>
<td>3.297668</td>
<td>0.06103</td>
<td>0.002688</td>
<td>3.115123</td>
<td>0.000982</td>
</tr>
</tbody>
</table>
Figure 4.10: OLS – Standard Residuals. Data Source: SIGOT (2012)
Table 4.2: OLS – Diagnostic statistics.

The coefficients and p-values of the independent variables in Table 4.1 provide the insight into their relationships with the dependent variable of the OLS model. The *Coca Density* variable expresses a positive coefficient of 0.549067 and a p-value of 0.000267. This relatively high p-value suggests that this variable is one of the less significant variables of the model. The *Presidential Voting Rate* variable expresses a negative coefficient of -0.01094 and a p-value of 0.000177. This variable expresses a relatively high p-value as well and suggests that it is not as significant as some of the others. The *Displaced Persons Rate* variable expresses a positive coefficient of 0.041316 and a p-value of 0.01697. This is the highest p-value among the independent variables and suggests that it is the least significant variable of the model. The *Standard of Living Index* variable expresses a positive coefficient of 0.023912 and a p-value of 0.000000.
This low p-value suggests that this variable is highly significant and one of the most important independent variables of the model. The *Terrain Ruggedness Index* expresses a positive coefficient of 0.006043 and a p-value of 0.000000. This p-value is very low as well and suggests that this variable is also one of the most important independent variables of the model. The *FARC Armed Actions Rate* variable expresses a positive coefficient of 0.087054 and a p-value of 0.00103. The p-value for this variable indicates that while it is significant, it may not be one of the most important independent variables of the model. The *Public Force Armed Actions Rate* variable expresses a positive coefficient of 0.16186 and a p-value of 0.000000. The low p-value for this variable indicates that it is one of the most important independent variables of the model.

The histogram in Figure 4.11 shows the distribution of the residuals from the OLS model. These residuals have a mean of zero and appear to be normally distributed, which confirms that the OLS model is properly specified. This indicates that the data transformations were successful in alleviating any bias that may have potentially affected the model.
Figure 4.11: OLS – Histogram of OLS Residuals. Source: ESRI (2012f)

The map of the OLS model standard residuals in Figure 4.10, however, shows that there are several clusters of standard residuals with red colors indicating underpredictions and blue colors indicating overpredictions. According to ESRI (2012d), these clustered patterns suggest that the model could be misspecified and that there may be one or more key explanatory variables still missing from the model (also indicated by the relatively low adjusted $R^2$). After conducting a GMI analysis of the OLS residuals, the results (Figure 4.12) confirm that the observations are spatially autocorrelated with a z-score of 13.35093 and a p-value of 0.000000. The analysis was then taken a step further by conducting an LMI analysis on the OLS residuals. The resulting map (Figure 4.13) confirms two rather large areas expressing high-value (underpredicted) residuals from the OLS model. This confirms that the model could be misspecified and that some key
explanatory variables may be needed to better explain the homicide rates of these two areas. In an effort to alleviate the spatially varying nature of the data, a geographically weighted regression (GWR) analysis was conducted thereafter.

Figure 4.12: Global Moran’s Index of OLS residuals. Source: ESRI (2012f)
Figure 4.13: Local Moran’s Index of OLS residuals. Data Source: SIGOT (2012)
4.5 Local Model Results (GWR)

The supplemental diagnostic statistics table for the GWR model in Table 4.3 demonstrates an improvement of approximately 15-20% over the Ordinary Least Squares (OLS) global model. The $R^2$ and adjusted $R^2$ values are 0.528666 and 0.458805 respectively. These values indicate that the explanatory variables of the GWR model account for roughly 50% of the homicide rates throughout the study area. The AICc value, another measure of model performance, is typically desired to be the lowest possible value between comparable models. For the OLS model, the AICc value is 1995.311351 while for the GWR model it is 1831.181733. This is a further indication that the GWR model is better suited for this study. Subsequently, a GMI analysis was performed on the residuals of the GWR model. The results in Figure 4.14 reveal a z-score of -0.608214 and a p-value of 0.543045. This p-value indicates that the null hypothesis cannot be rejected, that the residuals are randomly distributed throughout the study area, and that the GWR model is better specified than the OLS model (ESRI, 2012a).
Table 4.3: GWR – Supplemental diagnostic statistics. Source: ESRI (2012f)

Figure 4.14: Global Moran’s Index of GWR residuals. Source: ESRI (2012f)
The GWR analysis also generated an output feature class that contains the predicted values, local $R^2$ values, coefficients, residuals, standard errors, and coefficient standard errors for each feature within the attribute table. The local $R^2$ values represent how well each of the unique local regression models fit their surrounding features (ESRI, 2012j). These values are mapped in Figure 4.15 to illustrate the dynamic strength of the GWR model throughout the study area. The local $R^2$ values appear to be relatively weak in the western and upper-central regions of Colombia, indicating some key explanatory variables may still be missing in these areas, such as drug trafficking, racial heterogeneity, etc. It is also possible, while unlikely, that these low $R^2$ values are merely the result of random variations in homicide rates.
Figure 4.15: GWR – Local $R^2$ values. Data Source: SIGOT (2012)
The coefficient values for each explanatory variable are also rendered in Figures 4.15 - 4.23 to demonstrate the localized correlations between homicide rates and each respective independent variable. The Coca Density Coefficient map (Figure 4.16) shows that homicide rates tend to be positively correlated with coca cultivation density in the western region of Colombia. This positive correlation appears to be strongest in the Tolima and Valle de Cauca departments. Additionally, coca cultivation density tends to be negatively correlated with homicide rates in parts of the northern, southern, and eastern regions of Colombia. For some of these areas, there were simply no recorded instances of coca cultivation. For those areas that did have recorded instances of coca cultivation, this negative correlation may illustrate how coca cultivation in rural areas is often beyond the authority and enforcement of the state.

The Presidential Election Participation Coefficient map (Figure 4.17) shows that homicide rates tend to be negatively correlated with voter turnout rates throughout most of Colombia. The correlation appears to be strongest around the Valle del Cauca, Tolima, Quindío, and a few other departments in the northern and upper-central region. These departments represent some of the more urbanized areas of the country and include several major cities. There are also several municipalities in eastern region of Colombia where voter turnout rates tend to positively correlate with homicide rates. These municipalities appear to extend from the Cordillera Oriental into the more rural lowlands of the east.
Figure 4.16: GWR – Coca Density Coefficient.
Data Source: SIMCI (2005), SIGOT (2012)
Figure 4.17: GWR – Presidential Election Participation Coefficient.
The *Displaced Persons Coefficient* map (Figure 4.18) shows that homicide rates tend to be positively correlated with rates of displaced persons throughout much of the northern, southern, and eastern regions of Colombia. This correlation seems to be strongest around the Boyacá, Cundinamarca, and Meta departments. Conversely, it appears that the rate of displaced persons tends to be negatively correlated with homicide rates in several western departments including Valle del Cauca, Tolima, Risaralda, and Quindío. This may be due to the fact that many displaced persons see this region as a refuge despite its emergence as one of Colombia’s most dangerous areas (Romero, 2007).

The *Standard of Living Index Coefficient* map (Figure 4.19) shows that homicide rates tend to be positively correlated with the standard of living index throughout much of Colombia. This correlation appears to be strongest in the western region of the country around the Cauca, Valle del Cauca, Quindío, Chocó, and Tolima departments. While this relationship seems somewhat counterintuitive, it may be an indicator of violence that stems from income inequality in various regions of the country. Additionally, the standard of living index tends to be negatively correlated with homicide rates in the central region of Colombia. This correlation appears to be strongest in the Cundinamarca, Santander, and Boyacá departments that surround the national capital.
Figure 4.18: GWR – Displaced Persons Coefficient. Data Source: SIGOT (2012)
Figure 4.19: GWR – Standard of Living Index Coefficient. Data Source: SIGOT (2012)
The *Terrain Ruggedness Coefficient* map (Figure 4.20) suggests that homicide rates are positively correlated with terrain ruggedness indices in the northern and parts of the western regions of Colombia. This correlation appears to be been strongest in the Chocó, Antioquia, and Córdoba departments along with some of the municipalities in and around the Tolima department. In parts of the southern and upper-central regions of Colombia, however, homicide rates are negatively correlated with terrain ruggedness indices.

The *FARC Armed Action Coefficient* map (Figure 4.21) shows that the homicide rates are positively correlated with the armed actions initiated by the FARC throughout most of Colombia. This correlation seems to be strongest in the Nariño, Córdoba, Cundinamarca, and several other departments in the northern region. Conversely, homicide rates are negatively correlated with armed actions initiated by the FARC in some of the more central areas of Colombia. Some of these departments include Huila, Casanare, and Boyacá.

The *Public Force Armed Action Coefficient* map (Figure 4.22) suggests that homicide rates are positively correlated with the armed actions initiated by the Public Force throughout nearly all of Colombia. This correlation is strongest in and around the Meta, Huila, Caquetá, and Guaviare departments in the southern region of the country.
Figure 4.20: GWR – Terrain Ruggedness Index Coefficient. Data Source: CGIAR-CSI (2008), SIGOT (2012)
Figure 4.21: GWR – FARC Armed Action Coefficient. Data Source: SIGOT (2012)
Figure 4.22: GWR – Public Force Armed Actions Coefficient.
Data Source: SIGOT (2012)
The results of the GWR local model of Colombian violence provide valuable insights into the relationships between homicide rates and the spatial varying nature of the social, economic, political, and physical aspects of Colombia. While not true throughout the entire study area, the explanatory variables tend to reflect the same positive and negative relationships that were identified by the OLS analysis. Overall homicide rates tend to express positive correlations with such independent variables as *Coca Cultivation Density, Displaced Persons Rate, Standard of Living Index, Terrain Ruggedness Index, FARC Armed Actions Rate*, and the *Public Force Armed Actions Rate* while expressing negative correlations with the *Presidential Election Participation Rate*. 
CHAPTER 5: CONCLUSIONS AND FURTHER RESEARCH

5.1 Conclusions

The first objective of this thesis is to determine whether there are any discernible patterns in the geographic distribution of homicide rates across Colombia at the municipal level and if and how these patterns are distributed in geographic space. According to the Local Moran’s Index in Figure 4.9, there are indeed clusters of both high- and low-value homicide rates in several regions of the country. The major clusters of high-value homicide rates are located in the western and southern regions of Colombia. While these two regions appear to experience similarly high-levels of violence, they are considerably different in several ways.

The second objective of this thesis is to determine what combination of statistically significant predictors, if any, generates acceptable regression models for predicting the distribution of homicide rates throughout Colombia. Having discovered that there are indeed discernible patterns in the geographic distribution of homicide rates across Colombia, prior studies and other relevant information were assessed to identify the potential predictors that would most strongly explain the patterns of violence. These independent variables were then evaluated using an exploratory regression analysis to determine the best combination of statistically significant predictors of Colombian homicide rates. In the end, a global OLS model and a local GWR model were estimated with the final list of independent variables, including Displacement Rate, Coca Cultivation Density, Presidential Election Participation Rate, Standard of Living Index,
FARC Armed Actions Rate, and Public Force Armed Actions Rate as the independent variables.

The relatively low adjusted R² value (0.302492) of the global OLS model and the spatial clustering of its residuals indicate that some key explanatory variables could be missing from the model. The local GWR model, however, generates a dramatically improved adjusted R² value (0.458805) and does not exhibit spatial clustering among its residuals. In addition, the AICc value also decreases from 1995.311351 with the OLS model to 1831.181733 with the GWR model, suggesting that the GWR model is a better fitting model.

Regions that contained densely cultivated coca fields were selected to represent the territorial extents of guerrilla organizations that are outside the reach of government control and policy enforcement. Therefore, the Coca Cultivation Density is one of the selected predictors to represent the state presence in various regions of Colombia. As mentioned earlier, violence in modern-day Colombia is often perceived as a simple product of drug trade (Bergquist et al., 1992). Some scholars even draw a direct connection between illicit drugs and violence and believe that the violence would inexorably dissipate if illicit drugs were eliminated (Holmes et al., 2006). The elimination of the illicit drug trade, however, may only cripple these guerrilla organizations. It is far more likely that other revenue sources would be appropriated to compensate for any financial losses.

Unlike the western region of Colombia, coca cultivation is far more prevalent in the southern region (Figure 3.6). The Coca Density Coefficient map in Figure 4.16 indicates that there is a negative correlation between coca cultivation and homicide rates.
in the south-eastern region. While this may be true for this particular area, it does not hold true for all areas of the country. Coca cultivation density is positively correlated with homicide rates in the northern and south-western regions of Colombia. For that reason, the claim by Holmes (2006) that there is minimal geographic overlap between coca production and violence may only hold true for certain regions of the country or perhaps only at a more aggregated scale.

In addition to the density of coca cultivation, the state presence is also represented by the *Presidential Election Participation Rate* for the 2006 presidential election. The voter turnout rate tends to reflect the legitimacy and implicit endorsement of the current system of government. It can also be an important indicator for the effectiveness of government policy and enforcement. Overall, the results of the GWR analysis indicate that there is typically a negative relationship between the presidential election participation rates and homicide rates, that is, the higher the presidential election participation rates the lower homicide rates. While the highest participation rates occurred near major cities and urban areas (Figure 3.7), the coefficients (Figure 4.17) express a positive correlation with homicide rates in many of the more rural areas of the eastern region where voter turnout is relatively low. This suggests that political participation increases homicide rates in the rural areas of the eastern region while decreasing homicide rates elsewhere as people become more politically active. Since there is relatively low voter turnout in this eastern region, it may be an indication of little or no state presence. According to Waldmann (2007), this often encourages support for guerrilla organizations that then establish and provide basic law and order even in a relative vacuum of authority.
The Displacement Rate was selected as a predictor to represent the social conditions throughout Colombia. Displacement rates express a positive relationship with homicide rates throughout most of Colombia except for a particular area in the western region and a few municipalities in the northern and eastern regions (Figure 4.17). According to Romero (2007), the people displaced by armed conflicts in the countryside see the western region as a refuge despite its emergence as one of Colombia’s most dangerous areas. He further states that according to the federal government, the refugee population is mostly Afro-Colombians from rural areas.

The map in Figure 5.1 illustrates the territorial extent of the Afro-Colombian and indigenous populations between 1982 and 2005. It shows that the territory of the Afro-Colombian population tends to stretch around the homicide rate hotspot of the western region and extend along the Pacific coast, where the GWR model performs relatively poorly. This may suggest that some key explanatory variables related to ethnic heterogeneity are missing from the model. Likewise, there are settlements of the indigenous population in and around this same area. It is therefore intriguing that the clusters of high-value homicide rates are immediately adjacent, and with minimal overlap, to the Afro-Colombian and indigenous populations. This particular phenomenon, with regard to the indigenous population, appears to occur in the southern and eastern regions of Colombia as well. Unfortunately, ethnic heterogeneity indices are not included in this study as they were unavailable at the municipal level for 2005.
Figure 5.1: Distribution of Afro-Colombian and Indigenous populations, High- and Low-value LMI clusters of homicide rates. Data Source: IGAC, SIGOT (2012)
The *Standard of Living Index* was selected as a predictor to represent the economic conditions throughout Colombia. Surprisingly, this variable exhibits a rather counterintuitive relationship. Municipalities with high-value homicide rates tend to also possess some of the highest standards of living. This finding is difficult to explain, however it may indicate that the standard of living throughout Colombia may be a proxy variable for income inequality, an independent variable that is consistently and positively correlated with homicide rates in several cross-national studies (LaFree, 1999, pp. 127-132). As seen in Figure 4.19, the one area of Colombia that does not express a positive correlation with homicide rates appears to be in and around the national capital of Bogotá. This may identify the one area of Colombia where increases in the standard of living do not, in turn, increase the level of income inequality. While it would be interesting to actually incorporate income data, it was regrettably unavailable at the municipal level for 2005.

The *Terrain Ruggedness Index* was selected as a predictor to represent the physical geography throughout Colombia. This variable exhibits a relatively strong correlation among various regions of the country. The map in Figure 4.20 clearly marks the north-western region of Colombia where a more rugged terrain is conducive to increased rates of homicide. This coincides with the notion proposed by Baechler (1999) that the susceptibility of war, and presumably violence, is quite significant in mountain regions. Pickering (2011) also suggested that mountainous areas, like the principal ranges of Colombia, are imbued with certain qualities that encourage conflict. The southern region, however, expresses a negative correlation with homicide rates. Fafchamps and
Moser (2002) may therefore be correct in that the isolation of these areas serve as a sanctuary for rural populations and guerrilla forces alike.

The *FARC Armed Actions Rate* and the *Public Force Armed Actions Rate* were selected as predictors to represent the political stability throughout Colombia. For reference, the approximate territorial extents of the FARC, the ELN, and the paramilitaries in 2005 are illustrated in Figure 5.2. A majority of the FARC territory is in the south-west with other smaller territories scattered across the northern and central regions of Colombia. The ELN territory is mostly in the northern region with some areas in the western region as well. The paramilitaries are, understandably, scattered evenly throughout the northern, western, and southern regions to combat and weaken the guerrilla organizations.
Figure 5.2: Territorial extent of the FARC, ELN, and paramilitaries, 2005.
Source: UN Environmental Programme (2006)
The *FARC Armed Actions Coefficient* map (Figure 4.21) presents a positive correlation with homicide rates throughout nearly every region in Colombia except for a small area in the southern region. This largely positive correlation is also expressed in the *Public Force Armed Actions* map (Figure 4.22) except for a small area in the southwestern region. It is interesting, however, that these two maps appear to be the inverse of one another. Areas where the *FARC Armed Actions Rate* expresses a negative coefficient with homicide rates are the same areas where the *Public Force Armed Actions Rate* expresses a highly positive coefficient with homicide rates. This trend seems to hold true throughout the rest of the study area as well and could explain why the primary footholds of these groups are relatively stable and devoid of violence.

The metadata for the 2005 homicide rate data (SIGOT, 2012) actually notes that the different rates between municipalities may have been affected by the security and armed conflict within them. Of the independent variables, the *Displaced Persons Rate*, the *FARC Armed Actions Rate*, and the *Public Force Armed Actions Rate* are the variables most directly related to the state of security and armed conflict within the municipalities. It may be reasonable to conclude that the internal conflict is often a byproduct of opposing political and military ideologies. While an increase in homicide rates would understandably be an inherent and direct result of armed actions, the rate of displacement might be a more subtle indicator of municipal stability.

Regions of Colombia subjected to extended periods of conflict likely generate more insecurity and incertitude among the people. Those preferring the protections of the state may have been displaced and relocated to more secure areas outside the reach of
guerrilla control. As guerrilla organizations expand their territories, however, their escalation of crimes against the civil population becomes an inexpensive and effective strategy for furthering their ambitions and strengthening their areas of control. Consequently, the mere threat of impending violence, which often drives people from their land, may have become a reality reflected by the high-value homicide rates of the regions.

Working as a spatial magnifier, the Geographically Weighted Regression (GWR) local model is proven to be a more suited method to dissect and quantify spatial patterns of homicide rates across the study area. Estimating a separate and unique regression equation for each municipality, the GWR local model accounts for the spatial variations that were otherwise not easily revealed by the Ordinary Least Squares (OLS) global model. The strength of the GWR local model, however, varies throughout the study area and suggests that some of the homicide rates were either random or that certain important explanatory variables were not included, especially where the current model is the weakest. Nevertheless, those municipalities that express relatively low $R^2$ values actually provide valuable information as well. One can infer that these areas are exposed to circumstances that are unique among the Colombian municipalities. The violence in these areas may be the product of organized criminal activities and logistics that are simply too difficult to quantify with a selected a few explanatory variables. The entire region of low $R^2$ values could potentially represent the unofficial overlap between interurban conflict and drug-trafficking. As a result, a more intensive research approach may be necessary to fully elucidate the homicide rates of these particular areas.
Fortunately, it is these spatial variations in the model’s strength that reveal just how valuable the GWR local model truly is. Even municipalities that exhibit nearly identical conditions can be the result of drastically different geographic circumstances, providing a truly spatial assessment of a given set of features. This quality of the GWR local model allows for more accurate predictions throughout a study area where features are not entirely homogenous. Therefore, based on the results of this study, it is safe to conclude that the local GWR model is an undeniably valuable method of the assessment for criminal studies.

5.2 Future Research

In 2005, a general census was conducted by the Departamento Administrativo Nacional de Estadística (DANE) in an effort to acquire more accurate and up-to-date demographic and socioeconomic data for Colombia. As a result, additional datasets from a variety of other organizations were generated as a compliment to the general census datasets. Data collected in the year 2005 are therefore the most recent and complete representation of the social, economic, and political nature of Colombia. Until another general census is conducted, studies of subsequent years will likely include estimations rather than actual figures.

With regard to this study in particular, future research could focus on the regions of the study area that seem to be more inexplicable than others. Having applied an extensive research approach, this study is concerned with identifying the common properties and patterns of Colombian homicide rates. It is conducted at a national scale that includes descriptive statistics and numerical analyses that can be tested by merely
replicating the data and methods. The models developed in this study provide important insights into the general characteristics of spatially varying homicide rates and their dynamic relationships with certain explanatory variables. In particular, the GWR model is useful for examining specific places and problems but is of somewhat limited use for prediction purposes due to the overall low adjusted $R^2$ values in some areas. For this reason, additional key explanatory variables, if available, must be identified and included in future models, such as ethnic heterogeneity, income inequality, drug trafficking, gang activities, etc. As an example, the strength of the GWR local model, as seen in Figure 4.15, appears to be relatively weak in the western region and parts of the central region of Colombia. To explain this phenomenon, it may be necessary for future studies to exercise a more intensive research approach. At a relatively local scale, this would be more appropriate for explaining the conditions of a particular place rather than conditions that vary across space.

This study also includes the 2006 Colombian presidential election as an explanatory variable to depict the state presence throughout the country. Another presidential election was also conducted in 2010. If a future study were able to acquire datasets that were also from this year, it would be interesting to determine whether the voter participation rate continued to express the same relationship with homicide rates that were found in this study. Official datasets, however, may be difficult to obtain as estimations would likely be the only available data for several of the explanatory variables used in this study.

Another area of future research could focus on the potential decriminalization of illicit crop cultivation in Colombia. According to Colombia Reports (2012), an
independent news organization based in Medellin, the Colombian House of Representatives passed the first draft of a bill that seeks to legalize illicit crops on May 9th, 2012. The initiative calls for the decriminalization of growing plants such as coca, marijuana and opium poppies. While the bill would in fact legalize the cultivation of these plants, the processing and trafficking of drugs would remain subject to criminal sentencing. The article also notes that the Minister of Justice claims the government is strongly opposed to the proposed legislation, and that it is not yet time to make a policy change in the fight against drugs. Nevertheless, if the government were to decriminalize the cultivation of the currently illicit crops, it would provide an incredibly valuable research opportunity.

As noted earlier, the United States has viewed the violence in modern-day Colombia as a simple product of drug trade, a view that is promoted by the U.S. media and informs U.S. government policy and support (Bergquist et al., 1992). It would therefore be fascinating for future studies to determine whether the decriminalization of illicit crops has any effect on the homicide rates throughout the country. While guerrilla organizations often depend on the revenue generated by the illicit drug trade, this would unlikely affect their financial stability so long as other countries continue to prohibit various narcotics. Be that as it may, decriminalization may provide protection for farmers that were formerly overlooked or prosecuted by local authorities. This additional protection and security could potentially alleviate the threat of violence and displacement that often plagues the more rural areas of Colombia.
LITERATURE CITED


http://www.economics.ox.ac.uk/Members/marcel.fafchamps/Homepage/crimepo.pdf


http://geoportal.igac.gov.co/mapas_de_colombia/IGAC/Tematicos/34813.jpg


http://oasportal.policia.gov.co/portal/page/portal/UNIDADES_POLICIALES/Dir ecciones_tipo_Operativas/Direccion_de_Investigacion_Criminal/Documentacion/


SIMCI (Sistema Integrado de Monitoreo de Cultivos Ilícitos). (2005, December 31).
Publicaciones SIMCI. Retrieved November 11, 2012, from UNODC:


http://www.biesimci.org/SIMCI/SIMCI.html


