Bus Stops and Crime:
Do Bus Stops Increase Crime Opportunities in Local Neighborhoods?

By Sung-suk Violet Yu

A dissertation submitted to the

Graduate School-Newark

Rutgers, The State University of New Jersey

In partial fulfillment of requirements

for the degree of

Doctor of Philosophy

Graduate Program in Criminal Justice

Written under the direction of

Doctor Ronald V. Clarke

And approved by

Doctor Marcus Felson

Doctor Michael G. Maxfield

Doctor Jerry Ratcliffe

Newark, New Jersey
May, 2009

Abstract

Bus Stops and Crime: Do Bus Stops Increase Crime Opportunities in Local Neighborhoods?

By Sung-suk Violet Yu

Dissertation Director:

Doctor Ronald V. Clarke

Mass transit is often thought to be dangerous, eliciting concerns for personal security when waiting for and traveling on mass transit. One of the first steps in preventing crime in mass transit environments is to obtain accurate figures of crime. There are two mechanisms linking mass transit to crime: extension of offenders' journey-to-crime and development of crime attractors and generators. The first mechanism has been tested by comparing crime patterns when a new light rail system was expanded into new areas; these studies reported that new transit systems do not change crime patterns. Testing the second mechanism, whether mass transit enhances crime opportunities in the neighborhoods by attracting targets and offenders alike, however, faces more challenges; testing the impact of small places on larger areas such as transit stations on crime is fraught with difficulties due to the interactions with their surrounding environments. These difficulties are even more relevant for bus stops located without clear demarcation or controlled access.

Using frameworks of routine activity, crime pattern, and rational choice theories, the present study set out to investigate the relationships between bus stops and crime using Newark, New Jersey as the study area. To delineate the impact of bus stops on their surroundings from other possible covariates, the existence of commercial activities in the areas was also examined. The present research study examined five crime types: robbery, aggravated assault, motor vehicle theft, theft from motor vehicle, and burglary.

To better understand the impacts of spatial aspects of the data, several data analyses methods were utilized. First, the study examined the magnitude and structure of spatial dependence in the data. Second, spatial process models were performed and compared with the OLS regression results to examine the impacts of spatial aspects on the regression results. Third, to address non-normality and spatial dependence of the data, the count response model was run by adding spatial lag as one of the predictors.

The data analysis results showed that both bus stops and commercial establishments were associated with increased crime in the neighborhoods. Among the business types, some of them displayed more robust relationships than others. For instance, the category of food store was almost always significant to increased crime whereas banks were not statistically significant across crime types and regression methods.

Considering the fact that fear of crime plays a strong role as actual risk of crime in making travel decisions, it is suggested that the physical and social incivilities should be analyzed by performing an environmental survey using case-control design. In addition, future research should incorporate both short-term and long-term temporal analysis.

Acknowledgement

I have been fortunate and privileged to have the opportunity to pursue my doctoral degree under advisement and support from distinguished criminologists. Among them, I would like to express my deepest gratitude to the dissertation committee chair Dr. Ronald V. Clarke. Despite his busy schedule traveling around the world lecturing and attending conferences, he has been extremely encouraging and helpful. Having Dr. Clarke as a mentor and my dissertation chair has been a great pleasure and an honor. I would also like to thank Dr. Marcus Felson who does not fail to impress those around him with his creativity and honesty. Whether it comes to life lessons or academic topics, he offers advice without hesitation. I would also like to thank Dr. Michael Maxfield whose warmth and knowledge knows no bound. He shares his knowledge and experiences with those around him. Lastly but not the least, I would like to thank Dr. Jerry Ratcliffe who was integral in tackling some of the methodological problems in this dissertation.

I would also like to thank various members of city agencies. The NJ Transit Geographic Information Systems Team Director Louis Millan has been always responsive and helpful for the past several years. Despite the fact that we have never met, I feel as though I know him quite well. I would also like to thank Newark City Police Director Garry F. McCarthy and Compstat Unit Detective Steven Rodriguez for their prompt responses and cooperation. Detective Rodriguez always responded promptly to my requests and questions. I would also like to express my sincere thanks to Joseph F. Fox, Jr, a principal planner at City of Newark who provided me with the data and shared his in-depth knowledge about Newark, NJ.

There are so many friends and faculty members who I would like to express my gratitude and appreciation but the space does not permit this. Specifically I would like to thank Dr. George Kelling and Dr. Wayne Fisher for helping me get the crime data necessary for the dissertation. Special thanks to Ms. Phyllis Schultze who is always eager to help even though she is always busy with her other tasks. I would also like to thank my friends at Rutgers University who has encouraged me throughout the years.

Finally, I would like to thank my husband Jeffery for his loving support. Special thought and thanks go out to my family in Korea. As the first person who earned a graduate degree in the family, I am constantly reminded of their support, encouragement, and pride when I talk to them. When I think back, I am reminded of how lucky I am with the help and support I have received over the years from my mentors, friends, and family members.

Table of Contents

List of Tables

List of Figures

CHAPTER 1. INTRODUCTION

There has been growing concerns about environments, air pollution, dependence on foreign oil, and traffic congestion. Accompanied with these concerns, there are also increased interests in improving and expanding mass transit into new areas. By providing safe and efficient means of travel at economic costs, mass transit can improve quality of life in urban and rural communities.

The transit systems are an important infrastructure that greatly shape the business locations, and land use patterns (Vogel and Pettinari, 2002). Both the capacity and spatial patterns of transportation network influence and shape the structure of cities and densities of populations (Button et al., 2004). While most of the changes in the cities are usually incremental, significant changes can result from investments in transportation infrastructure. One of the often assumed benefits of public transportation is an economic one: by improving one's ability to commute to work, mass transit is hypothesized to increase job opportunities (Cervero et al., 2002). Indeed, it is thought that the spatial mismatch, the lack of public transportation in poor areas to travel to other areas, is a root cause of unemployment that translates into physical and social isolation leading to intergenerational poverty.

The assumed cyclical relationships between transportation investments and potential economic growths led some city planners, transit officials and citizens to welcome mass transit into their neighborhoods by perceiving it as a device to achieve economic developments in inner-city or under-developed areas (Loukaitou-Sideris, 2000; Loukaitou-Sideris and Banerjee, 2000). Termed as Transportation Oriented Developments (TODs), this is intended to induce economic growth by developing transit villages around the stations. An envisioned idea in TODs is that developing residential and commercial areas in close proximity to mass transit stations would make commuting to work easy and convenient. This convenience would increase employment opportunities as well as expand economic growths by drawing commercial activities to the area.

Table 1. Ridership based on the mode of travel in New Jersey

Mass transit	2001*	2002	2003	2004	2005	2006	2007**
Percent							
Buses	67.9	68.2	67.8	66.5	65.6	64.3	62.8
Commuter rail	27.9	28.0	28.0	28.1	28.3	28.8	29.4
Light rail	4.2	3.8	4.2	5.4	6.1	7.0	7.8
Total number of ridership	215,641	218,600	215,879	226,341	236,324	246,898	254,259

Note: Ridership is defined as the unlinked passenger trips. The table is compiled using data from APTA quarterly reports.
* 2001 ridership is tabulated until December 3 of 2001.
**2007 ridership is tabulated until December 7 of 2007.

Research studies in many countries show that more than a half of travels on mass transit is made on the buses (Crime Concern, 2004). This is also true in the state of New Jersey (NJ). Table 1 shows the ridership by mode of travel reported by NJ Transit Corporation to American Public Transportation Association (APTA). Overall, there was increased ridership over the years. The increased ridership in light rail is primarily due to the expansions of two new light rail systems in NJ; one in Jersey City in 2000 and one in Newark in 2006.

Despite the fact that the bus services make up the majority of the trips made on the mass transit systems, the buses are often seen as the last resort among the available modes of travels (Lusk, 2001). Even the use of aesthetic vehicles for bus services still renders the term “loser cruiser” in referring to the bus services (Ibid). Unfortunately, the buses never garnered the sense of privilege or longing as other modes of public transportation such as commuter train or street cars have done. Instead, the buses are seen as the cheapest and least attractive means of travel serving its utilitarian purposes: moving people from places to places.

Notwithstanding the low regard held by the public, the bus services are practical and economical solutions when expanding mass transit services into new areas (Lusk, 2001). The operating costs of buses are relatively inexpensive and its routes can be easily modified unlike the light rails or commuter trains where the lands for the tracks need to be purchased and the routes are permanently fixed (Lusk, 2001). It is also argued that buses can provide places for social interactions due to its small settings in comparison to trains or subway vehicles. In addition, buses are small enough for the driver to monitor the inside of the vehicle unlike the trains or light rails where the operators are often separated from the passengers. These characteristics of bus services should make buses more desirable than other modes of mass transit.

It is a common reality that many mass transit passengers encounter and tolerate the fear of crime when using mass transit. The fear of crime is influenced by the actual as well as vicarious victimization experiences such as media reports and other people's experiences (Lusk, 2001). The fear of crime is also influenced by social and physical environments in and around mass transit systems. Others, however, may avoid the use of mass transit out of fear altogether (Lusk, 2001; Vogel and Pettinari, 2002; Crime Concern, 2004). In fact, the prior research studies show that those who do not ride on the buses perceive the buses to be two to four times more dangerous than those who ride buses regularly (Lusk, 2001).

The majority of passengers report using mass transit to go to work, school, or shopping (Crime Concern, 2004). For many, use of mass transit is a routine activity and for others it may well be the only mode of transportation available due to their financial or health reasons. For those without alternative means of travel, avoiding the use of mass transit may be impossible or can lead to social exclusion, no access to health care, leisure and other social facilities (Bailey, 2004). The studies show that there are identifiable characteristics of the mass transit passengers (Tucker, 2003; Bailey, 2004; Crime Concern, 2004). Higher proportions of mass transit patrons are ethnic minority, women, the elderly, and the economically disadvantaged.

The level of personal security while walking to and from the bus stop, waiting for the bus, and traveling on the vehicle are all important aspects to consider for the mass transit patrons when they make travel decisions (Vogel and Pettinari, 2002). The first step in addressing concerns for bus stop related crime is a systematic analysis of crime patterns to assess the prevalence of crime (Newton, 2004). This in turn can be used for crime prevention strategies as well as dissemination of the accurate information. For this reason, the primary focus of the present research study is to examine the factors contributing to the crime in Newark, NJ. More specifically, the main focus of the study will be placed on the influences of bus stops and commercial establishments on spatial patterns of crime in Newark, NJ.

Using frameworks of three opportunity theories— routine activity, crime pattern, and rational choice theories, spatial crime patterns in relation to its environment will be examined at micro-level. According to routine activity theory, criminal incidents are most likely when motivated offenders intersect with suitable targets in the absence of capable guardians. Crime pattern theory further informs that the areas where offenders spend their time are more likely to have high levels of crime opportunities. Based on these two theories, the areas served by mass transit can easily become high crime areas due to their exposure to the offender population. At the same time, the areas where many commercial establishments are located may attract a large number of people increasing crime opportunities in the areas. However, the phenomenon of hot dots implies that the target suitability operates on a smaller scale than areas or neighborhoods. In understanding target suitability, rational choice theory informs that offenders make their decisions to commit crime influenced by various situational factors.

Overall, crime opportunity theories are seen to be more relevant to predatory crimes which are thought to be influenced by situational factors rather than individual factors. Thus, the primary focus of present research study is to compare areas with and without bus stops and commercial establishments in relation to predatory crimes. As such, the research questions are focused on the two independent variables –number of bus stops and commercial establishments-- and their impacts on criminal incidents; robbery, aggravated assault, motor vehicle theft, theft from motor vehicle, and burglary. The two research questions are summarized below:

Q1. Do areas with bus stops experience higher numbers of criminal incidents?
Q2. Do areas with commercial establishments experience higher numbers of criminal incidents?

The present research study utilizes several data sources; the 2007 crime data obtained from the Newark City Police Department, Newark, NJ; the 2007 Verizon's YellowBook NJ database as a measure of commercial establishments; the 2007 NJ Transit bus stop locations; and the 2006-2007 Newark parcel and zoning data obtained from the City of Newark. To investigate spatial patterns of crime and their covariates, utilized techniques include spatial econometrics techniques in addition to count response model estimation methods. It also employs Geographic Information Systems (GIS), and Stata in its data analyses.

The present research study begins with a literature review on mass transit and crime in Chapter 2. The literature review includes literature on light rail as well as subway primarily due to the sparse empirical research studies examining the relationships between buses and crime. Many of the literature on buses and crime place their emphasis on the perception of crime rather than the empirical crime patterns. After presenting empirical research findings and covariates of crimes in Chapter 2, Chapter 3 lays out the theoretical frameworks and the research questions. The utilized theories include routine activity theory, crime pattern theory, and rational choice theory in conjunction with situational crime prevention techniques. Presented in Chapter 4 are discussions of the study area, data sources, unit of analysis, and spatial aspects of data. In discussing spatial aspects of the data, the following areas will be covered—spatial dependence, spatial heterogeneity, and spatial weight matrix. Chapter 5 presents the data analysis plan and their intended purposes. Chapters 6 to 8 are the results sections. Chapter 6 contains information on the geocoding results and further examination of the study area. Explanatory spatial data analyses are presented in Chapter 7, divided by point data and polygon data analyses. Chapter 8 contains the main findings of the present study. Included in Chapter 8 are the detailed analyses on the influences of bus stops and commercial establishments on crime. The inferential statistics in Chapter 8 largely relies on three methods. The first planned step is to perform the Ordinary Least Square regression with diagnostics in GeoDa to assess the magnitude and structure of spatial dependence. Based on the diagnostics, either spatial lag or spatial error model will be performed and the results will be compared with the Ordinary Least Square regression results. To address non-normality of the data, the data analysis will move to count response model estimations; negative binominal regression and zero inflated negative binominal regression models. To account for spatial dependence, spatial lags will be added as one of the predictors in the count response models. Chapter 9 concludes the study with the discussions on limitations of the present research study, suggestions for future research, and crime prevention initiatives.

CHAPTER 2. LITERATURE REVIEW

The existing research studies show that busy places often generate higher numbers of criminal incidents. These busy and high crime areas are often served by mass transit. Perhaps for this reason, the areas related to mass transit are frequently perceived to be dangerous places, eliciting fears for personal security. However, in general the mass transit areas are shown to be no more dangerous than other public places (LaVigne, 1996) or even one's own residence (Walker et al., 2006). This applies to subway stations (Kenney, 1987) as well as on the transit vehicles (Loukaitou-Sideris, 1999). In fact, when counting serious crimes only, mass transit is considered to be safer in providing personal security. For instance, less than three percent of serious crimes occurred on the New York City (NYC) Subway (Kenney, 1987). In addition, mass transit often has higher levels of police presence than other public spaces including city streets (Kenney, 1987; Loukaitou-Sideris et al., 2002). Both the research studies and anecdotal evidence show that the police presence in mass transit has increased after the September 11, 2001 tragedy. Nevertheless, the perceived links between mass transit and crime persist.

There are two causal mechanisms that are thought to increase crime in and around mass transit. The first is through the exposure of new areas to the offenders by mass transit (Poister, 1996; Liggett et al., 2003; Sedelmaier, 2003). In this perspective, mass transit may increase accessibility to remote areas or help offenders overcome distances by extending their journey-to-crime (Liggett et al., 2003). This will lead to increased numbers of crime in the previously low crime areas following the introduction of transit service. There are mainly two ways to test this hypothesis: by comparing crime data before-and-after the inception of mass transit and examining offender travel patterns.

The second mechanism linking mass transit to crime is the development of crime attractors and generators around mass transit (Brantingham et al., 1991; Block and Block, 2000; Sedelmaier, 2003). Mass transit may attract offenders and targets to the local areas by providing a variety of legitimate and illegitimate activities. This may also increase the number of targets in the areas. Testing this hypothesis seems to be fraught with practical difficulties. As much as the risks of victimization vary greatly by the nature and characteristics of the places, crime in and around mass transit are products of interactions between the mass transit and its surrounding environments (Block and Block, 2000). For this reason, investigating the impact of mass transit on crime needs to account for their bigger environments. In this section, these two mechanisms are discussed in more detail.

1. Formation of Crime Hot Spots

The plans to expand light rail or subway systems to affluent areas often raise concerns about crime and property values in the neighborhoods (Poister, 1996; Liggett et al., 2003; Sedelmaier, 2003). Indeed, there is a pervasive belief that mass transit will provide access to the inner city offenders to suburban areas where undiscovered attractive crime opportunities are abundant (Poister, 1996). This would imply that without mass transit these relatively low crime areas will continue to be safe from offender population. In the outset, this mechanism seems to be more relevant to light rail where the volume of passengers are greater than the buses. It is often the case that the changes made on the bus services are incremental whereas the changes made on light rail services are more dramatic such addition of new lines or stations.

1.1. Crime Patterns Before and After New Mass Transit System

One of the concerns related to offenders' journey-to-crime lie on the possibility that crime may be pulled toward or transported into the areas surrounding the stops and stations of mass transit (Liggett et al., 2003; Sedelmaier, 2003). These concerns stem from the reasoning that mass transit helps offenders develop the area knowledge while traveling on mass transit (Poister, 1996). By analyzing crime patterns before and after the inception of mass transit, it can be examined whether mass transit has led to formations of new crime hot spots. This kind of research studies on buses are lacking because most of the areas are already served by buses. Even an addition of new bus line in the neighborhoods does not make a big impression since its capacity and desirability are seen as less than those of light rails. The changes on the bus services are often seen as small, too insignificant to bring noticeable impacts on crime or property values.

Nevertheless, it is still possible to speculate how bus stops might influence crime opportunities. Since bus stops are located in open streets, waiting for a bus allows motivated offenders the opportunities to observe their surroundings. In addition, riding a bus affords opportunities to observe the outside along the routes. Buses frequently stop for various reasons: to pick up passengers, for traffic lights or due to traffic congestion. This means that bus passengers will develop the knowledge of areas along the bus routes and around the bus stops. This is also true for light rail or subway systems. While not as open or accessible as bus stops, traveling on light rail or subway vehicles may provide psychological comforts to offenders to overcome psychological barriers.

Furthermore, subways or light rail systems are free from traffic congestions, traveling faster and further than buses do.

The hypothesis that mass transit creates new crime hot spots can be tested by comparing crime data before and after the introduction of mass transit. Fortunately, there has been an increased interest and actual expansion of light rail services in recent years (Department for Transport, 2006), and these new light rail systems offer opportunities to examine their impacts on crimes. There are several studies examining the impacts of new light rails on crime. One of the studies was carried out by Poister (1996) who examined a total 11 types of crime data within 10 to 15 minute walking distance from the light rail stations. The data ranged 38 months prior to and 15 months after the operation of the light rail began. The data analysis suggested that the new rail stations may had increased crime temporarily.

A similar study was carried out by Sedelmaier (2003) on a new light rail system in Jersey City, NJ which opened in April 2000, adding a few more stations until November 2000. Examined crime types included both index crimes and quality-of-life offences. By comparing data which ranged from January 1996 to November 2001, Sedelmaier (2003) found that the new light rail system neither increased nor pulled crime toward the areas where the stations were located. Another similar research was carried out by Liggett and her colleagues (2003) in Los Angeles (LA), California. Liggett and her colleagues (2003) examined neighborhood crime levels and municipality-wide crime trends for five years before and after the inception of the LA Green Line. They also found that the new light rail service did not have any significant impact on crime trends. These research findings refute the concerns raised by opponents of expanding mass transit services to the suburban neighborhoods; crimes do not seem to be transported from the inner-city to the suburban areas via light rail as is frequently feared.

1.2. Offender Travel Patterns on Mass Transit

The distance between any two locations is an obstacle to be overcome for spatial interaction to occur. In general, the greater the spatial separation from the origin to the destination, the less attractive the travel becomes. The importance of distance in travel decisions is recognized by criminologists studying spatial patterns, and many have analyzed distances traveled by offenders in committing crime. Termed as distance decay, it was reported that there is a decreased level of both legitimate and criminal activities as the distances from offenders' residences is increased (Brantingham and Brantingham, 1993; Canter and Larkin, 1993; Eck and Weisburd, 1995; Wright and Decker, 1997; Wiles and Costello, 2000; Kent et al., 2005; Laukkanen and Santtila, 2005). This phenomenon has remained the same despite the increased ease in mobility in recent years (Wiles and Costello, 2000).

There is a lack of information on the offender travel patterns on mass transit. Most crime data do not differentiate between the locations where crime occurred and where it was reported (Newton, 2004). Other relevant information such as offenders' boarding locations, their residences, and their intended destinations are not routinely collected (Belanger, 1999; Newton, 2004). Combined with low clearance rates, these data limitations pose challenges in learning offender travel patterns using offender arrest data. However, the studies on offender travel patterns on light rail and subway show that offenders do not travel long distances on using mass transit (Belanger, 1999; Sedelmaier, 2003; Tilley et al., 2004)

By examining 252 arrested felons within the subway system between 1990 and 1995, Belanger (1999) found that the most prevalent type of trips (56 percent) was within offenders' borough of residences. In addition, these offenders tended to commit their offences in the same subway stations (23.9 percent) or the same subway lines (64.5 percent). This finding indicates that offenders' travel patterns are dictated by routine activities, and mass transit does not lengthen the distances traveled by offenders to unknown areas.

Sedelmaier (2003) also examined offender travel patterns using arrest data, four years before and one year after the light rail operation began and concluded that the new light rail system did not change offender travel patterns. A large number of offenders were found to reside in the same service zones or within one mile from their arrest locations. When offenders were arrested in different service zones from their residence service zones, they were often arrested in shopping center areas than any other areas. Furthermore, the arrestees who were arrested outside of their residential areas tended to come from other cities where the new light rail system did not serve. This negated the possibility that the offenders arrested outside of their own neighborhoods traveled on mass transit to the crime sites. Belanger (1997) reported a similar finding; when offenders did venture out of their own residential areas using the subway, they tended to travel toward city centers rather than suburban areas. Indeed, the most prevalent form of trips out of their own borough was to Manhattan (29 percent) rather than to any other areas.

Belanger (1997) offered two possible reasons for the observed travel patterns. First reason is that Manhattan may be attracting offenders due to its seemingly high number of targets in the areas. Another reason may lie on the subway system design: traveling to Manhattan is the most convenient trip compared to traveling to any other boroughs. Indeed, the least common types of trips (12 percent) which happened to be the longest trips were from one borough to other boroughs except to Manhattan. Overall, research studies show that even when offenders use public transportation system, they usually do not travel to suburban areas. In fact, offenders tend to travel toward the city centers due to the convenience influenced by system designs and abundance of targets.

2. Development of Crime Generators

It is consistently shown that crimes are disproportionately clustered in a few locations (Brantingham and Brantingham, 1993; Farrell et al., 1995; Eck, 1997; Shaw and Pease, 2000; Farrell and Pease, 2001; Farrell and Sousa, 2001; Eck et al., 2007). For this reason, the locations where offenders travel to commit crime are one of the focal interests for the opportunity criminologists. The destination attractiveness determines the volume of people traveling to any given area, and one way to measure the destination attractiveness is by number of possible activities in the areas. These highly attractive areas tend to be downtown, entertainment, shopping, and recreational districts which are easily accessible by public transportation. In addition, many retail businesses are likely to locate themselves along the major roads where it is convenient for the prospective customers to reach. This suggests that there would be increased crime along the major roads, busy areas, and mixed land use areas.

These highly attractive areas to offenders are divided into two groups; crime attractors and crime generators (Brantingham and Brantingham, 1995). The distinction between crime attractors and generators is whether the criminal incidents occur during normal legitimate activities or whether offenders travel to the areas with intent to carry out criminal transactions. According to Brantingham and Brantingham (1995), crime attractors are criminogenic places that draw motivated offenders into the areas due to their reputation for ease of committing crime. On the other hand, crime generators are places which offer many activities thereby drawing general population to the areas.

Consequently, the high volume of criminal incidents in crime generators is almost a secondary characteristic to their primary characteristic. For instance, shopping malls or downtown areas have abundant targets, and they can easily become crime hot spots. Indeed, research studies show that offenders often travel toward city centers and shopping centers when they commit offences outside of their own localities (Belanger, 1999; Sedelmaier, 2003).

To examine whether crime generators are formed, the area characteristics and types of activities influencing the volume of ambient population need to be collected and analyzed in relation to crime. To conclude whether crime attractors are formed or not, it may be necessary to obtain the intimate knowledge of the areas including area reputations or population characteristics to deduce why offenders display preference for the given areas. In addition, this type of process question may be better answered by qualitative or longitudinal studies.

2.1. Crime in and Around Mass transit

The real and perceived risks of victimization are an important consideration when making travel decisions. Some crime types are facilitated by over-crowding while other crimes are facilitated by isolation. Both of these crime types are the result of lack of resources: over-crowding occurs when there are not enough vehicles in ratio of patrons, and isolation occurs when staff supervision is low in ratio of patrons (Smith and Clarke, 2000). The crimes such as pick-pocketing, bag opening and low level sex crimes are facilitated by high target densities. For this reason, rush hours provide suitable situations for bag opening (Clarke et al., 1996) and low level sex crimes (Beller et al., 1980).

Other crimes are more conducive when the areas are relatively deserted which leads to lack of guardianship (Block and Davis, 1996). Decreased levels of guardianship are often used to explain increased risks of victimization when the levels of ambient population or target density are low. For instance, Clarke and his associates (1996) examined robberies on the NYC subway platforms. By standardizing platform robberies using ridership for 206 NYC subway stations, the authors found that the risks of robberies continued to increase as the densities of passenger decreased. Often called a late-night effect (Nelson, 1997), robbery rates on the subway platforms were found to be particularly high between 9 p.m. and 5 a.m. when the areas were relatively deserted. This late night effect is also observed for the bus stop crimes. While the majority of crimes against persons occurred during the day time, most serious crimes occurred between 10 p.m. and midnight when few people were expected at the bus stops (Loukaitou-Sideris, 1999). These research findings suggest that high numbers of pedestrians or bystanders who can act as capable guardians may be related to lower crime rates (Liggett et al., 2001).

The study on street robberies in relation to subway stations also showed the same result; the volume of street robbery was not the highest during the morning and evening rush hours when high levels of formal and informal surveillance were speculated. On the contrary, street robberies outside of mass transit stations were most common during the late night hours between 11 p.m. and midnight and around 2 a.m. where the streets were not booming with pedestrians (Block and Davis, 1996).

The lack of guardianship is also used to explain the peak in robbery incidents a short distance away from the transit stations (Block and Davis, 1996). The assumption is that the volume of people in the areas decreased as the distances from the stations increased. For instance, Block and Dvais (1996) examined street robberies in Chicago in relation to rapid transit stations and found that street robberies peaked a short distance (650 feet) away from the stations rather than the immediate vicinities of the stations.

After the secondary peak (1,200 feet), the existence of rapid transit station did not seem to influence occurrences of street robberies. Block and Block (2000) also examined robbery patterns in the Bronx in NYC and Chicago in relation to rapid transit stations and reported a similar finding; street robberies peaked a few hundred feet away from the rapid transit stations rather than the areas immediately surrounding the stations.

Transit related facilities may become hot spots of property crimes when the properties are left unattended for a period of time (Barclay et al., 1996). Subway station parking lots or hallways within mass transit systems are known to suffer higher levels of criminal incidents. Parking facilities can also provide places to commit other types of crime such as assault or robbery. As found by Loukaitou-Sideris and her colleagues (2002), in one of two high crime Green Line Stations, 60 percent of Type I crime occurred in the park-and-ride lots, while only 20 percent of them occurred on the platform.

It is said that mass transit riders experience higher levels of minor or quality-of-life offences than people on other public spaces (Loukaitou-Sideris et al., 2002). These quality-of-life offences are usually crimes against the mass transit system which include various offences such as fare evasion, vandalism, graffiti, littering, and various disorderly conducts. These crimes were traditionally thought to be victimless and thus not serious crimes. However, it was recently recognized that the quality-of-life offences have pervasive impacts on the fear levels and desirability of the services (Weidner, 1996; Nelson, 1997; Morgan and Cornish, 2006). Indeed, patrons of mass transit perceive quality-of-life offences on mass transit as precursors of more serious crimes and loss of control (Sloan-Howitt and Kelling, 1990; Nelson, 1997).

Research studies show that disorderly conducts or even innocent youthful activities performed in large numbers may incite fears among the patrons, conveying the message that the area is out of control (Nelson, 1997; Morgan and Cornish, 2006). These far-reaching impacts of quality-of-life offences are demonstrated by various studies. For instance, a household based survey of both riders and non-riders of buses showed that the three primary concerns related to personal security in using buses were obscene language, panhandling, and disorderly conducts, rather than any grave bodily harms (Lusk, 2001). The survey result in Britain also showed that one of the most common victimization in transit environments were being stared at in a threatening manner or being deliberately pushed (Crime Concern, 2004).

Existing research studies also show that mass transit crews face higher levels of victimizations than employees in other occupation (Department for Transport, 2002). Reflecting actual risks of bodily harms, almost a half of mass transit workers reported having concerns of being assaulted at their work (Budd, 2001). It is shown that the transit crews who have direct contacts with patrons such as bus staff suffer more victimization (Department for Transport, 2002). One of the main causes of assault on bus staff was due to disputes over fares. The fare system where the fares need to be collected directly from passengers usually produces higher numbers of criminal incidents against the staff and crew members. Recently, a stabbing death of a NYC bus driver by a fare evader grappled the attention of nation.[1] Another growing concern with the bus operators' safety is disorderly behaviors on the transit vehicles (Department for Transport, 2002). About a third of assailants on bus staff were aged between 13 to 16

---

[1] On December 1st of 2008, a bus driver driving a Brooklyn line in NYC was stabbed to death by a fare evader. On the said day, a passenger got on the bus without paying his fare which the driver did not intervene following a policy guideline by the bus company. Later on, the same passenger requested a free transfer which was denied. Enraged, the fare evader punched and stabbed the driver, who was pronounced dead shortly after. New York Times article written by McFadden, Robert D. (December, 2nd, 2008 available online: http://www.nytimes.com/2008/12/02/nyregion/02driver.html)

years old, validating concerns for increasing hooliganism in general on the transit vehicles.

It is difficult to assess the impacts of mass transit on crime. When crimes occur in close proximity to transit station or stops, it is not clear what proportion of offenders used mass transit to travel to the crime location. While offenders tend to select targets located in areas easily accessible by public transportation, this does not mean that offenders *used* public transportation to travel to their crime sites (Brantingham and Brantingham, 1993; 1997; Wright and Decker, 1997; Desroches, 2002). In addition, research studies on offender travel patterns show that offenders display directional biases when they travel; they tend to travel toward areas with many targets such as city centers or shopping centers (Brantingham and Brantingham, 1993; Eck and Weisburd, 1995; Clarke, 1999). Since the busy areas are most often served by mass transit, crimes often seem to cluster in areas with mass transit or along major roads. In fact, high crime bus stops are frequently concentrated along the main streets (Liggett et al., 2001) or city centers (Loukaitou-Sideris, 1999). Newton (2004) also reported a similar finding. The bus shelters with highest levels of criminal damage tended to be located in main roads in relation to the city center.

2.2. Does Mass Transit Crime Reflect Area Crime Rates?

There are disagreements regarding whether mass transit crime rates simply reflect the above ground crime rates or not. For bus stops located without access control to the areas, it is likely that the bus stop crime rates reflect crime rates of their surroundings. Indeed, the prior research on bus stops and crime found that the area crime rates were related to crimes at bus stop (Liggett et al., 2001) or light rail stations (Loukaitou-Sideris et al., 2002). For instance, Liggett and her colleagues (2001) used regression model to examine impacts of environmental characteristics in relation to bus stops. Using unit of analysis as 150-foot radius of an intersection with bus stops, both transit police data and city police crime data were aggregated to the intersection level. Data on three groups of environmental characteristics were collected (2001: p21):

> “(a) urban form characteristics around intersection, which included information on the land use and condition of the surrounding area; b) bus stop characteristics such as the existence of bus shelters, visibility, and lighting; and c) street characteristics such as street and sidewalk, on-street parking and traffic levels”

Liggett and her colleagues (2001) reported that the most important predictor of crime concentration was location, meaning that high crime bus stops were located in dangerous areas to begin with. These dangerous areas were characterized as being disadvantaged in social and compositional characteristics, and often closely located from undesirable facilities inflicted with litters (Liggett et al., 2001). Newton (2004) reported a similar finding. Vandalism on bus shelters showed positive correlation with volumes of recorded crime in the areas.

A similar finding was reported in relation to light rail stations. For instance, the crime rates of the Green Line in LA, California was examined in relation to the social and physical characteristics of their neighborhoods (Loukaitou-Sideris et al., 2002). The LA Green Line has a total of 14 stations in residential, industrial, mixed land use areas as well as inner-city areas with graffiti and litters. The data analysis showed that the low crime stations were located in affluent suburban communities with low crime rates. Except for motor vehicle related theft, the station crime rates were directly related to the crime rates in the station neighborhoods. The authors explained the high volume of motor vehicle related offences in the station parking lots by pointing out that there were not many parking lots existed other than around the mass transit stations. It is also noteworthy to point out that the LA Green Line is a light rail system employing an honor system without any barriers to the stations.

However, it can be hypothesized that that well designed mass transit stations with access control can deter criminals from the above-ground. LaVinge (1996) hypothesized that there would be enough differences in crime rates between above and below ground if the mass transit system was able to fend off the criminals from invading to the underground transit environments. Comparisons of Washington D.C. Metro systems and crime rates of above-ground showed that assault was the only crime type displaying significant positive correlations with the above-ground crime levels. Clarke and his colleagues (1996) reported a similar finding. They found that the NYC Subway station robbery rates were not correlated with the above-ground robbery rates at the precinct level. However, they also found substantial variations in robbery rates among the stations within each precinct (Clarke et al., 1996). Since NYC subway systems were not designed uniformly as the Washington D.C. Metro was, this finding is probably not surprising.

In some cases, the crime rates of surrounding areas are shown to influence crime concentrations around mass transit areas. In a separate study, Block and Davis (1996) examined spatial patterns of street robbery in four Chicago Police Districts. In two districts with low robbery rates, street robberies were concentrated near rapid transit stations while in the two high crime districts the concentrations were less pronounced. In these high crime areas, robberies were most likely to occur along main streets; at least every block had some robberies during the 1993 and 1994. In a separate study, Block and Block (2000) examined street robbery in the surrounding areas of elevated rapid transit stations in Chicago and in the Bronx, NYC, and found that street robberies were concentrated around the stations. The authors hypothesized that the existence of both legal and illegal activities around the transit stations explained clustering of street robberies around the stations.

3. What Factors Increase Crime Opportunities?

3.1. Environmental Characteristics Related to Crime

Research studies show that there seems to be relationships between certain types of commercial establishments and crime. Many commercial establishments such as convenience stores (Hunter, 1999), banks, sports facilities, parking structures (Loukaitou-Sideris et al., 2002), restaurants, liquor outlets, and other types of retail businesses often end up becoming crime targets themselves or attract offenders to the areas (Taylor and Mayhew, 2000 March; Eck et al., 2007).

Among the commercial establishments, liquor-licensed establishment is one of the facility types which received widespread attention from criminologists in their impacts on crime. It is often thought that liquor-related places or high levels of alcohol consumption are linked to criminal behaviors (Block and Block, 1995). Block and Block (1995) examined spatial patterns of street robberies in relation to taverns and liquor stores in Chicago. Due to the high frequency and being located on major streets, Block and Block (1995) proposed that neither density nor existence of liquor-related places should be used as an indication of high criminal activities. The study found that high-crime levels at liquor-related places generally reflected the crime pattern where they were located. Interestingly, high-crime liquor-related establishments within hot spot areas were far more likely to be near an elevated rapid transit station.

The complexity in identifying causal relationships or even correlates of crime is often exemplified by conflicting empirical findings. For instance, Newton (2004) reported that the number of liquor-licensed establishments in the area was associated with lower incidents of bus shelter damage. On the other hand, other researchers found that the liquor-licensed establishments were associated with higher numbers of crimes on bus stops (Loukaitou-Sideris, 1999; Loukaitou-Sideris et al., 2001).

In general, there seems to be three types of environmental factors increasing crimes in close proximity to bus stops. One of the factors is types of commercial establishments in the areas; facilities such as liquor store and check cashing stores were related to high crime rates (Liggett et al., 2001; Loukaitou-Sideris et al., 2002). In addition, bus stops in close proximity to parking lots with no attendant displayed higher number of criminal incidents. For the crimes in close proximity to the bus stops, the fenced parking lots had positive correlation with crime rates while unfenced parking lots had negative correlation (Liggett et al., 2001). To summarize, liquor outlets are usually associated with higher crime rates in the areas although this seems to be reflections of crime rates of their surroundings (Block and Block, 1995; Liggett et al., 2001; Loukaitou-Sideris et al., 2002).

The second factor is related to the layout of the street and land use patterns such as alleys, midblock passages, vacant building, and multi-family housing (Loukaitou-Sideris et al., 2001). Yet in another study, unit vacancy rates did not show association with crimes at light rail stations (Loukaitou-Sideris et al., 2002). Third factor that seems to be associated with increased crime is physical incivilities such as graffiti and littering supporting Broken Windows approach to understanding crime (Loukaitou-Sideris, 1999; Loukaitou-Sideris et al., 2002; Newton, 2004).

3.2. Area and Resident Characteristics Related to Crime

There are some supports that both resident and ambient population influence crime rates in relation to mass transit. Belanger (1999) found that subway stations where crimes occurred usually had significantly higher ridership than on average. It was the opposite for the offenders' origin subway stations. Newton (2004) also reported a stronger influence of ridership on bus shelter damage than the size of resident population. Yet in another study, light rail stations in residential areas had higher crime rates while stations located in primarily office and industrial areas with seemingly high ambient population had lower crime rates (Loukaitou-Sideris et al., 2002).

In addition, area deprivation is significantly related with the variations in crimes at light rail stations (Loukaitou-Sideris et al., 2002), bus stops (Loukaitou-Sideris, 1999), and bus shelter damage (Newton, 2004). Some of the resident characteristics shown to influence crime levels are often social disorganization variables such as poverty, family stability, residential mobility, ethnicity, immigration status, percent of renters in the areas, youths, and unemployment rates (Wikstrom, 1995; Wang and Minor, 2002; Andersen, 2006). However, when the Washington D.C.'s Metro's Green Line was expanded to serve inner-city areas by adding six stations, this did not lead to increased crime rates on the Metro system (LaVigne, 1996). This lends support to the claim that a well-designed system is able to fend off criminal invasions from the above-ground.

Angel (1968) also found that the middle income areas or the business areas catering middle class had little to no robberies. This is perhaps not surprising since the wealthy individuals or businesses alike have easier time in adopting security measures against crimes than the disadvantaged. Furthermore, research shows that offenders are not familiar with middle-class areas, and the areas known to offenders overlap with each other due to their personal or past ties (Wiles and Costello, 2000).

The characteristics of areas, such as being residential or industrial, determine the volume and characteristics of people in the areas. Some commercial and residential areas where mass transits stations are located suffer from high crime rates while others do not. Therefore, examining the surrounding environments such as what kind of businesses or activities are available in the vicinities would help confirm whether the mass transit stations indeed create crime opportunities in the areas.

3.3. Lack of Capable Guardians

The mass transit crews' ability to monitor the transit vehicles seems to increase the passenger's sense of security as well as providing some protection. When crimes are examined in relation to buses, 67 percent of them occurred at bus stops whereas only 33 percent occurred on buses (Loukaitou-Sideris, 1999). In addition, the reported fear levels are lower while travelling in the transit vehicles in comparison to waiting at the bus stops (Crime Concern, 2004).

The late-night effect, increased risks of robbery both inside and outside of transit facilities when there are fewer number of targets, is explained by decreased levels of guardianship due to isolation of targets. Newton (2004) also reported a positive relationship between bus shelter damage and open areas without place managers such as parks and children's play areas. In addition, the presence of school with high truancy rates within 100 meters (328 feet) from the bus shelter displayed a positive correlation with bus shelter damage. This led Newton (2004) to hypothesize that the unsupervised youths may be responsible for bus shelter damages. Furthermore, contrary to the common belief, there were negative relationships between bus shelter damage and the presence of alcohol-related establishments. An explanation offered by Newton (2004) was that while alcohol-related facilities were managed and controlled by employees, open areas such as parks and playground lacked guardianships to prevent undesirable activities. It was also found that high crime bus stops were often located in isolated areas with lack of natural and formal surveillance (Loukaitou-Sideris, 1999). This also applies to light rail stations where high crime stations tended to be located in areas with poor visibility and lack of natural surveillance (Loukaitou-Sideris et al., 2002).

CHAPTER 3. THEORETICAL FRAMEWORK AND RESEARCH QUESTIONS

1. Theoretical Background

The question of why crimes occur where they do has attracted the interests of criminologists with ecological orientation for decades. Beginning with Shaw and McKay (1942)'s seminal work which analyzed the juvenile offender residence locations, the opportunity theories in the early 1980s regenerated interests in ecological perspective in crime analysis (Paulsen and Robinson, 2004). These opportunity theories are differentiated from the earlier ecological perspective due to their focus on smaller areal units such as places rather than neighborhoods (Anselin et al., 2000). In addition, these theories shift their focus to criminal offences from offenders (Felson and Clarke, 1998).

Opportunity theories recognize the importance of crime opportunities to explain occurrences of crimes (Felson and Clarke, 1998). In addition they also recognize that crime opportunities are not equally distributed: they are spatially and temporally clustered (Felson and Clarke, 1998). This is because crime opportunities are shaped by routine activities at spatial and temporal levels.

In the present research, three opportunity theories are employed to explain spatial patterns of crimes: routine activity theory, crime pattern theory, and rational choice theory in conjunction with situational crime prevention techniques. The frameworks of routine activity and crime pattern theories are used to explain formation and exposure levels of crime opportunities to the potential offender populations whereas rational choice theory is used to explain differing target vulnerability at micro-level.

1.1. Routine Activity Theory

There were sharp increases in crime rates during the 1960s and 1970s in both violence and property. At the same time, the nation also witnessed increased socio-economic indicators (Cohen and Felson, 1979). The increased crime rates accompanied with improved quality-of-life contradicted a dominant explanation of criminality at that time; that poverty was the main cause for people committing crimes. By analyzing the changes occurred in routine activities at macro levels, Cohen and Felson (1979) reconciled this seemingly contradictory phenomena: increased crime rates accompanied with increased socio-economic indicators.

In routine activity perspective (Cohen and Felson, 1979), crime opportunities are created when motivated offenders intersect with suitable targets in the absence of capable guardians. The formation of crime opportunities, Cohen and Felson (1979) suggest, is influenced by people’s routine activities. Cohen and Felson (1979) supported their assertion by examining the macro level changes in routine activities and crime rates. They claimed that changes in routine activities at macro level resulted in increased crime opportunities by leaving targets without proper protections.

Routine activity theory is also relevant to temporal variations of crime since routines dictate *where and when* people carry out their daily activities. For instance, the timeframes of when stores are occupied, when the houses are left empty, when the youths are unsupervised influence when crime opportunities are mostly likely to be created. In recent years, the concept of capable guardian is developed into three types of controllers: intimate handlers, managers, and guardians (Eck and Weisburd, 1995; Felson, 1995). Intimate handlers are those who are capable of controlling offenders; they may be parents, teachers, or even employers. Place managers are responsible for particular facilities; they have control over how the property is managed and protected. Lastly, the guardians are those who can protect potential targets and victims from victimizations. Naturally crimes are most likely when all of these three controllers are absent when the offenders and potential victims intersect.

1.2. Crime Pattern Theory

Crime pattern theory utilizes the concepts of paths, nodes, and edges to bring spatial aspects to crime analysis (Brantingham and Brantingham, 1984). In this perspective, everybody has awareness space defined as the areas known to him or her usually through routine activities. The awareness space consists of three types: nodes, paths, and edges. The nodes are often the focal activity points such as residences or workplaces. The paths are the streets used to travel from places to places. The edges are the areas near boundaries where mixed land use begin. This makes it difficult to distinguish outsiders from insiders, making it easier for offenders to blend in.

Offenders develop their awareness space through their daily movements like any other people (Brantingham and Brantingham, 1993; Eck, 1997; Wright and Decker, 1997). According to Brantingham and Brantingham (1993), offenders' activities, whether legal or illegal, often occur within their awareness space. This is because the offenders have substantial knowledge of the areas in addition to having more psychological comforts. The higher activity levels in the awareness space will lead to increased risks of criminal incidents in the neighborhoods. In this perspective, it is expected that the areas used by many people will experience a high volume of crime. This explains why central business district, recreational, and entertainment areas frequently become crime hot spots; these areas are exposed to a large number of people through routine activities and may end up becoming offenders' activity spaces (Lu, 2003).

1.3. Rational Choice Theory

If the targets in one's awareness spaces are equally suitable for commission of crime, then crime should be somewhat evenly distributed in a given area. However, research studies report observations of hot dots or repeat victims even in crime hot spots. This indicates that suitability of targets even in relatively small areas is highly unequal.

While both routine activity and crime pattern theories provide useful premises regarding crime patterns at macro and meso-levels, rational choice theory (Cornish and Clarke, 1986) in conjunction with situational crime prevention techniques (Cornish and Clarke, 2003) help explain the phenomenon of repeat victims or hot dots. From a rational choice perspective, offenders' decision to commit specific criminal act are often made in a relatively short time period with limited information at that time. In this decision making process, offenders weigh the costs and benefits of committing crimes. As implied by the term *situational factors,* target suitability is situation specific and highly fluid at the micro level.

To understand why some victims are repeatedly victimized even in high crime areas, situational crime prevention techniques provide which targets are more likely to be victimized than others are. The targets requiring more efforts and risks while offering lower provocation, rewards, and excuses will be less often targeted. Some of these situational factors may be embedded in the physical structures or environments while some may be influenced by the management styles or maintenance of the facilities (Cornish and Clarke, 2003).

For instance, it is well known that big corporations and banks are located in areas with easy accessibility. However, they may employ various security measures such as private security guards, controlled access, and CCTVs. Due to the heightened security measures of guardianship, these targets may not be attractive to (unskilled) potential offenders (Loukaitou-Sideris *et al.*, 2002). In addition, target suitability can be influenced by a host of other characteristics such as the store layouts, product displays, business hours (Eck et al., 2007), number of employees and customers, and cash handling practices (Hunter, 1999). Due to the highly specific nature of the situational factors influencing target suitability, it may be difficult to measure them without carrying out security audit on crime specific targets.

The above mentioned three opportunity theories suggest that crimes will cluster in areas used by many people with a high number of vulnerable targets. Within their activity space, offenders will select suitable targets influenced by various situational factors. To summarize, routine activities influence formation of crime opportunities while crime pattern theory explains the areas' exposure levels of crime opportunities to the offenders in the areas. Lastly, rational choice theory explains the process of target selection and the existence of hot dots.

2. Spatial Concentrations of Crime

Crimes, whether they occur between strangers, intimates, or against properties, are known to cluster in space over an extended time period. Crimes displaying spatial concentrations include assault (Brantingham and Brantingham, 1998; Lockwood, 2007), robbery (Brantingham and Brantingham, 1993), burglary (Brantingham and Brantingham, 1998; Ratcliffe, 2003), motor vehicle theft (Clarke and Goldstein, 2003; Mayhew and Braun, 2004; Plouffe and Sampson, 2004), and even spousal assaults (Miles-Doan, 1998; Cunradi et al., 2000; McQuestion, 2000; Benson et al., 2003; Yu, 2004). Crimes in context of mass transit is not an exception in displaying spatial concentrations (Loukaitou-Sideris et al., 2002; Liggett et al., 2003; Newton, 2004). For instance, three percent of bus shelters made up 20 percent of bus shelter damage, and 12.5 percent for 50 percent (Newton, 2004). For bus routes, only two percent of bus routes made up 60 percent of all bus incidents on the routes (Ibid).

Empirical studies show that crime concentrate in smaller scale such as places or facilities rather than at neighborhood levels (Tseloni and Pease, 2003). From the previous research findings, it is often concluded that crime hot spots are composed of hot dots or repeat victims (Tseloni and Pease, 2003). Consequently, an area where crime is evenly spread out seems to escape empirical observations. Furthermore, even in the same business category, a small number of facilities house a disproportionate number of criminal incidents. The places which generates disproportionate number of crimes are termed as risky facilities (Eck, 1997; Pease, 1998; Burrows et al., 1999; Eck et al., 2007).

Despite the enduring perception that links mass transit to crime, there is very little empirical support for causal relationship between mass transit and crime. There are several reasons that it is not suitable to draw a causal relationship between mass transit and crime. First, new transit systems do not seem to extend the distances traveled by offenders (Belanger, 1999; Sedelmaier, 2003) or create new crime hot spots in suburban areas (Poister, 1996). Research studies on offender travel patterns show that offenders do not venture out far, and when they do, they tend to travel toward city centers rather than outward areas. Second, research studies on the impacts of new light rails on crime have found very little to no changes in the existing crime patterns (Liggett et al., 2003; Sedelmaier, 2003).

Third, there are mixed research findings regarding whether crime rates of mass transit stations correlate with crime rates of their surrounding areas. While some researchers report finding that mass transit crime rates do not correspond to the surrounding areas' crime rates (Clarke *et al.*, 1996; LaVigne, 1996), others report that crime rates between the surrounding areas and the mass transit stops reflect each other (Loukaitou-Sideris *et al.*, 2002; Newton, 2004). However, it is important to note that the former group of researchers studied the subway stations with access control while the latter group of researchers examined bus stops or light rail stations without access control.

Lastly, not all mass transit stops and stations suffer from high crime concentrations (Block and Davis, 1996; Loukaitou-Sideris, 1999; Liggett *et al.*, 2001; Newton, 2004). The fact that mass transit does not uniformly increase crime in the areas may signal that the crimes in the areas due to the interaction between mass transit and their bigger environments.

3. Research Questions

The main goals of the present research study are to examine the influences of bus stops while taking into account their surroundings to examine causes of crime concentrations. In this research study, the covariates of crime are divided into two groups; bus stops and commercial establishments.

Many activities, including criminal activities, are intertwined with legitimate activities. Bus stops are often located on the open street without clear demarcation or access control. If they are located in high crime areas, it is difficult to separate the impacts of bus stops on crime from other factors. Therefore, whether bus stops increase crime rates in the surrounding areas need to be examined in the context of larger areas; its bigger environments and with other available legitimate activities (Felson et al., 1996). This requires relating places to one another in order to understand how crime opportunities are linked to legitimate routine activities. This is particularly relevant because many bus stops are located along major roads or commercial areas where crime concentrations are frequently observed (Brantingham and Brantingham, 1997; Block and Block, 2000).

Research studies show that some types of businesses may provide natural surveillance leading to lowered crime while others may have the opposite impacts by attracting offenders to the areas (Felson et al., 1996). For instance, it is not likely that crime will cluster around banks due to various security measures employed, yet it seems plausible that crime may cluster around neighborhood convenience stores as it is often depicted in the movies. In addition, special purpose facilities influence levels and characteristics of ambient population, and different ambient population influences crime differently. For instance, areas around schools are often thought to have higher criminal incidents while this is not the case for beauty shops. Since cities are divided by different land use, spatial patterns of crimes against fixed targets will be heavily influenced by target locations. Also commercial establishments tend to locate themselves closely to the mass transit stations and they may attract targets to the areas rather than bus stops themselves (Block and Block, 2000).

To examine whether bus stops help develop crime generators or crime attractors, a few things need to be examined. First, whether crimes cluster around bus stops need to be examined. Second, if bus stops are shown to increase crime in the areas, then the area characteristics relevant to crime generators need to be surveyed. If there are many services and activities offered in the areas attracting a large number of people, then it is likely that crime generators are developed. However, if there are crime concentrations around bus stops despite lack of legitimate activities available in the areas, then it is likely that crime attractors are developed. Finding out whether crime attractors are formed would require intimate knowledge of the areas to relate which factors are attracting offenders. This step may also require environmental survey and observation of the high crime areas which are beyond the scope of the present research study.

3.1. Research Question One and Hypotheses

The focus of the present research study needs to be placed on crime types where the two types of independent variables (bus stops and commercial establishments) are thought to influence crime opportunity levels. To answer the research questions, five crime types including robbery, aggravated assault, motor vehicle theft, and theft from motor vehicle will be analyzed. The 2007 crime data were obtained from Uniform Crime Report (UCR) data compiled by Newark City Police Department. To achieve the goals of the proposed research, two research questions will be investigated. Table 2 summarizes the research questions. The first research question is concerned with whether the existence of bus stops increases the occurrences of robbery, aggravated assault, motor vehicle related offences, and burglary in the areas.

In general, existence of bus stops may increase crime opportunities if they provide targets for offending. If offenders can expect concentration of targets in the vicinity of bus stops, then this will likely attract offenders to the areas. Empirical research studies show that robberies often cluster in and around mass transit stations and along the major routes. Therefore, it is hypothesized that the presence of bus stops will increase risks of robbery. It is also shown that the vehicles parked in parking facilities have higher risks of victimization than vehicles parked in residential areas (Mayhew and Braun, 2004; Plouffe and Sampson, 2004), reflecting offenders' desire not wanting to waste their time by traveling to low target density areas. The roads where bus stops are located tend to be business areas with commercial parking facilities or street parking. Since there are usually abundant parking spaces where bus stops are located, motor vehicle related offences will cluster around bus stops.

Table 2. Research questions and hypotheses

Research question	Hypotheses
1. Do areas with bus stops experience higher numbers of criminal incidents?	• Bus stops will increase robbery, motor vehicle theft, and theft from motor vehicle in the areas. • It is not clear whether bus stops will increase crime opportunities for burglary or assault.
2. Do areas with commercial establishments experience higher numbers of criminal incidents? In particular what type of businesses seems to influence what kinds of crime?	• Commercial establishments will increase robbery, assault, motor vehicle theft, and theft from motor vehicle in the areas. • It is not clear whether bus stops will increase crime opportunities for burglary.

It is not clear whether burglars in urban areas use mass transit or private vehicles to carry out their offences. However, in crime pattern theory, offenders are more likely to notice suitable targets within their awareness space composed of nodes, paths, and edges. This means that the residential units around bus stops or bus routes will be placed at higher risks than other residences. In reality, the bus stops are often linked to and located in business land use areas rather than the residential areas. Therefore, it is not clear whether bus stops will increase the risks of burglaries. The same can be said about the aggravated assault. There seems to be no theoretical or empirical basis to hypothesize that the existence of bus stop increases the risks of being assaulted.

To summarize, it is hypothesized that crimes of robbery, motor vehicle theft, and theft from motor vehicle will display positive correlations with the existence of bus stops. Hypotheses are not formed for burglary and aggravated assault due to lack of theoretical and empirical basis.

3.2. Research Question Two and Hypotheses

The second research question is related to whether commercial establishments increase crime in the areas. A particular interest would be what kind of businesses seems to increase crime in the areas. Often called street crime tend to show concentrations in city center areas (Wikstrom, 1995). It is found that high crime bus stops are often concentrated along the main streets (Liggett et al., 2001) or city centers (Loukaitou-Sideris, 1999). Newton (2004) also reported a similar finding. The bus shelters with highest levels of vandalism tended to be located in main roads in relation to the city center. If businesses are located heavily around mass transit, then concentrations of robberies in these areas might not be surprising.

The motor vehicle related offences may be a good indication of areas where the levels of ambient population are high. It is likely that it is the visitors who park their vehicles in the parking lots or around bus routes, rather than the resident populations. It is also reasonable to expect that crimes against businesses cluster around major roads since businesses tend to locate themselves nearby major roads. Indeed, the majority of auto related theft are found to occur in the mass transit parking lots (Loukaitou-Sideris et al., 2002) or near major activity centers or along major transit routes where ample parking spaces exist (Barclay et al., 1996).

Existing research studies show that robbery and motor vehicle related offences often show concentrations in the city centers, shopping centers, or central business districts. So it is expected that existence of commercial establishments will increase incidents of robbery, motor vehicle theft, and theft from motor vehicle.

A recent study shows that offenders arrested with violent offences are more likely to be under the influence of alcohol (Valdez et al., 2007). This means that areas where alcoholic beverages are served may have higher numbers of assaults but not necessarily influenced by bus stops in the areas. This provides a hypothesis that the entertainment areas, particularly if alcoholic beverages are served, will have higher numbers of assault. However, for crimes against fixed targets requiring no contact such as residential burglary, it is not clear whether and how ambient population will have an impact on their occurrences.

To summarize, robbery, assault, motor vehicle theft, and theft from motor vehicle are expected to display positive correlations with number of commercial establishments. Due to the land use patterns, the commercial establishments will not increase burglaries: in fact, negative correlation is expected since the areas where businesses are located would not have residences in the areas.

CHAPTER 4. RESEARCH DESIGN

The present research study will be carried out by performing a series of hypothesis tests. The first proposed step is to examine whether crime cluster around bus stops. When comparing crime based on bus stop, it is necessary to expect homogeneity in the influences of bus stop within the areal units. The same is true for other independent variables; it is necessary to assume homogeneity in the influences of commercial establishments within the areal units. Small areal units is one of the minimum requirements to assume area homogeneity, since things that are closer to each other are more likely to be similar than dissimilar.

This research employs an ex post facto design which independent variables (i.e., number of bus stops) are not manipulated. Since the crime data include information on locations, the data analysis should and will include theories and techniques incorporating spatial information in the explanations or interpretations. Spatial econometrics, a subfield of regional science, allows the researcher to address the problems and challenges caused by spatial effects in data analysis (Anselin, 1988). There are two special qualities of spatial data which requires different statistical techniques from the traditional statistics. First is the spatial dependence, and the second is the spatial heterogeneity

This section begins with an overview of the study area. The other topics discussed include the data source and the unit of analysis. This section concludes by addressing the spatial aspects of the data analysis in relation to the present study.

1. Study Area: Newark, New Jersey

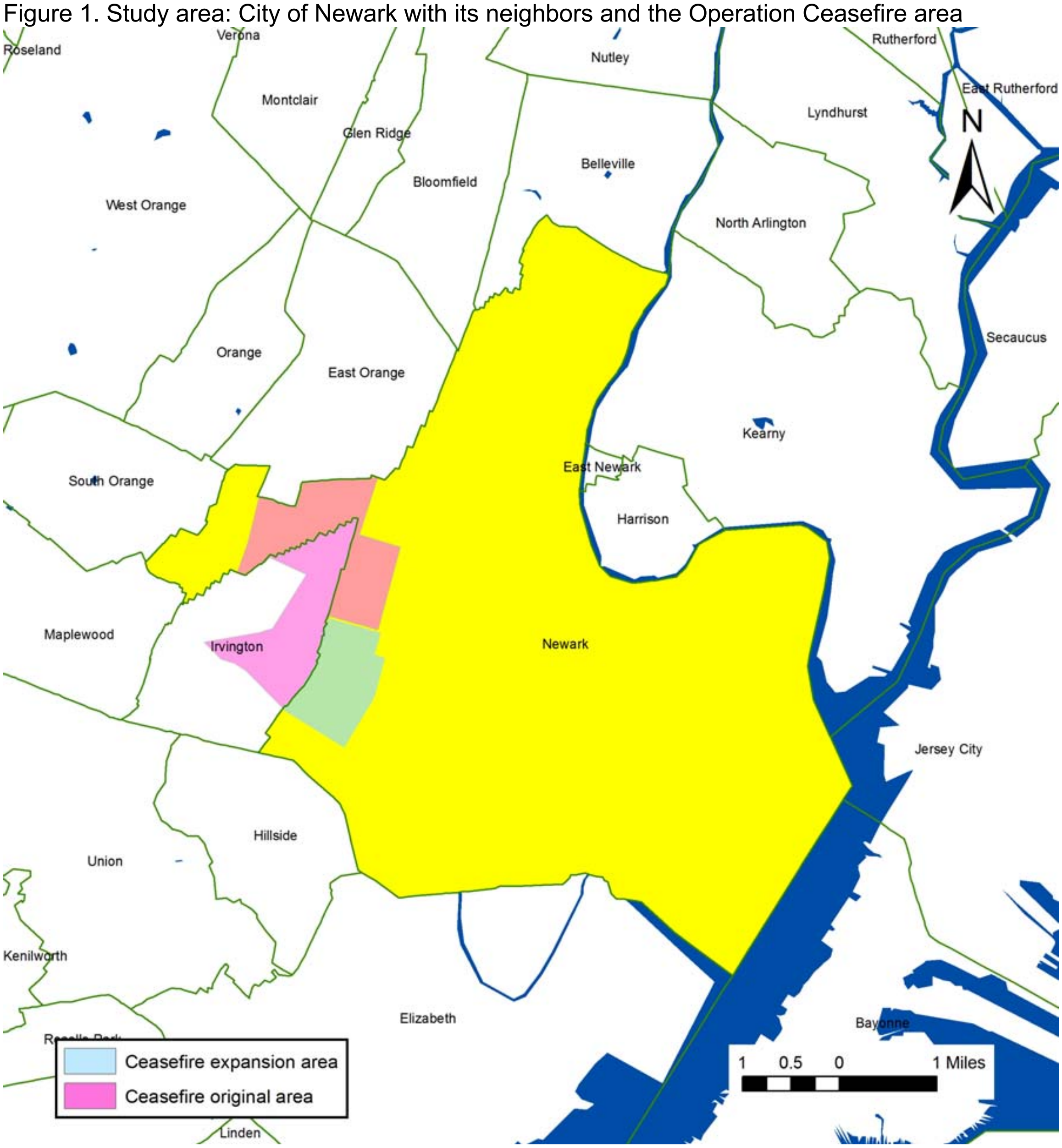
The study area of the present study is Newark, NJ. Newark shares its boundaries with eight other municipalities. Figure 1 shows the map of Newark with its neighbors. The perimeter of Newark is about 32 miles. The east side of Newark is surrounded by Passaic River and Newark Bay. Due to these geographic barriers, Newark does not have geographically connected neighboring municipalities in its east side. Also these geographic barriers block most of the spatial influences from the east side barring several bridges and highways. The size of Newark is 26 square miles, and the U.S. Census Bureau estimated its population to be 265,357 in 2007. Newark is also known for its ethnic diversity with more than hundreds bus lines traveling within or to and from Newark.

In addition to bus services, there is one PATH station (Newark Penn station), three rapid rail stations (Newark Penn, Broad Street, and Newark Liberty International Airport stations), and 15 light rail stations composed of two lines in the city. The rapid rail stations and light rail stations are excluded from the data analyses given their relative infrequency in the city and the study's focus at the micro-level. Examining the impacts of one or two stations on crime will not likely produce meaningful or generalizable results. Furthermore, bus stops are always located right outside of the rapid transit stations or subway stations. Therefore, it can be argued that bus stops could approximate the level of mass transit in the city.

Newark is the largest city in NJ and boasts the largest police forces in the State. In 2006 Newark City Police Department reported having 1,286 sworn officers: 1,065 male and 221 female officers (State-of-New-Jersey, 2007). The city also houses five higher educational institutions and several financial institutions. This indicates that

Newark has high levels of commuters in the area. The factors such as high volume and variations in bus stops and businesses provide an ideal setting for the present research study.

Figure 1. Study area: City of Newark with its neighbors and the Operation Ceasefire area



The socio-demographic characteristics of Newark and its neighboring municipalities are summarized in Table 3. In the nation, Newark is a disadvantaged city.

While the national figure for percent of family below poverty level was 9.8 percent for 2007, Newark's figure was more than double of that proportion. In 2007, the Census Bureau estimated that 20.5 percent of families in Newark fell below the poverty level. Also the median family income was lower than the national average; in 2007, the median family income in the nation was estimated to be $60,374 whereas it was only $40,583 for families in Newark. In addition, Newark has a reputation as being one of the most dangerous cities in the nation based on official crime figures.

Table 3. Socio-demographic characteristics of Newark's neighboring municipalities

Name of municipality	Shared boundary in mile	Location	Median family income ($)	Percent family below poverty level	Population
US	n/a	n/a	60,374	9.8	298,757,310
Newark	n/a	n/a	40,583	20.5	265,375
Belleville	3.1	N	60,348	6.3	35,712
Bloomfield	.5	N	83,321	4.7	47,066
East Orange	4.0	NW	42,027	23.2	62,240
South Orange	1.5	W	107,641	1.9	16,964
Maplewood	.7	W	111,725	1.7	24,588
Irvington	5.5	W	51,433	12.4	59,532
Hillside	1.8	SW	74,178	3.7	19,932
Elizabeth	2.8	S	46,026	15.4	126,538

Source: U.S. Census Bureau, 2005-2007 American Community. The figures of South Orange Township are from the 2000 census data.

The characteristics of Newark's neighboring municipalities are quite diverse. While South Orange and Maplewood share their boundaries with Newark, their median family incomes were a lot higher and percent of families below poverty line were quite lower than the figures of Newark. However, the three municipalities on the west and south sides of Newark (East Orange, Irvington, and Elizabeth) are just as disadvantaged as Newark.

In particular, there seems to be quite a spatial interaction on violence crimes between Irvington and Newark. This is reflected in Operation Ceasefire which began in May of 2005 and concluded at the end of 2008. The expansion of Operation Ceasefire occurred in year 2006 since it was perceived that violence in this area transcended the administrative boundaries.

2. Data Sources

There are four main data sources used in the present research study; the 2007 Newark City Police crime data; the 2007 Verizon's YellowBook NJ database; the 2007 NJ Transit Corporation mass transit data; and the 2006-2007 Newark land use and zoning district information. In this section, the data will be examined in preparation for data analyses.

2.1. Newark City Police Crime Data

The Newark City Police are responsible for keeping records of criminal incidents occurring within the city. There are two other police forces in Newark; the NJ Transit Police who are responsible for criminal incidents occurring inside the mass transit stations or on its vehicles, and the Port Authority Trans-Hudson Corporation (PATH) Police who are responsible for the criminal incidents at the Newark Liberty International Airport as well as at the Pennsylvania PATH station.

The crime data from NJ Transit Corporation and PATH are not available for the present research study. While the incorporation of these data are desirable, previous research studies show that citizens may simply call local police even if they are victimized within mass transit environments (Liggett et al., 2001; Crime Concern, 2004). This is more likely when demarcation is not clear such as buses as opposed to underground stations. While it is safe to assume that the Newark City police will not be called to the airport for criminal incident due to the distance, it is probable that passengers of mass transit, whether they use buses, light rail or commuter rail, may simply call local police for help.

For the purpose of the present research study, the Uniform Crime Report (UCR) classification data for year 2007 were obtained from the Newark City Police Department. The crime data contained information on the date, time, weapons used, location, and type of crime premises. Table 4 summarizes the selected crime types. The original dataset separated pick-pockets and snatch incidents (n= 95) from robbery. To follow the UCR classification, these incidents were combined with robbery. In addition, robbery incidents also included 127 car-jacking incidents. For aggravated assault, there was no information on the relationships between victims and offenders. Therefore, it is impossible to find out the proportion of intimate partner violence in aggravated assault. The weapons used for this offence ranged from firearms, knives, blunt objects to brute force.

Table 4. Number of crime in Newark

Crime type	Number
Robbery*	1,239
Aggravated assault	1,089
Motor vehicle theft	4,367
Theft from motor vehicle	2,605
Burglary	1,231

* Robbery incidents include pick-pocket and snatch incidents (n=95) which are classified as violent theft.

The crime type of burglary is unique among the selected criminal offences due to the target immobility. Since residences and businesses are often separated by zoning codes, it is important to differentiate residential burglaries from commercial burglaries. For burglary incidents, Newark City Police Department separate burglaries into three categories: residential, commercial and quasi-public. Quasi-public classification included premise types such as churches or a vague description of “building” or “lot”. At times there seems to be some disagreements regarding what types of premises are quasi-public as opposed to public. Sometimes, hospital was classified as quasi-public and other times as public places. To understand this different category, it may be necessary to read the criminal incident report. For instance, the hospital lobby may be public but the patient's room in the hospital may be quasi-public. As it can be seen in Table 5, the majority of burglary was residential (75.4 percent). Only a handful cases were classified as quasi-public. Due to small numbers of non-residential burglaries, comparison of residential burglary to non-residential burglary is not suitable.

Table 5. Burglary incidents separated by premise type in Newark

Premise type	Number (%)
Residential	928 (75.4)
Quasi-public	64 (5.2)
Commercial	237 (19.3)
Missing information	2 (.2)
Total	1,231(100.0)

2.2. Mass Transit: Buses, Subways, and Commuter Trains

The routes and stops of mass transit are good measures to construct awareness space in crime pattern theory to approximate the areas known to a larger number of people. The mass transit data were obtained from New Jersey Transit GIS Team. The mass transit data are composed of routes and stops of buses, light rails, and commuter rails. Figure 2 shows the mass transit systems in Newark, NJ.

Upon the examination of the routes and stops, in September of 2007, there were more than 242 bus lines traveling through or to Newark. Among these, 28 bus lines were identified to have their center in Newark. There were a total of 863 separate bus stop locations. When allowing double counting (i.e. if more than one bus line shared the same bus stop, then this was counted accordingly by how many bus line stopped at that location), it was added up to 2,118 bus stops.

Figure 2. Mass transit stops and stations in Newark

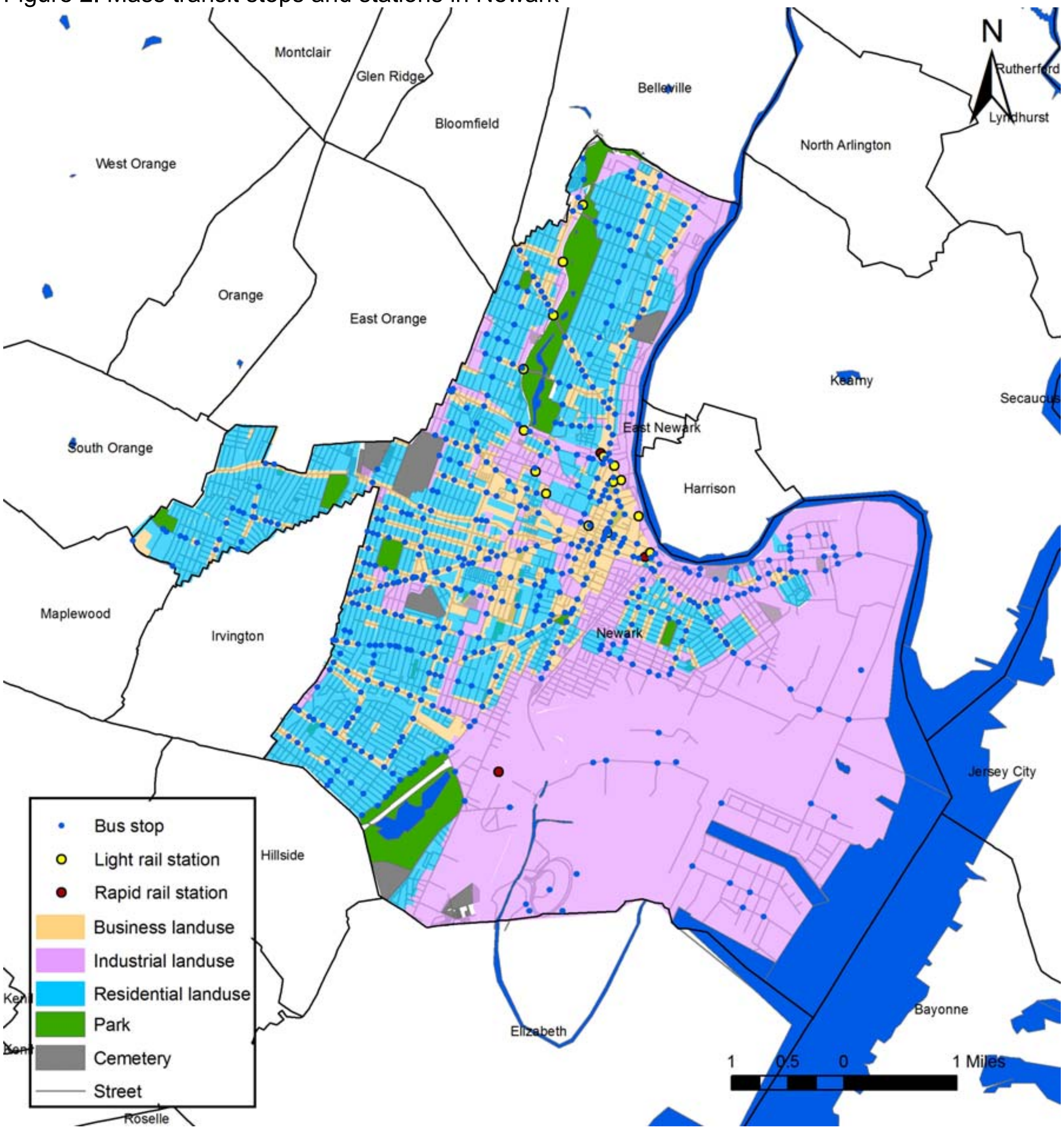


Table 6. Distribution of mass transit stops and stations in Newark by zoning code

Zoning district	Bus stops(n=2,118)	Subway stations(n=15)	Rail stations(n=3)
Business	1,483	10	0
Industrial	606	6	3
Residential	653	2	0
Mixed land use	583	9	0

Note: The mass transit stops and stations may be counted more than once depending on their locations.

The locations of bus stops in relation to zoning code were examined to obtain a better picture of bus system in Newark. An examination of bus stops revealed that many bus stops were located on the main arterial roads. These main roads, however, often did not belong to any zoning code. Therefore, simply counting the number of bus stops within each zoning code led to not counting bus stops located on main roads. To address this problem, bus stops as far as 100 feet away from the zoning code were counted to obtain number of bus stops in each zoning districts. For instance, suppose that there is a bus stop located on a main road without land use designation. Further, suppose that one side of the main road is residential area and the other business. Since the bus stop is located within 100 feet of residential and business zoning codes, this bus stop would be counted twice and classified as being located in mixed land use areas. This method led to counting every bus stop at least once but no more than three times; residential, business and industrial land use.

As it can be seen in Table 6, the majority of bus stops (1,483 bus stops out of 2,118) were located in business zoning areas whereas only 653 bus stops were located in residential areas. This supports the earlier suspicion regarding the influences of bus stops on burglaries; that many bus stops are located in business areas rather than residential areas. Therefore it would be difficult to hypothesize the bus stops will increase crime opportunities for burglary in reality despite some theoretical supports.

A total of 583 bus stops (27.5 percent) were located in areas where different zoning code began within 100 feet away from its location. Among these 583 bus stops, 217 bus stops (37.2 percent) were located in areas between business and residential, 133 bus stops (22.8 percent) for industrial and residential, and 192 bus stops (32.9 percent) for industrial and business zoning codes. Only 41 bus stops (7 percent) were located in close proximity of all three zoning codes.

2.3. Verizon's YellowBook Data

It is often thought that certain types of facilities or commercial establishments increase crime in the local neighborhoods. Since bus stops are located in bigger environments, it is important to examine what types of facilities and businesses exist around the bus stops in relation to crime. Criminologists with ecological orientation often pay attention to the impact of facilities linked to vice or cash handling businesses. The facilities shown to increase crime opportunities include liquor-related establishments (Block and Block, 1995; Loukaitou-Sideris et al., 2001), pawn shops, and check cashing facilities (Loukaitou-Sideris et al., 2001).

Obtaining the official records of business locations turned out to be an obstacle if not impossible. Scores of phone calls to city hall to different departments to obtain the records on the business locations were proven to be fruitless. One of the most discouraging information was that the business license permit data were not computerized. Not only obtaining a paper copy of business locations would be prohibitively expensive but also the data entry process would be very time-consuming even if paper copies of the locations were to be obtained. Therefore, it was decided to use the phonebook database to measure the business locations instead.

The Verizon's YellowBook data used in the present study was compiled in October 2007. Use of the YellowBook data to measure commercial activities in the areas poses some concerns in relation to accuracy and representativeness of the population. For representativeness of the data, the percentage of businesses subscribing to the listing service is simply unknown. Regarding the accuracy of the data, there seemed to be many duplicates or incomplete business listings in the data.

After a closer examination of the YellowBook data, there were a couple of questions concerning the utility of the database. First question was related to the acceptable matching rates between the official data and the YellowBook data. The comparison of official data and the YellowBook data was not possible since it was the inaccessibility to the suitable government data that led to use of the YellowBook data. Second question was the likelihood of obtaining YellowBook services for different business categories. In other words, it was not clear whether subscription rates to YellowBook service would be the same regardless of the business categories. The fact that YellowBook listing service offers multiple pricing plans ranging from free to more than $100 a month indicates the needs and demands for the listing services are different depending on the type and size of businesses.

The YellowBook data used in this research originally contained a total of 18,782 business listings in Newark. However, it was soon discovered that there were many duplicate records in the data. There are several reasons for duplicate records. First, the basic listing is free - when there was an error or missing information, it appeared that customers simply added a new listing rather than modifying the existing records. Second, in some cases, one listing would be in English but another listing would be in second languages such as Spanish or Italian, etc. Third, businesses with more than one phone number appeared to be listed multiple times displaying several different phone numbers. The cases identified with having the same address, the same name or the same phone numbers (n= 2,893) were deleted from the dataset using SPSS 15.0.

Once the duplicate records were cleaned, the next step involved deleting the cases that could not be geocoded. There were 505 cases without addresses and 732 cases with PO Box addresses. Often the ones with PO Box addresses were one of multiple listings belonging to the same company or franchise. For instance, Burger King may have several locations where food is served but one of the listings would include a PO Box where customers call or send documents to participate in some sort of promotions. These PO Box addresses cannot be geocoded and it is unlikely that these locations would influence either legitimate or illegitimate activities in the area. Therefore, a logical decision was to delete these cases from further analyses. The listings without addresses appeared to belong to contractors who travel to multiple areas to perform services. For instance, contractors such as plumbers or carpenters often visit the sites to give estimates and the customers may not necessarily know where and whether these individuals have physical business establishments. For these contractors, there is no need to advertise where their actual business locations are.

In addition, there were 4,914 cases (26.2 percent) without phone numbers. Since YellowBook data are phonebook data, it was not clear listing the businesses without phone numbers would prove useful. It is possible to argue that a determined prospective customer will drive to the specified location even if the listing may not provide a phone number. However, it is unlikely that there are many prospective customers with such determination. These business listings, if they indeed were in operation, are not likely to get benefits from YellowBook listing service since one of the first question can easily be whether the business in question is indeed in operation. Also many of these records could be duplicate cases with missing information.

After deleting duplicate records, listings without phone numbers, and non-geocodable business listings[2], YellowBook database was left with a total of 10,031 cases. Table 7 summarizes the steps taken to clean up the data.

---

[2] The cases which contained information on the street or the area name where it was located were not defined as non-geocodable cases. For instance, a business listing displaying its address as “Port Newark”

Table 7. Steps taken to clean up YellowBook database

Steps taken for data cleanup	Number (%)
Initial number of records	18,782(100.0)
Duplicates (same address, same name or same phone number)	2,893 (15.4)
No address given (address left blank)	505 (2.7)
PO Box address	732 (3.9)
No phone number	4,914 (26.2)
Cases with geocodable addresses	10,031 (53.4)

Table 8. Selected business category listings in Newark

SIC classification	Number (%)
Restaurant (with or without liquor license)	514 (5.1)
Grocery store and convenience store	228 (2.3)
Liquor store	66 (0.7)
Automobile parking	61 (0.6)
Drinking place	53 (0.5)
Gasoline service station	45 (0.4)
National commercial bank	30 (0.3)
Pawn shop and used good store	4 (0.0)
Sub total	1,001 (10.0)
Total number of YellowBook cases	10,031(100.0)

Table 9. Business subdivision using SIC classification method (n=10,031)

Business division	Sub division	Number (%)
Retail trades n = 1,854 (18.45%)	Building materials, hardware, garden supply, and mobile home dealers	57 (0.6)
	General merchandise stores	84 (0.9)
	Food stores	324 (3.3)
	Automotive dealers and gasoline service stations	185 (1.9)
	Apparel and accessory stores	163 (1.7)
	Home furniture, furnishings, and equipment stores	123 (1.3)
	Eating and drinking places	567 (5.8)
	Miscellaneous retail stores, (not classified)	348 (3.6)
Finance	Depository institutions (Banks and credit unions)	94 (1.0)
Services n=1,788 (17.82%)	Hotels, rooming houses, camps, and other lodging places	27 (0.3)
	Personal services	559 (5.7)
	Business services	335 (3.3)
	Automotive repair, services, and parking	413 (4.2)
	Miscellaneous repair services	309 (3.2)
	Motion pictures	43 (0.4)
	Amusement and recreation services	102 (1.0)
Other types of business listings		6,298(64.2)

Types of facilities that were theorized to influence crime opportunities in the areas are summarized in Table 8. As it can be seen, some of the selected business categories turned out to be very few. Except for the restaurants (n=514) and grocery stores (n=228), none of the business categories was more than 100. Considering the

---

with a phone number was not deleted. The same is true for the listing displaying its address by street intersection or only with street name.

size of Newark and the quantitative nature of the study, this measurement did not seem to be appropriate.

Table 10. Business categories under major SIC groups

Major group	Business category
Personal services	- Power Laundries, Family and Commercial - Garment Pressing, and Agents for Laundries and Drycleaners - Linen Supply - Coin-Operated Laundries and Drycleaning - Drycleaning Plants, Except Rug Cleaning - Carpet and Upholstery Cleaning - Industrial Launderers - Laundry and Garment Services, Not Elsewhere Classified - Photographic Studios, Portrait - Beauty Shops - Barber Shops - Shoe Repair Shops and Shoeshine Parlors - Funeral Service and Crematories - Tax Return Preparation Services - Miscellaneous Personal services, Not Elsewhere Classified
Business services	- Advertising Agencies - Outdoor Advertising Services - Radio, Television, and Publishers' Advertising Representatives - Advertising, Not Elsewhere Classified - Adjustment and Collection Services - Credit Reporting Services - Direct Mail Advertising Services - Photocopying and Duplicating Services - Commercial Photography - Commercial Art and Graphic Design - Secretarial and Court Reporting Services - Disinfecting and Pest Control Services - Building Cleaning and Maintenance Services, Not Elsewhere - Medical Equipment Rental and Leasing - Heavy Construction Equipment Rental and Leasing - Equipment Rental and Leasing, Not Elsewhere Classified - Employment Agencies - Help Supply Services - Computer Programming Services - Prepackaged Software - Computer Integrated Systems Design - Computer Processing, Data Preparation and Processing Services - Information Retrieval Services - Computer Facilities Management Services - Computer Rental and Leasing - Computer Maintenance and Repair - Computer Related Services, Not Elsewhere Classified - Detective, Guard, and Armored Car Services - Security Systems Services - News Syndicates - Photofinishing Laboratories - business services, Not Elsewhere Classified

As an alternative, a broader business category than the ones in Table 8 was used. Table 9 shows the selected business division according to the Standard Industry Code (SIC). The SIC classification is a four digit code used to differentiate business categories within specific industries. In Newark, there were 1,854 (18.5 percent) retail establishments, and 1,788 (17.8 percent) service establishments excluding social services and membership organizations in October 2007. Among the retail trades, three types of retail categories were chosen for the present study; eating and drinking places, food store, and automotive related retail establishments. For the service industry, three types of business categories were chosen; personal, business, and automotive related services. For the subdivisions of SIC, most of the categories were self-explanatory except for the categories of personal and business services. The sub-business categories for Personal and business services major groups are provided in Table 10.

2.4. Zoning Districts: Land use Elements

Existing research studies show that facilities without place managers such as parks (Newton, 2004) or vacant buildings (Loukaitou-Sideris et al., 2001) are associated with increased number of criminal incidents. Other environmental features linked to higher number of crimes include multi-family housings (Loukaitou-Sideris et al., 2001), public housings (Loukaitou-Sideris, 1999), and schools (Newton, 2004). Therefore, it would be useful to account for land use patterns and facilities in Newark.

The Land Use Element mandated by the State’s Municipal Land Use law provides a basis for the City’s zoning ordinance which designates permitted land use (City of Newark and Philips Preiss Shapiro Associates, 2006). Newark zoning ordinances contain mainly three types of land use elements: residential, industrial, and commercial. The land use elements exclude designation on roads, parks, rivers, and cemeteries, therefore, its total area size does not add up to the estimated area of Newark, 26 square miles. Figure 3 shows the zoning districts in Newark. From Figure 3, it can be seen that the business land use elements are located along the major roads, resembling an image of veins running through a body. The central business district is located in the east side of Newark at the center. The majority of the west side is residential areas, and the lower east side area is mostly industrial areas. This signals the heterogeneity of Newark as a whole.

Figure 3. Land use zoning districts of Newark

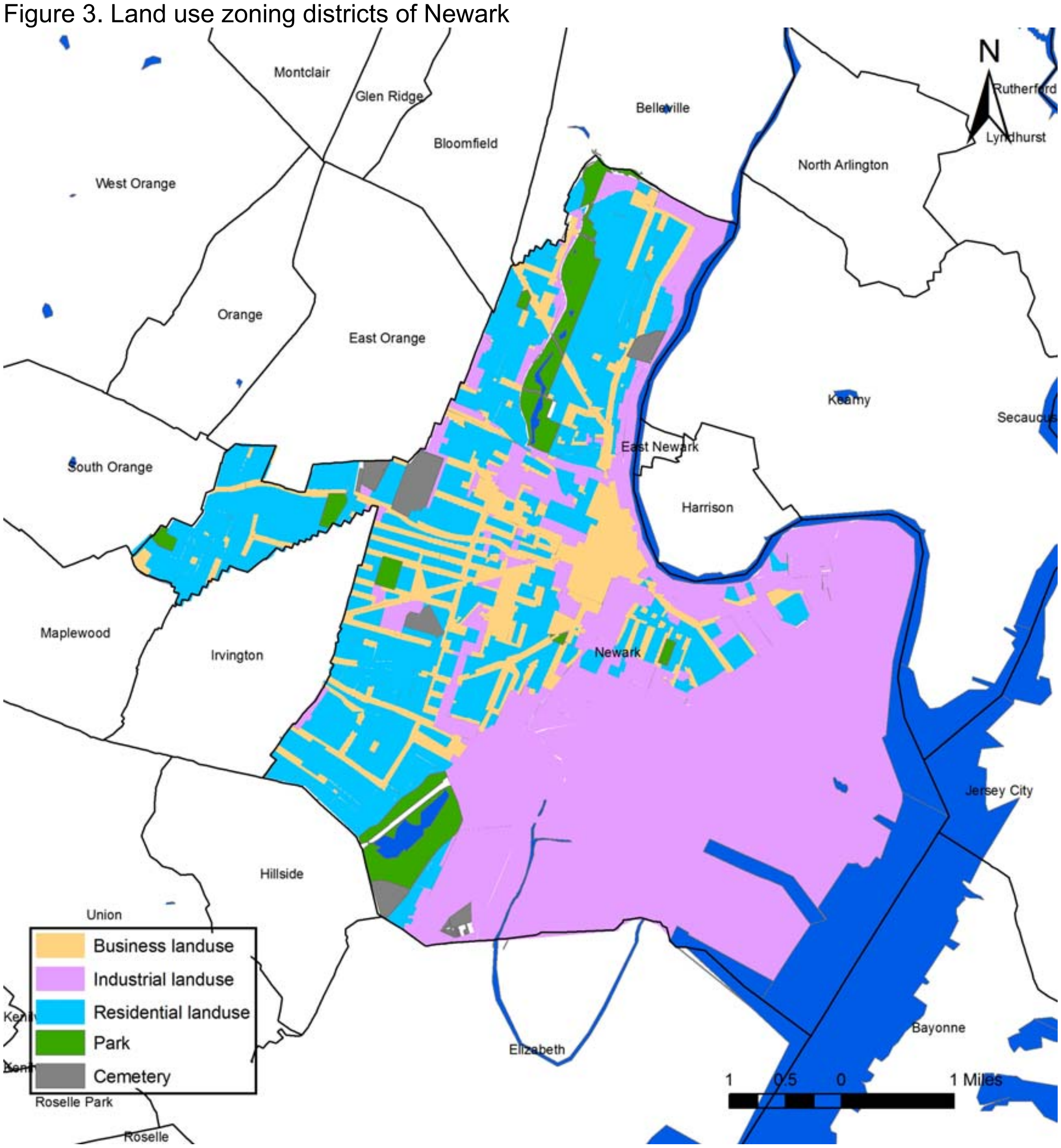


Figure 4. Heavy industrial land use designations and areas without residents in Newark

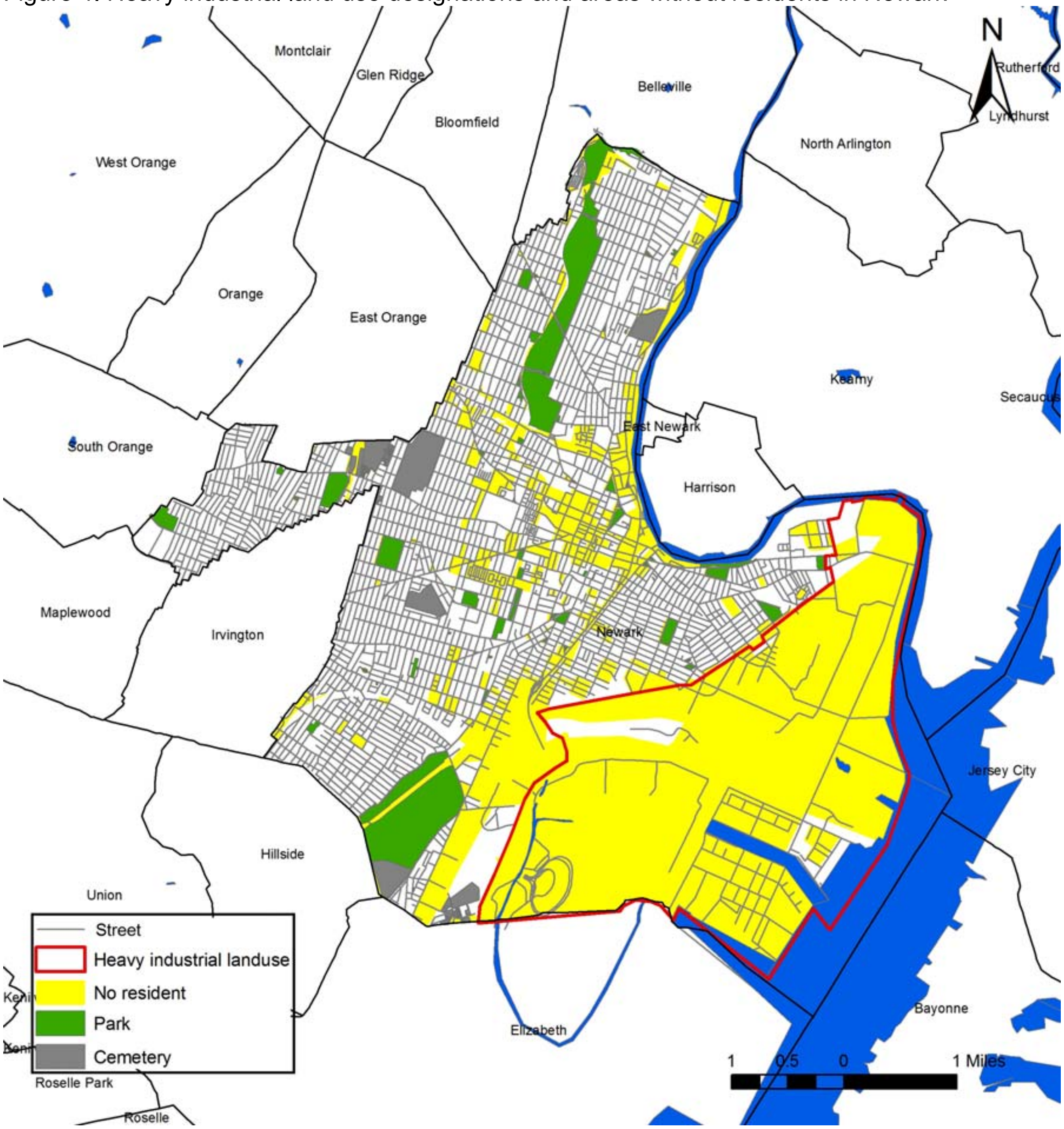


Table 11 shows the summary of land use elements in Newark. In Newark, the designation of business land use is the smallest (13.6 percent) in its size and the largest land use category is industrial, 56.8 percent in its size. Only about 30 percent of Newark is designated as residential areas. Each of the land use element is again divided to permit low, medium, and heavy use of respective categories (City of Newark and Philips Preiss Shapiro Associates, 2006).

Table 11. Newark city land use zoning district and their sizes in square miles

Zoning district	Number of plots (%)	Size in square miles (%)
Business	199(35.9)	3.1(13.6)
Neighborhood commercial	24(12.1)	.2(7.7)
Community commercial	119(59.8)	1.7(54.8)
Regional commercial	45(22.6)	.5(14.8)
Shopping center commercial	11 (5.5)	.7(22.6)
Industrial	127(22.9)	13.0(56.8)
Light use	91(71.7)	2.2(17.3)
Medium use	24(18.9)	1.0(7.6)
Heavy use	12 (9.5)	9.8(75.1)
Residential	229(41.3)	6.8(29.7)
Low-density	18 (7.9)	2.3(33.6)
Medium-density	53(23.1)	2.4(35.9)
High density	113(49.3)	2.4(35.9)
Multi-family	45(19.7)	1.2(17.1)
Total	555(100.0)	22.9(100.0)

* Parks, street, roads, cemeteries, and rivers are excluded from the zoning designations.

There are four subgroups under business land use designation. Most notable designation is community commercial designation which is designed to offer convenient access to the residents. This designation radiates major arterial and retail corridors in Newark (City of Newark and Philips Preiss Shapiro Associates, 2006) and this can be seen in Figure 4 here business land use designations are observed along the major roads.

The heavy industrial land use designation in East Ward contains the Newark Liberty International Airport and Port Newark neighborhoods. Notably, one heavy land use element exceeds the size of eight square miles. Figure 4 shows city of Newark with heavy industrial land use element and areas without any residents according to the 2000 census data.

Not surprisingly, the area with heavy industrial land use designation in East Ward seems drastically different from other areas in two respects. First, the local street networks are not well developed in this area. The poor street network is reflected with the addresses recorded by the police as well as in the business listings in YellowBook data. The addresses in this area were often written as foot of or beginning of street name (i.e. foot of Pacific Street) rather than having regular numbered addresses[3]. The cases with poor addresses will not be geocoded whether they are criminal offences or businesses. Therefore, the geocoding results will be biased due to the exclusion of the cases with vague or with highway addresses.

Second, there are no residents in the areas with heavy industrial land use designation since no residence is allowed in the areas. Figure 4 shows the areas without resident according to the 2000 census. There are smaller industrial land use elements in other parts of Newark. However, these smaller heavy industrial land use elements are quite different from the heavy industrial land use elements containing the Newark Liberty International Airport and Port Newark. First, these smaller land use elements are in close proximity to other land use designations. Second, these smaller industrial land use elements have good street networks within the plots, and are easily accessible within the city. These facts, their pocketed nature and well developed street network, make the high levels of spatial interaction within the city reasonable and probable. However, the area in the East ward is somewhat isolated from the city's everyday activities due to its distance and barren landscape. In this area, there are no residents, and under-developed street networks. Due to the heterogeneity of the area from the rest of the city, this heavy industrial land use area will be excluded from the data analyses.

[3] Businesses in this area often list their addresses in vague terms by listing intersection or the general area where they are located such as Port Newark. Some of addresses in this area have highway addresses which may not be accessible from the local roads. All of these incidents are not geocodable, therefore, they present a bias in geocoding results.

3. Selecting Appropriate Areal Unit: Thiessen Polygon

Examining the impacts of small locations or places on crime using larger units may be problematic if heterogeneity exists within the selected areal unit (Boyd et al., 2007). The areal unit selected need to be appropriate in theoretical frameworks. In addition, it needs to be possible to assume homogeneity in dependent and independent variables within the areal unit.

3.1. Modifiable Areal Unit Problems

Spatial pattern analyses often require counting number of crimes within the given area to allow researchers to perform statistical tests or make meaningful comparisons with other areas. Yet spaces can be divided in many ways. These spatial boundaries are often arbitrary and they may not be meaningful in respect to the phenomena of our interests. Whether the phenomena in our interests occur respective to the areal boundaries or not, the interpretation and policy implications are often made based on these arbitrary administrative boundaries (Anselin, 1988).

One of the concerns of aggregating crime data to areal unit is inconsistency in spatial patterns. Depending on the areal units used, spatial patterns may be different or even distorted (Wong and Lee, 2005). This problem, known as Modifiable Areal Unit Problem (MAUP), needs to be considered in statistical analyses or in its visualization due to their pervasive impact. Another related problem with MAUP is that the statistical measures and interpretation of the test results may differ depending on the spatial units used. For instance, the magnitude of autocorrelation and parameter coefficients in models may be different depending on which areal units are used (Anselin, 1988).

There are at least two suggestions made in addressing MAUP. One is using more than one areal unit in the analyses to compare and contrast the results (Wong and Lee, 2005). The other suggestion is using relatively scale insensitive spatial analytical techniques (Wong and Lee, 2005). Due to the limited success of spatial techniques attempting to address MAUP, some have suggested simply acknowledging the existence of MAUP and identifying the sources of the inconsistency.

3.2. Thiessen Polygon Using Street Intersection as the Center

There are two things to consider when deciding the spatial unit. First is the theoretical focus: what is the appropriate spatial resolution needed to answer the research questions? One of the opportunity theories used in this research is rational choice theory which brings the study focus to small places rather than neighborhoods levels. Therefore, there is a need to use spatial unit that is suitable for the micro-level focus of the present research study. Second, the selected spatial unit needs to allow the assumption of homogeneity in independent variables such as accessibility to bus stops or businesses within the unit. In both cases, use of smaller size of areal unit is warranted.

A review of the previous research studies produced no clear guidance in selecting spatial units. For subways, the areas thought to be influenced by the subway stations was about a quarter mile (Block and Davis, 1996) and up to one mile radius (Poister, 1996). For bus stops, the areal units used by researchers are quite small, a 150 feet radius using the street intersection as the center (Liggett et al., 2001).

Use of existing boundaries such as census tract and census block presents various problems. The scale of census tract is bigger than the focus of the study. Census block seems to fit the theoretical focus of the study but its use will most likely separate crimes on the street corners or intersections into different areal units. The same will be true for commercial establishments. However, the criminal offences and businesses at the corners are influenced by and influence other activities in adjacent street blocks. Therefore, the selected areal unit needs to be able to place criminal offences or businesses at the street corners to the same areal units rather than to different ones.

One way to address this problem is dividing the study area using street intersection as the center of the areal unit. This can be accomplished by drawing buffers around the intersections to create unit of analysis. However, use of buffers as areal unit may lead to two different problems. If the street segments are shorter than the defined buffer radius, then the buffers will overlap with each other. On the other hand, if the street segments are long, then the buffers will exclude the middle of the street segments. This will lead to the exclusion of crimes located on the middle of the streets which was the case with Liggett and her colleague's study. To avoid these possible complications, it was decided to create an areal unit using the street intersection as the center. In this case, each street segment will be divided in the middle to be assigned to the closest intersection. This process will continue until the Thiessen polygon areal unit is created. The use of Thiessen polygons as the areal unit will create no overlap between the areal units without excluding any portion of the street segments.

Thiessen polygon is unique in its characteristics that any point within the polygon is the closest to where it belongs. This means that crime and business will be assigned to the closest street intersection. Indeed, there is a recent development in the use of Thiessen polygon using street intersections as the center. For instance, in a study analyzing gang-related drug offences, Ratcliffe and Tangiuchi (2008) used Thiessen polygon using street intersection as the anchor point.

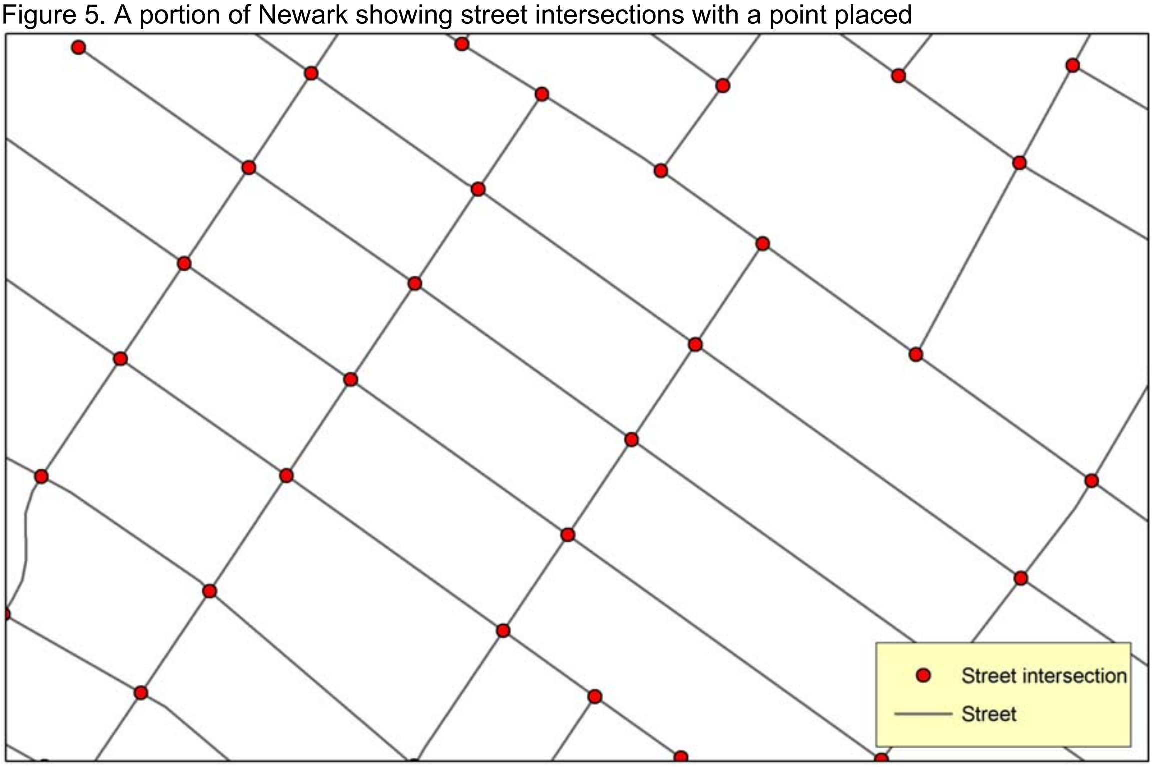
The first step in creating Thiessen polygon was to place a point on every street intersection. Using X-tools Pro 5.2, a point was placed on each street segment longer than 50 feet on both start and end locations. This produced many duplicate points since different street segment touch each other at the intersections. The duplicate points (i.e. the points with the same x- and y- coordinates) were deleted using SPSS 15.0. Figure 5 shows a portion of Newark where a point was placed on street intersection.

After deleting duplicate points placed on the street intersections, a Thiessen polygon layer was created using an arcscript downloaded from ESRI user forum. When the created Thiessen polygon layer was examined, there were two visible problems. First, some Thiessen polygons were very small having no more than a total of 50 feet of street length within the unit. Two causes were identified. First cause was short street segments producing closely located street intersections. These small Thiessen polygons were merged with their neighbor. Second cause was the 240 private streets where public access are generally blocked off and the city is not responsible for upkeeping of them. These streets are often the street within an apartment complex that connects between parking lots. These small Thiessen polygons covering private streets were combined together to have a few number of polygons rather than having 10 or 20 small polygons covering one apartment complex.

Second problem was that some Thiessen polygons included streets that were physically disconnected. This problem was due to the fact that Thiessen polygon was created using Euclidian distance without considering whether it is possible to travel from one street to the other. The logic of using Thiessen polygon as the areal unit was to assign crime and business to the nearest street intersection. For this reason, the intersection within the areal unit needs to be connected to the street where crime occurred even if it may not be the closest intersection using the straight line. Figure 5 shows the problem where several Thiessen polygons contain physically disconnected streets within the unit violating logic behind the decision making process.

The Thiessen polygons containing disconnected streets within the units were edited to have only physically connected streets[4]. This process ensured that the areal unit employed made a logical sense that it did not include physically disconnected or separated streets in the same areal units. Figure 7 shows Thiessen polygons after completion of the editing process. After the editing process, there was a total of 2,750 Thisseen polgyons.

Figure 5. A portion of Newark showing street intersections with a point placed



[4] This process requires problematic Thiessen polygons be divided into two or more parts, then combined with their neighbors to have only physically connected streets in the units.

Figure 6. A portion of Newark showing Thiessen polygons containing physically disconnected streets within the same areal unit

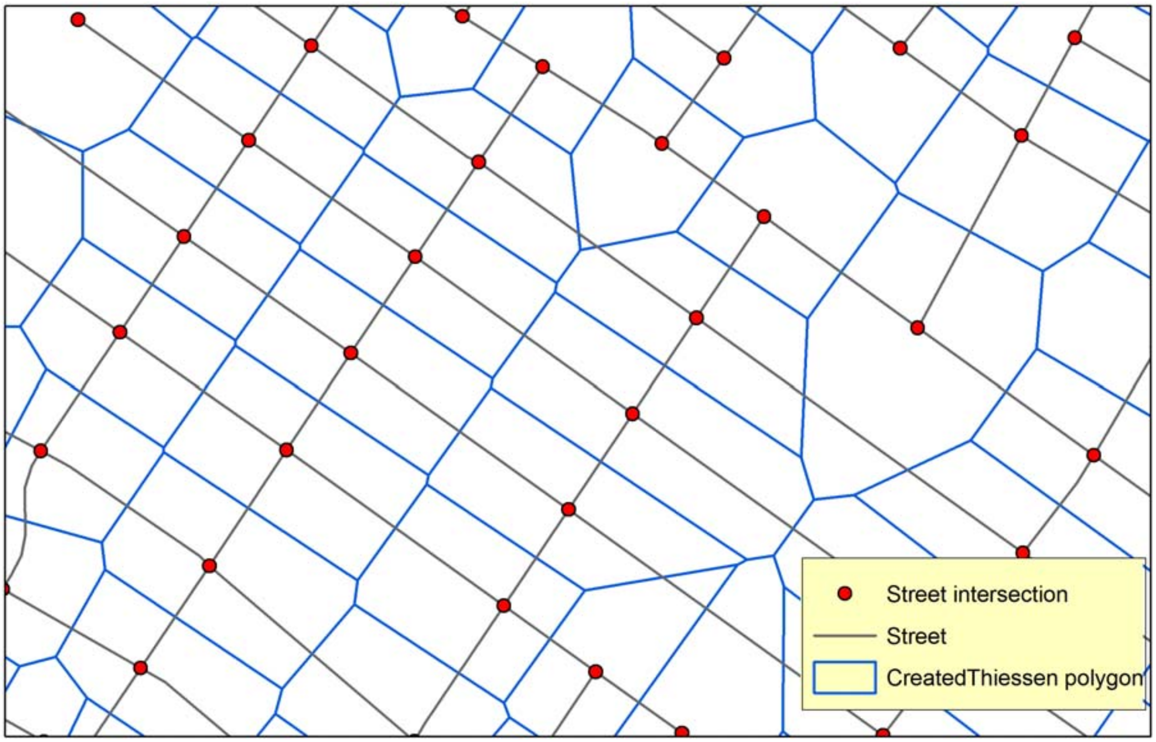
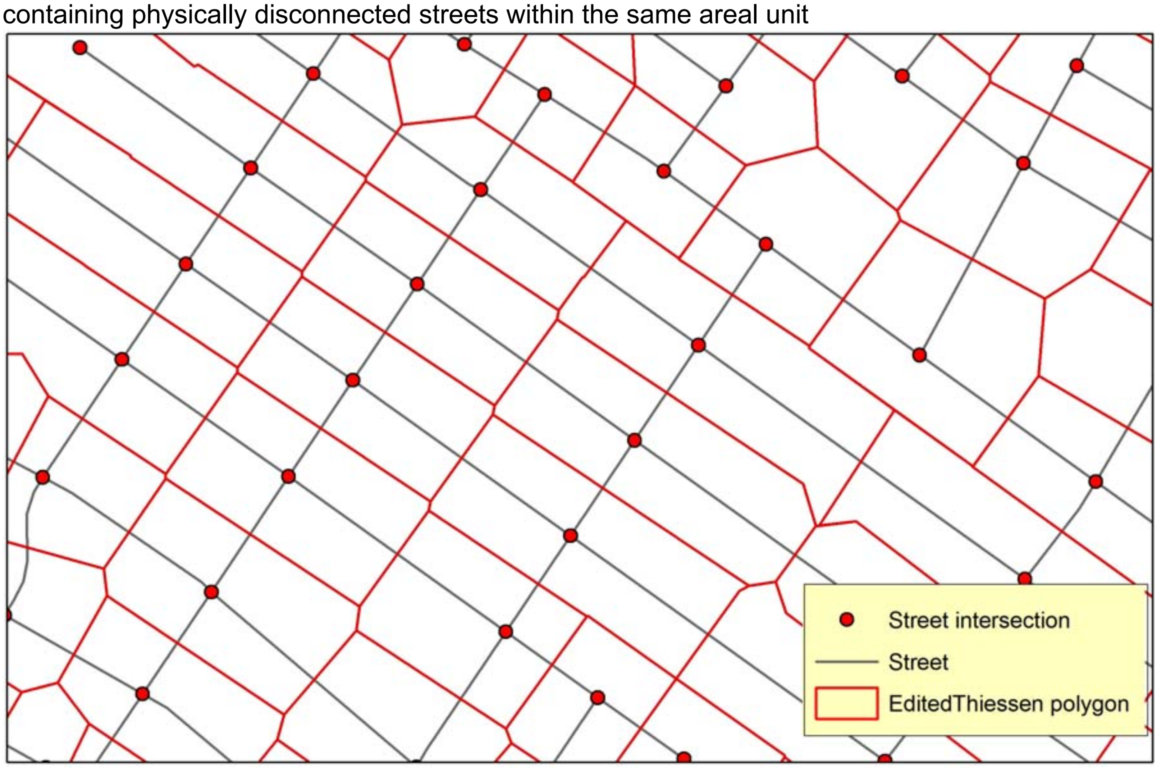


Figure 7. A portion of Newark showing Thiessen polygons after editing process to avoid containing physically disconnected streets within the same areal unit



4. Addressing Spatial Aspects of Data

Use of classical statistics is often inappropriate for spatial data for two reasons. First, spatial data usually contain information on locations, and the classical statistics do not use spatial information in their analysis (Wong and Lee, 2005). Not analyzing geographical features or impacts of the data may defeat the purposes of collecting and analyzing spatial information. Second, most classical statistical tests assume independence of observations while most spatial data display some level of spatial dependence (Anselin, 1988). A significant level of autocorrelation produces the same effect as having duplicate observations, possibly leading to an incorrect conclusion from statistical tests (Wong and Lee, 2005). Therefore, when the assumption of independent observation is violated, it is necessary to use spatial econometrics.

The field of regional science recognizes the importance of the location and distance in human activities (LeSage, 1998). When investigating spatial patterns, it is pertinent to examine whether the geographical locations play a role in the attribute distribution. If there is a significant level of spatial dependence in the data, then spatial statistics such as spatial regression modeling or spatial econometric modeling needs to be performed accordingly (Graaff et al., 2001). Spatial econometrics is a collection of tools designed to take account of spatial effects of data (Anselin, 1988). The spatial effects are divided into spatial dependence and spatial heterogeneity. In this section, these two concepts will be discussed in detail.

4.1. Spatial Dependence and Spatial Heterogeneity

Spatial dependence, also called spatial autocorrelation, occurs when the neighboring values influence the observed value (Wong and Lee, 2005). In general, spatial dependence should be related to the distances; that spatial units closer to the observation should have greater influences than the spatial units located farther away (LeSage, 1998). In other words, there should be distance decay function working between the distance and the strength of spatial dependence.

Spatial dependence can be caused by a couple of reasons. First, spatial dependence may be due to measurement errors when data related to individual characteristics are collected at an aggregate areal unit. If there is a disjunction between the underlying process of the data collected and areal unit used, this may cause the observed characteristics to spill over across different areal units, possibly causing spatial dependence or spatial autocorrelation (Anselin, 1988).

The second cause of spatial dependence is related to human behavior and human geography (Anselin, 1988). The locations and distances are important factors influencing spatial interaction, and they may lead to interdependencies of human behavior in space. For this reason, an observation of any given space is influenced by what happens in other places. This will likely cause some level of spatial dependence.

Spatial heterogeneity basically refers to the fact that the relationships under investigation may display variation over space (LeSage, 1998). There are two aspects to spatial heterogeneity: structural instability and heteroskedasticity (Anselin, 1988). Many things we study display structural instability over space (Anselin, 1988). In addition, the use of ad hoc spatial units may cause measurement errors that may vary with locations, and characteristics of the spatial units. This type of spatial heterogeneity is based on spatial structure or areal unit. Also spatial heterogeneity is often reflected in measurement errors in forms of missing values or functional misspecification, and they may lead to heteroskedasticity, violating the assumption of constant variance in error term (Anselin, 1998). Ignoring these aspects may bring threats to statistical conclusion validity due to misleading significant levels (Anselin, 1988). To summarize, structural instability or non-homogeneous spatial units may cause heteroskedasticity (Anselin, 1998).

Spatial dependence and spatial heterogeneity are often linked to each other (Graaff et al., 2001). In strictly observational sense, both spatial dependence and heterogeneity may be due to the geographical features (Graaff et al., 2001). For instance, robbery may cluster as well as bus stops in different areas in the study region. In this case, spatial heterogeneity as well as spatial dependence would be a type of spatial effect. In addition, spatial dependence is a form of heteroskedasticity. Due to this interconnection, sometimes, tests for spatial dependence or heteroskedasticity may not be capable of distinguishing between the two.

4.2. Constructing Spatial Weight Matrix

One of the most important step in data analysis lies on how to define neighbors since this definition brings several implications for the estimations and test statistics (Anselin, 1988; Ward and Gleditsch, 2008). How the neighbors are defined will influence not only the value of spatial lag but also the level of autocorrelation measured. The positive value of Moran’s I statistics indicate clustering, and the negative values indicate the opposite (Wong and Lee, 2005).

To incorporate spatial dependence in modeling, it is necessary to construct spatial weight matrix. The spatial weight matrix contains information on which spatial units in the system are thought to influence the observed value of the spatial units (Anselin, 1988). This weight matrix can be used to create the spatially lagged variable which represent an average value of the neighboring areal units in addition to assessing the spatial dependence levels (Ward and Gleditsch, 2008).

Although deciding which areal units will be considered as one's neighbors is arguably one of the most important step in data analysis (Ward and Gleditsch, 2008), this decision is largely arbitrary, and there are several methods in constructing spatial weight matrix. Most often used methods in constructing weight matrix are based on either distance or contiguity (Anselin, 1988). However, the guiding principle should be the nature of the phenomena being modeled (LeSage, 1998).

When computing the spatial lag using shared boundaries, there are mainly two options (Wong and Lee, 2005). Rook contiguity option uses common boundaries to define neighbors which exclude corner neighbors. On the other hand, Queen contiguity uses all common points in the definition which include corner neighbors in computing spatial lag (Anselin, 2003). The spatial weight matrix created using Queen contiguity results in higher number of neighbors compared to the Rook contiguity weight matrix.

Another method in defining one's neighbors is to use distance calculated using the center of polygons (Wong and Lee, 2005). In this method, centroid (the center point) of the polygon is calculated for each polygon. By defining a distance band to be used in defining the neighbors, the polygons where their centroids fall within the user-defined distance are considered as neighbors.

Use of distance between centroids is useful if there are spatial interactions transcending the boundaries of areal units (Wong and Lee, 2005). One drawback of using distance method is that the centroids' locations are affected by the shape of polygons although recent improvements in computation methods have lessened the weakness (Wong and Lee, 2005). Nevertheless, the different sizes of the Thiessen polygons and importance of spatial interaction among areal units encourage the use of distance method for constructing spatial weight matrix.

CHAPTER 5. DATA ANALYSIS PLAN

The data analyses will consist of mainly two stages. The first stage involves exploratory spatial data analysis using point and polygon data. In this step, the spatial patterns of criminal offences will be examined. The next step involves several estimation methods to answer the research questions. For this, the first step is to perform diagnostic Ordinary Least Square (OLS) regression estimations. Based on the diagnostic OLS regression results, it will be decided whether to perform spatial lag or spatial error model estimations. Due to the non-normality of the data, another suggested estimation method is count response data estimation methods. As the last step, there will be comparisons of the results from different regression methods to examine the impacts of spatial effects on the data.

1. Exploratory Spatial Pattern Data Analysis

Before embarking inferential statistics to find out the causes of spatial patterns or distribution, it is useful to perform descriptive statistics to understand the nature of the data. The first step of exploratory spatial pattern data analysis is analyzing the point data. This will provide an important information on where criminal incidents clusters or how dispersed they are. Once the point data analyses are completed, then the next step is to analyze data using areal units. This will permit the comparison of the areal units and use of inferential statistics.

1.1. Mean Center and Standard Directional Ellipse

One of the first steps in data analysis is often obtaining descriptive statistics. As with the traditional statistics, measures of central tendencies may be useful to find out the center of the distribution. The mean center in mapping correspond to the average or mean value in classical statistics (Wong and Lee, 2005). The mean center is useful for two purposes; tracking changes in the distribution, and comparing the distribution of different features (Mitchell, 2005). For this research, the mean center of each crime type will be calculated and compared with other type of offences.

The standard distance corresponds to standard deviation in classical statistics (Wong and Lee, 2005). Spatial data often display directional bias that may be influenced by the shape of the city or other geographical features. It is important to consider the influence of geography for the present research since the study area forms the shape of a crooked cross. To take account of this irregular shape of the study area, standard directional ellipses, a logical extension of the standard distance circle, will be used to capture the directional bias in crime distributions (Wong and Lee, 2005). The standard deviation ellipse of different crime types can be used to evaluate the relative dispersion of the point data (Wong and Lee, 2005). In addition, the standard directional ellipse can reveal whether the crime display directional bias as well (Mitchell, 2005).

1.2. Cluster and Distance Analyses

Identifying the mean center needs to be differentiated with identifying the cluster, or hot spot (Levine and Associates, 2004). Often, the mean center is not crime hot spot since the mean center is the average of X and Y coordinates of all observations (Mitchell, 2005). One way to identify crime hot spots is to perform cluster analysis (Levine and Associates, 2004). Cluster analysis identifies areas where criminal incidents occur in close proximity to each other. In this research study, two techniques will be used to analyze spatial clusters: the fuzzy mode, and Nearest Neighbor Hierarchical (NNH) clustering routine.

Both the mode and fuzzy mode identify the clusters of point locations (Levine and Associates, 2004). However, because the mode uses unique X and Y coordinates when finding the clusters, the mode does not group incidents closely located to each other with different coordinates. This means that even if one street segment may have several criminal incidents, if these criminal incidents have different X and Y coordinates from each other, they will not be grouped together. The fuzzy mode, on the other hand, counts the number of points within the user-defined search radius to calculate the frequency of the point data. This allows the identification of small hot spots where number of criminal incidents occurred in close proximity to each other even if they are not at the exactly same locations. However, in fuzzy mode point data may be counted multiple times as long as they meet the user-defined criteria.

The NNH clustering is another way of investigating crime hot spot using point data (Levine and Associates, 2004). The CrimeStat's NNH routine groups point data to each other using the user-defined distance band. The NNH routine will continue until all clusters either meet the criteria and be grouped together or fail to meet the user-defined criteria.

In nearest neighbor analysis, average distance is calculated between the nearest neighbors (Levine and Associates, 2004). The observed average nearest neighbor distances are compared with the expected distance, leading to a conclusion of whether the data are dispersed or clustered than expected.

1.3. Moran's I Statistics and Local Indicators of Spatial Autocorrelation

The Moran's I statistics shows how similar the observed value is with its neighboring values (Ward and Gleditsch, 2008). When the Moran's I value is positive, this indicates clustering whereas a negative value indicates dispersion (Wong and Lee, 2005). Moran's I statistics is a global measure which assesses the average correlation of the distribution, revealing whether there is spatial concentration in the distribution as a whole. To detect local variations, it is necessary to perform local measures of spatial autocorrelation (Wong and Lee, 2005).

The Local Indicator of Spatial Autocorrelation (LISA) is a local version of Moran's I statistics. As with the Moran's I statistics, a positive LISA value indicates the clusters of similar values. For instance, if the observed robbery value in the unit is high, and its neighboring values are also high, then this is positive local spatial autocorrelation. On the other hand, if an areal unit with few to no robbery incident is surrounded by areal units with high numbers of robbery incidents, then this is negative spatial autocorrelation, also termed as spatial outliers (Anselin, 2005). In LISA, the spatial cluster is identified when the value of an areal unit is more similar to its spatial lag than expected under spatial randomness assumption. As a default, the LISA statistics are calculated using 99 permutations producing a significance level of $p = .05$. However, this tends to lead to somewhat unstable results as observable by the changing results of the LISA maps when LISA is performed more than once. One recommendation suggested by Anselin (2005) is to increase the number of permutations to obtain more stable results. For this reason, the LISA maps will be created using 999 permutations. It is also important to remember that LISA identifies the core of the cluster, rather than identifying all individual locations of high-high or low-low areal units.

While LISA uses areal unit to identify the core of clusters, crime density is calculated using point data. Use of density functions allows the use of absolute occurrences as they are, without the restriction or distortion of areal unit (Mitchell, 1999). Since density functions are calculated using a uniform areal unit as defined by the user, it is particularly useful when the size of areal unit is irregular as is the case here in this study.

2. Spatial Process Model Estimations in GeoDa

To answer the research questions, regression model estimations will be performed using the OLS regression, spatial process model, and count response models. There are several reasons for performing different estimation methods. First, the OLS regression will be run as a base model that assumes no spatial variation in the relationship between predictors and crime (Cahill and Mulligan, 2007). Another reason for performing the OLS regression is to find out information on spatial aspects of the data. The OLS results in GeoDa present diagnostic sections which show structure of spatial dependence among others. Second, the spatial process model will be performed to examine the impact of spatial dependence on the estimation. Comparing the spatial process results and the OLS results will make it clear how the spatial dependence influences the coefficients of the predictors. Third, count response model will be performed to address non-normality of the data. The spatial process model is still a linear model although it takes into account the spatial effects in its estimation. On the other hand, the count response models are non-linear models. By adding spatial lags as one of the predictors, it is thought that the spatial dependence can be accounted for when using count response estimation methods.

Table 12 shows the four models which will be performed using several regression methods. The predictors are shown using variable name first followed by variable label. For ease of use, the result section will use variable names in their interpretation.

The first research question is related to the relationship between bus stops and crime. Therefore, a logical step is to build a model examining the relationships between bus stops and crime. The second question is related to the relationships between businesses and crime. The second model will examine the influence of commercial establishments on crime. The third model combines model 1 and model 2 to examine any interaction effects among the predictors, bus stops and commercial establishments. The last model (model 4) will be performed to take into account other land use patterns. The included variables are existence of mixed land use, vacant lands or buildings, public housing, college, school, and open space. For each model, spatial lag will be added for the spatial regression and count response estimation methods.

Table 12. Four regression models using predictors of bus stops, commercial establishments, and land use information

Predictor	Model 1	Model 2	Model 3	Model 4
spatial lag for spatial process model and count response model estimations	X	X	X	X
bstops : bus stops	X		X	X
sic58xx : eating/drinking places		X	X	X
sic55xx : auto dealers/gas stations		X	X	X
sic54xx : food stores		X	X	X
sic75xx : automotive repair, services/parking		X	X	X
sic73xx : business services		X	X	X
sic72xx : personal services		X	X	X
finance : banks		X	X	X
mixeduse : mixed land use (yes/no)				X
vcf1land : vacant land (yes/no)				X
phousing : public housing (yes/no)				X
univcoll : college (yes/no)				X
schools : schools (yes/no)				X
openspace: parks and cemeteries (yes/no)				X

2.1. OLS as Diagnostics in GeoDa

The spatial autocorrelation in the dependent and independent variables may strongly affect the OLS regression results. Therefore, performing the OLS regression on spatial data presents threats to statistical conclusion validity. More specifically, ignoring spatial autocorrelation frequently lead to rejection of the null hypothesis by underestimating the real variance in the data (Ward and Gleditsch, 2008).

However, it is useful to run OLS regression to perform diagnostics to assess spatial effects of the data. The diagnostic table from the OLS regression will show whether there are problems with spatial autocorrelation, multicollinearity, and heteroskedasticity. Based on the diagnostics, it will be possible to determine which spatial process model is suitable for the data to be performed. In addition, the OLS regression results can be compared with spatial process models to find out how spatial aspects of data influence the predictors' coefficient levels. For this reason, the first planned data analysis step is to perform diagnostic OLS regression in GeoDa.

One of the diagnostics produced by the OLS regression in GeoDa includes multicollinearity condition number (Anselin, 2005). If the indicator is over 30, this suggests multicollinearity between the explanatory variables. If this is the case, the model predictors should be examined to find a source of multicollinearity. For correct inference, assumption of normality is essential for many regression methods (Anselin, 2005). Therefore, another diagnostic test in GeoDa includes a test of non-normality (Jarque-Bera) accompanied with p value (Anselin, 2005). GeoDa also offers three diagnostics for heteroskedasticity; Breusch-Pagan, Koenker-Bassett, and White (Anselin, 2005).

In general, there are a couple of specification tests that are relevant to spatial analysis; assessing the magnitude of the spatial dependence and computing spatial lag (Anselin, 1988). The peak of spatial dependence can be identified by performing spatial autocorrelation test using varying distances (Wong and Lee, 2005). Depending on the process which spatial dependence occurs, there is a need for different solutions with different implications and inferences.

2.2. Spatial Lag and Spatial Error Estimation

When the modeling includes spatial dependence, either as a spatially lagged variable or as spatially dependent error terms, the estimation becomes more complex (Anselin, 1988). Spatial dependence may occur due to substance process or as a nuisance (Graaff et al., 2001). If the spatial dependence is due to substantive process, it requires development of spatially lagged process model. The Spatial Lag Model (SLM) requires a theoretical explanation in the spatial interaction between variables. The SLM is similar to time series model: spatial dependence is addressed by including a spatial lag in the predictor side (Ward and Gleditsch, 2008).

In the nuisance or Spatial Error Model (SEM), the spatial dependence is considered to be due to missing correlated variable, and this omission of predictor is reflected in the error term, causing spatial dependence in the regression error term (Anselin, 1988). Another possible reason for using SEM is that the study area does not coincide with the phenomena being investigated (Graaff et al., 2001).

There seems to be disagreements regarding which model to use for spatial process models. Some seems to think that SEM is by far the most relevant to cross-sectional data because most data are expected to have spatial dependence in error term when the data are collected for contiguous space and aggregated to areal units (Anselin, 1988; Graaff et al., 2001). For this reason, the SLM is seen to be applicable in investigating only specific spatial processes. However, Ward and Gleditsch (2008) suggest using SLM to consider the spatial dependence as a result of substantive process rather than seeing it as a statistical nuisance. Nevertheless, there is no priori reason to decide on which spatial process model to run before the data analysis (Anselin, 1988). For this reason, this decision will be based on the OLS regression diagnostics in GeoDa.

While a significant Moran's I statistic suggests spatial autocorrelation, it does not suggest how to address the observed spatial dependence. The Lagrange Multiplier (LM) test statistics are used to find out how to address the spatial dependence to the estimation (Anselin, 2005). When both the statistics of LM (lag) and Robust LM (lag) are statistically significant, the suitable alternative is SLM to address the spatial dependence. The statistically significant LM (error) and Robust LM (error) suggest SEM as the alternative. For this statistics to be useful, both standard and robust versions of the statistics need to be significant.

To summarize, SLM is used when there is spatial autocorrelation in dependent variable whereas SEM is suggested when there is spatial autocorrelation in error term. In SLM, the spatial dependence is added as an additional variable whereas in SEM, the OLS regression is seen to be inefficient in its estimation but unbiased (Graaff et al., 2001). This involves an adjustment to statistical inference such as using the spatial autoregressive moving average process (Graaff et al., 2001).

2.3. Assessing the Model Fits from OLS Regression to Spatial Process Model

There are mainly two steps in assessing the model fit improvements. First is to assess model fits within the same regression method using different models. Second is to compare model fits from different regression methods. To assess model fits, it is inappropriate to compare coefficient estimates from the OLS regression and spatial estimations (Ward and Gleditsch, 2008). The proper measures of model fit are Log Likelihood (LL) number, Akaike Info Criterion (AIC) and Schwarz Criterion (SC) (Anselin, 2005).For LL statistics, the bigger numbers (closer to the real line) suggest model fit improvements. For AIC and SC, the opposite is true; the smaller numbers suggest a better model fit.

In addition, there are three classic specification tests which can be used to assess the model fits from the classic OLS regression model to the spatial process model (Anselin, 2005). One of them is Likelihood Ratio (LR) test for spatial dependence produced in the GeoDa diagnostic section. This LR test is a test on the spatial autoregressive coefficient, not a test on spatial autocorrelation. In another word, the LR test is a classic specification test comparing the classic regression specification to the alternative SLM or SEM. The other classic test is Wald test which is obtained by squaring the asymptotic t-value or z-value of the spatial lag or spatial error term. The remaining test is LM (lag) or LM (error) tests.

These three tests are asymptotically equivalent but they typically produce different values, making it unclear to interpret the results (Anselin, 1988). However, the three statistics test values should produce the following result: Wald ≥ LR ≥ LM. If the three test statistics are not in the expected order, this would suggest that there is another source of misspecification (Anselin, 2005). To address this issue, it may be necessary to include new predictors or use new spatial weights.

3. Count Response Model Estimations with Spatial Lag

It is expected that spatial regression models are useful in addressing spatial dependence as well as having power against nonlinearity and heteroskedasticity (Graaff et al., 2001). However, spatial process models still belong to linear regression models unlike count response models which assume nonlinearity of the model. In nonlinear models, the change on the dependent variable with the respect to change in the predictor is not seen as constant. This varying impacts of predictors on the dependent variable makes the idea of perfect prediction unrealistic (Long and Freese, 2006).

The count outcomes are very common even though the use of regression models on count variables is relatively recent (Long and Freese, 2006). It is now widely recognized that the linear regression methods on count variables can be inefficient, inconsistent, and may lead to biased estimates. Due to the limitations of traditional statistics which assume normality of the data, the use of nonlinear modeling has increased in recent years (Graaff et al., 2001).

To use count response as dependent variables, there are several assumptions that need to be met which are relevant to the present research study (Hilbe, 2007). First, the existence of zero values needs to be confirmed to choose an appropriate regression model such as Poisson regression or truncated models. This assumption is met since the current dataset contains the areal units without criminal offence for 2007.

Second, Poisson distribution assumes the mean to be equal to the variance. If the variance is bigger than the mean, this is called overdispersion. If there is overdispersion, then negative binominal regression is suitable over Poisson regression method. Third, when there is an excess of spatial units with zero criminal incidents, it needs to be examined whether the data can be separated into two distributions to address inflated zero values (Long and Freese, 2006). In this case, the zero-inflated Poisson or the zero-inflated negative binominal regression models can be useful.

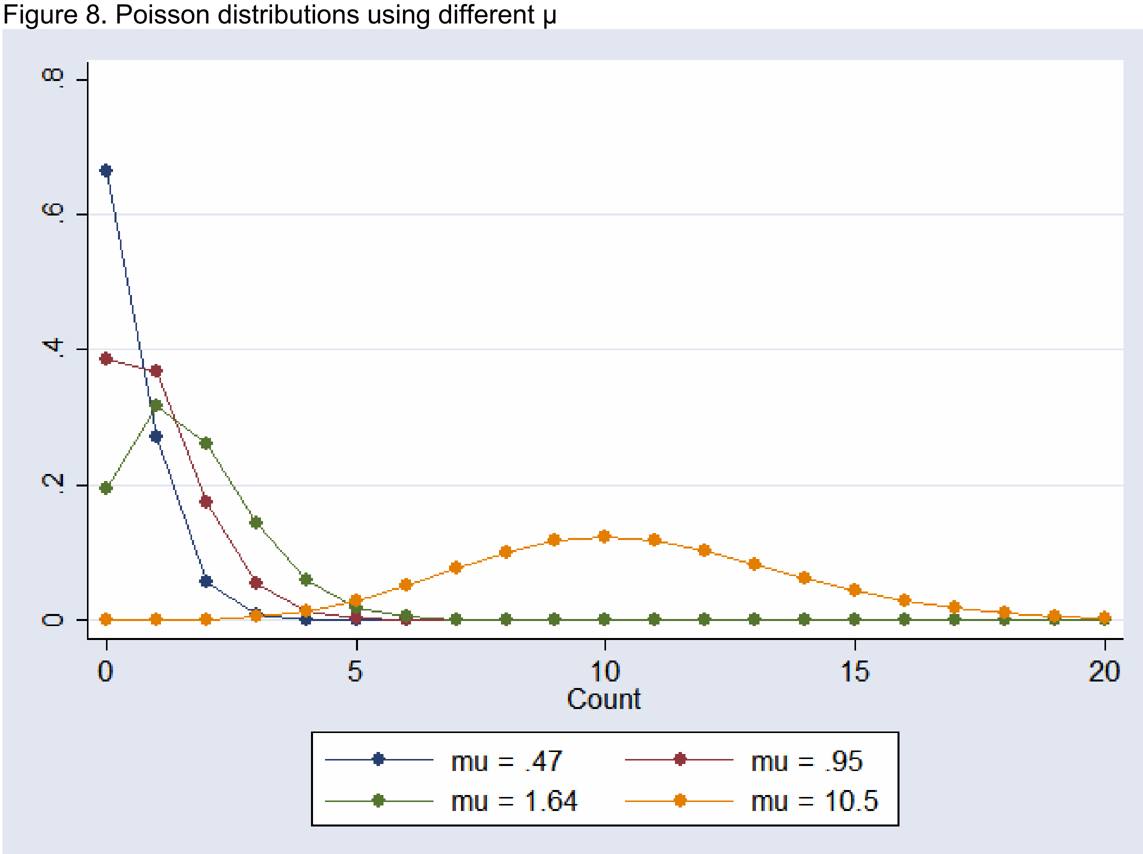
The inflated models assume that the zero counts are caused by two sources by the existence of two unobserved latent groups; one group is always zero and the other group is not always zero even though zero value is possible. For instance, it is possible that the areal units with commercial establishments may decrease the opportunities for burglary mainly because there is no residence in the unit. In this case, the variables measuring commercial establishments will influence group membership: the areal units without residences and with residences. The areal units without any residences will inflate zero values for burglary occurrences.

There are several ways to interpret nonlinear models (Hilbe, 2007) (Long and Freese, 2006). First is to compute predictions for each observation. To this end, the observed distributions will be compared with count response model distribution. Second is to compute the discrete change in the outcome based on the value of predictor. For this, the expected change in the dependent variable by predictor will be calculated. Third is to transform a nonlinear model to a linear model for easier interpretation, which will not be performed in this study.

The present study uses number of criminal offences as dependent variable. The Poisson Regression Model (PRM) is the standard method used to model count outcome data (Hilbe, 2007). Figure 8 shows various Poisson distribution using approximate mean values of five selected crime types for the present research study (see Table 21 for non-spatial descriptive statistics).

There are several important characteristics of the Poisson distribution. As the μ (the mean of the distribution) increases, the peak of the distribution shifts to the right and number of expected zeros decreases (Long and Freese, 2006). This leads to the Poisson distribution to resemble a normal distribution as μ increases. This is apparent from Figure 8 with the Poisson distribution with mean value 10.5. One way to assess the model fit is to compare the observed distribution with a univariate Poisson distribution (Long and Freese, 2006).

Figure 8. Poisson distributions using different μ



The PRM allows each observation to have a μ which is estimated from observed characteristics. The Poisson distribution is also equidispersion, meaning that the mean and the variance is the same (Long and Freese, 2006). In reality, it is often found that the Poisson distribution does not neatly fit the data, violating the assumption of equidispersion. This is because the Poisson distribution assumes that the sample members are homogeneous. For instance, the Poisson distribution assumes that all areal units in Newark have the same number of bus stops which is unrealistic.

The different numbers of crime in the spatial units can be due to other factors such as bus stops or businesses which are identified as predictors. Consequently, the next step is to allow differences in μ based on observed characteristics of the areal units in the estimation process (Long and Freese, 2006). This has led to developments of count models without assuming equidispersion (Hilbe, 2007).

In practice, the Poisson distribution often does not fit due to overdispersion and under-prediction of zero occurrences. When the variance is larger than the mean, this distribution is described as overdispersion (Hilbe, 2007). When there is an overdispersion, it may lead to a misleading conclusion by deflating the p values. This may make it appear as though the predictor is statistically significant when it is not.

When there is an overdispersion, an important question is whehter the overdispersion is statistically significant to require a different count response model. There are two ways to answer this question (Hilbe, 2007). Graphing the distribution is one way to discern whether the data are over-dispersed. Another method of detecting overdispersion is to use Pearson statistics divided by the degrees of freedom. If this value (the dispersion statistic) is greater than 1.0, then the distribution is over-dispersed. If the dispersion statistics is greater than 1.25 for a moderate sized model, then the overdispersion needs to be addressed. For models with large numbers, the dispersion statistic of 1.05 may signal overdispersion (Hilbe, 2007).

The Negative Binominal Regression Model (NBRM) is a standard method used on over-dispersed count data (Hilbe, 2007). The NBRM addresses the underperformance of the PRM by adding a parameter (alpha) which reflects heterogeneity in the sample members. If the test of the Likelihood Ratio test of Alpha = zero (the Poission model) is statistically significant, this indicates that there is statistically significant difference between the PRM and NBRM estimations.

When there is an under-prediction of zeros in NBRM (i.e. there is an excess of zeros than predicted by negative binominal distribution), Zero Inflated Negative Binominal Regression Model (ZINBRM) may be required (Hilbe, 2007). In ZINBRM, the data are assumed to come from two separable distributions; one group of zeros is caused by structure that comes from a binary distribution. The structural zeros are deemed to be caused by latent group membership, and usually logistic or probit regression is used. This process is referred as the binary process, and it is important to determine which predictors lead to inflation of zero counts. Another group that may have zero counts come from a count distribution. For the count outcomes (not always zero group), NBRM is used to model the distribution.

3.1. Assessing the Model Fits and Improvements

As is the case with OLS regression and spatial model estimations, there are two ways to assess the model fits: within and between model fit comparisons. To assess the model fit improvements, AIC and Bayesian Information Criterion (BIC) tests will used (Long and Freese, 2006). The AIC statistics “is based on the log-likelihood function”, and BIC “is usually based on the deviance value” (Hilbe, 2007; p27). Again, the better-fitted models will have lower AIC and BIC statistic values when parameter estimates have equal significance.

In selecting the final model, scalar measures of model fit can aid the decision making process (Long and Freese, 2006). When using the scalar measures of the fit, Raftery (1996)'s guideline which is summarized in Table 13 in Long and Freese (2006) is useful. However, it is important to remember that the model with the best scalar measure of the fit is not necessarily the optimal or best model. The final model needs to be chosen in conjunction with scalar measures as well as within the context of the theoretical frameworks of the study (Long and Freese, 2006).

Table 13. Raftery (1996)'s guidelines in using BIC as scalar measure of model fit

Absolute difference	Evidence of support
0 – 2	Weak
2 – 6	Positive
6 – 10	Strong
> 10	Very strong

CHAPTER 6. DATA PREPARATION

1. Geocoding Results

The following section presents the geocoding results of two datasets: Newark City Police crime data, and the YellowBook data. Other datasets are obtained in GIS shapefiles, having no need for geocoding.

The geocoding was performed using Newark street file obtained from the city's survey team. There are two advantages in using Newark street file obtained from the city rather than using the census tiger files. First, the NJ Transit Corporation GIS team uses Newark city street file in constructing mass transit files. This made the NJ Transit bus stops neatly lined up along the Newark city streets. This was not the case with the census tiger file as the bus stops appeared to be located between the streets rather than along the streets. Second, Newark city street file is quite up to-date, and the local street files are generally considered to be more accurate than other files distributed enemas.

One of the reasons for maintaining the Newark city street file is for emergency purposes. For this reason, the highways inaccessible from the Newark local roads are not contained in the file. This presents a problem for geocoding the cases with the highway addresses. With the current form of Newark street files, there are two choices in addressing cases with highway addresses. First option is to force these cases to have approximate X and Y coordinates. Second option is to exclude these areas from further analyses. The first option is suitable when the homogeneity within the city can be assumed. If the areas without good address systems are deemed to be considerably different from the rest of the city, these areas can be excluded as discussed in Chapter 4, Section 2.4.

1.1. Newark City Police Crime Data

The Newark City Police crime data were geocoded in ArcMap 9.3[5] using Newark city street file. Overall geocoding success rate was over 98 percent using an option of minimum matching score of 80. The cases with tied addresses were geocoded after their locations were verified using Google Maps. In many cases, a tie occurred due to typing errors in the street shapefile. While the street shapefile is maintained and updated regularly, this file is not used for geocoding purpose in general. Examination of unmatched addresses revealed that some addresses were simply incorrect or non-existent. However, many of the unmatched cases with valid addresses appeared to have highway addresses.

Table 14 shows geocoding results by crime categories. The table shows that motor vehicle theft and theft from motor vehicle have the lowest geocoding match rates, as low as 96 percent for theft from motor vehicle. This may highlight the fact that motor vehicles are not as restricted by street networks or bus routes.

Table 14. Geocoding result of 2007 Newark City Police crime data

Crime type	Matched (%)	Unmatched	Total
Robbery	1,219(98.4)	20	1,239
Aggravated assault	1,066(98.7)	14	1,080
Motor vehicle theft	4,260(97.5)	107	4,367
Theft from motor vehicle	2,506(96.2)	99	2,605
Residential burglary	917(98.8)	11	928
Non-residential burglary	299(98.7)	4	303
Total	10,267(97.6)	255	10,522

[5] The cases with tied addresses were geocoded after their locations were verified using Google Maps. In many cases, a tie occurred due to typing errors in the street shapefile. While the street shapefile is maintained and updated regularly, this file is not used for geocoding purpose in general.

1.2. The Verizon's YellowBook Data

Among the 10,031 geocodable YellowBook cases, almost 98 percent were geocoded. The geocoding result of the YellowBook database is shown in Table 15. As was the case for the crime data, the unmatched cases were examined to find any patterns or systematic bias in geocoding. Many of the unmatched records had highway addresses located in or nearby the Newark Liberty International Airport or Port Newark. These business categories were often hotel, construction or transportation, and heavy equipment companies.

In other cases, the businesses recorded to be located in Newark were actually located outside of Newark albeit in close proximity. It appeared that this was based on a rational decision; that the business owners may think that listing their business as located in Newark would increase their chances of being searched rather than listing the actual locations. For instance, if a restaurant is located about 500 feet from Newark boundary, it makes sense to list their business as being located in Newark in a hope to attract more customers.

Table 15. Geocoding result of the YellowBook data by major SIC group (n=10,031)

SIC subdivision and code	Matched (%)	Unmatched	Total
sic54xx: food stores	324 (99.1)	3	327
sic55xx: auto dealers/gas stations	179 (96.8)	6	185
sic58xx: eating/drinking places	560 (98.8)	7	567
finance: banks	93 (98.9)	1	94
sic72xx: personal services	558 (99.8)	1	559
sic73xx: business services	328 (97.9)	7	335
sic75xx: automotive repair, services/parking	400 (96.9)	13	413
Other types of business listings	7,373 (97.6)	117	6,437
Total	9,815 (97.9)	216	10,031

2. Redefining the Study Area

The examination of non-geocoded data showed that there was a systematic bias. The cases with highway addresses were always not geocoded since the street file did not contain highways. Also the businesses located along the highway seemed to be different from those located in the rest of the city. They were often hotels catering to tourists or those involved in transportation businesses. For the reasons mentioned in Chapter 4, it was decided to exclude the heavy industrial land use designation areas where the airport and Port Newark are located. A total of 148 Thiessen polygons out of 2,750 were identified to be completely within the heavy industrial land use element without residents. One notable exception was the area where a state prison is located nearby the airport[6]. This captive population is isolated from the general public, and their movements are not influenced by businesses or public transportation in the areas. Figure 9 shows the included and excluded areas from further analyses.

The use of street intersection as the center to create the Thiessen polygon led to creation of small areal units when streets were dense and large areal units when streets were sparse. For this reason, the 148 Thiessen polygons in non-residential heavy industrial land use areas occupied about 9.2 square miles and the rest of the Thiessen polygons (n=2,602) occupied about 17.3 square miles.[7] Table 16 shows the comparison of geographical characteristics of included and excluded areas. The big Thiessen polygons in included areas were mostly located nearby excluded areas or included open spaces such as park or cemetery within the city. Therefore, these relatively big Thiessen polygons do not necessarily contain long street segments.

---

[6] In Year 2000, it was reported that there were 2,652 male and 7 female residents.
[7] The Thiessen polygon layer includes open spaces such as parks, rivers, and cemeteries.

Table 16. Comparison of included and excluded areal unit characteristics (n=2,750)

Number of areal units	Included(n=2,602)	Excluded(n=148)
Minimum size in foot	15,552.3	30,234.0
Maximum size in foot	3,414,950.8	12,487,856.0
Average size in foot	185,359.2	1,740,715.1
Standard deviation in foot	180,106.8	2,483,394.4
Total area size in mile	17.3	9.2

To perform statistical analysis for meaningful comparisons, it is necessary to aggregate point data to an areal unit. First step was to aggregate point data to the Thiessen polygon areal unit. In this step, point data located on the Thiessen polygon boundaries presented a challenge in the process. Use of spatial join in ArcMap led to double counting of the point data on the polygon boundaries while the use of Hawths Tools led to exclusion of them. Simply assigning these point data to one of the Thiessen polygons did not seem to be appropriate since some of these point data were repeat addresses. The number of point data within Thiessen polygon was therefore created by dividing the sum of the numbers by ArcMap's spatial join function and by Hawths Tools by two. As a result, it is possible that some Thiessen polygons have .5 criminal incidents or commercial establishments. None of bus stops were located on the Thiessen polygon boundaries.

The city parcel data were obtained to gather the locations of schools, public housings, open spaces, vacant lands or buildings, and mixed land use areas. Some of these parcels may take up the whole street block or even bigger areas. Also some facilities such as cemeteries or educational institutions consisted of more than one parcels. Using city parcel polygon data, binominal variables were created to indicate the presence of chosen land use types and facilities. According to the city, there are total five higher educational institutions and 106 academies and charter schools in Newark. About 2 percent and 10 percent of areal units contained higher educational institutions and grade schools respectively. None of the educational institutions were located in the excluded areas. For the 81 public housing buildings, none of them was located in the excluded area. About 11 percent of the included areal units contained public housing. Table 17 shows number of Thiessen polygons in both included and excluded areas with selected facility types. The variable label is presented following the variable name shown in lower cases. With the exception of cemetery and vacant land, none of the selected facilities were located in the excluded area.

Figure 9. Refined study area: Thiessen polygons to be excluded

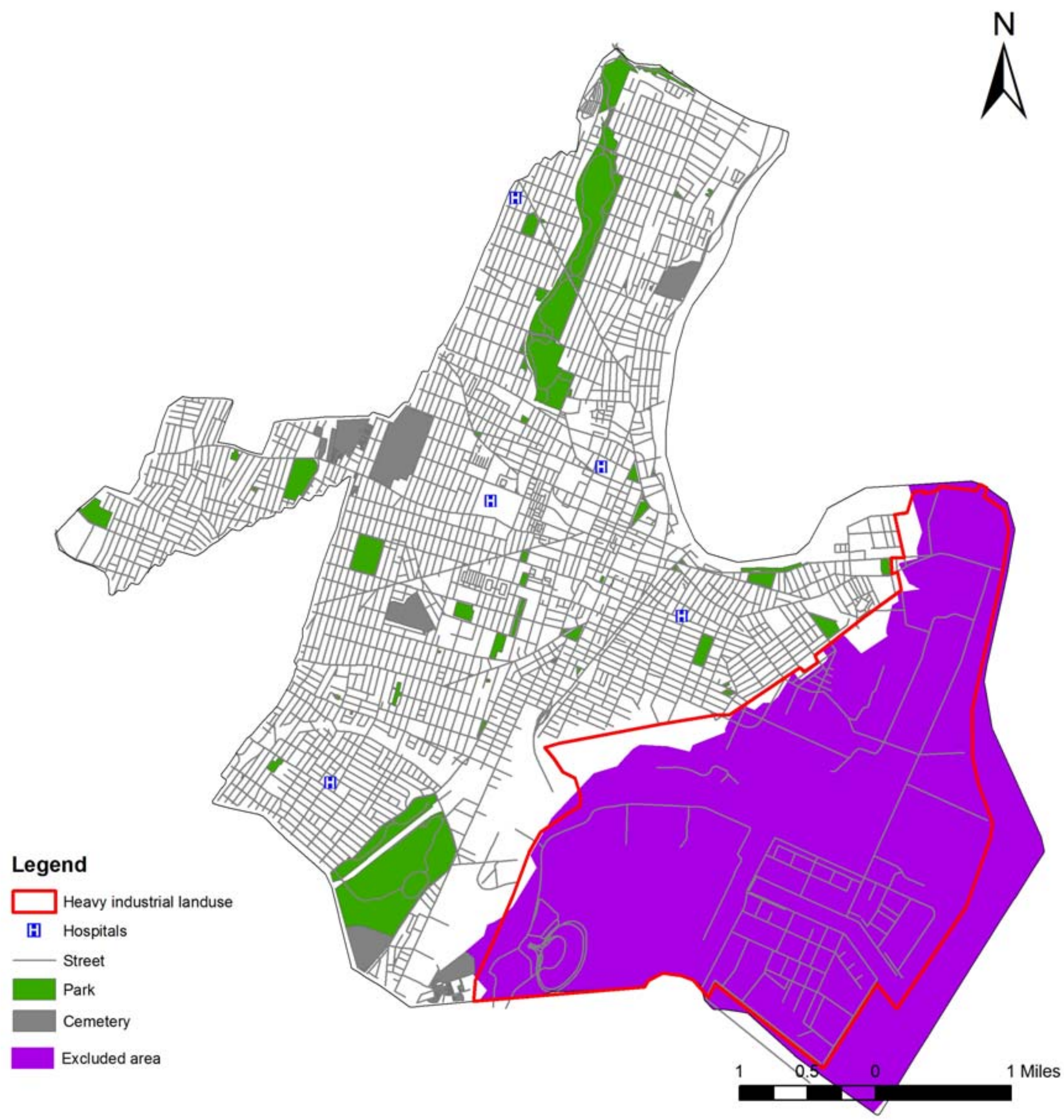


Table 17. Comparison of land use data by included (n=2,602) and excluded (n=148) areal units

Facility or land use type	Number of included area (%)	Number of excluded area (%)
univcoll: higher education(n=5)	55 (2.1)	0
schools:: grade K-12 (n=106)	259 (10.0)	0
phousing: public housing (n=81)	270 (10.8)	0
openspace: park and cemetery	395 (14.4)	1
mixeduse: mixed land use	291 (10.6)	0
pc1fvland: vacant land/ building	1,656 (63.6)	74

According to the city, there are total five higher educational institutions[8] and 106 academies and charter schools in Newark. About 2 percent and 10 percent of areal units contained higher educational institutions and grade schools respectively. None of the educational institutions were located in the excluded areas. For the 81 public housing buildings, none of them was located in the excluded area. About 11 percent of the included areal units contained public housing.

There are two types of open spaces in Newark in addition to vacant lands; parks and cemeteries. There are 65 city parks which tend to be small in their size and 14 county parks which tend to be larger. These parks were located in 333 Thiessen polygons (11.7 percent), and none of them were located in the excluded areas. For cemetery, obtaining the exact number of cemeteries was not straightforward since one cemetery may consist of several parcels. Only 66 Thiessen polygons (2.5 percent) contained cemeteries in their unit. By combining parks and cemeteries, a new variable *openspace* was created. About 14 percent (n=395) of areal unit contained park or cemetery.

There were many land parcels classified as vacant in year of 2007 in both included and excluded areas. In fact, a total of 1,656 Thiessen polygons had some sort

[8] The five higher educational institutions are Rutgers University-the State University of New Jersey, New Jersey Institute of Technology (NJIT), University of Medicine and Dentistry of New Jersey (UMDNJ), Seton Hall School of Law, and Essex County College. Most of them are located in University Heights area in Newark.

of vacant lands or buildings in their units. The mixed land use was defined by having all three types of land use designation within the areal unit. This was due to the fact that the majority of the areal units had more than one type of land use designations. Since there was no residential land use designation in the excluded area, the designation of mixed land use fell only within the included area. About 10 percent of the areal unit (n=291) had all three types of land use designation within their areal units.

The next step was to examine how many bus stops, criminal incidents, and commercial establishments were located in the excluded areas. Table 18 shows the crime data separated by excluded and included areas. Despite the fact that the excluded area is about a third of Newark, the majority of crimes (98.7 percent) occurred within the included area. All of the residential burglary occurred within the included study area. This makes sense since the excluded areas are with heavy industrial land use designation. The highest crime figure in the excluded area was theft from motor vehicle where almost three percent of them occurred in the excluded study area. It appeared that non-geocoded motor vehicle related offences occurred in the parking lots in the hotels around the airport. The comparison of the included and excluded areas confirmed the suspicion that the city as a whole is not homogeneous. Therefore, the decision to exclude the portion of Newark does not seem to comprise the research design.

Table 18. Comparison of crime types by included (n=2,602) and excluded (n=148) areal units

Crime type	Number of included area (%)	Number of excluded area (%)	Total(%)
Robbery	1,217(99.8)	2(0.2)	1,219
Aggravated assault	1,063(99.7)	3(0.3)	1,066
Motor vehicle theft	4,227(99.2)	33(0.8)	4,260
Theft from motor vehicle	2,437(97.2)	69(2.8)	2,506
Burglary	1,207(99.3)	9(0.7)	1,216

For commercial establishments, 118 (4.8 percent) geocoded cases belonged to the excluded areas. The high percentage was primarily due to the fact that many of these establishments were located in the Newark Liberty International Airport. This may accentuate the importance of excluding this area from further analysis; while the commercial establishments at the airport were successfully geocoded, the crimes at the airport were not included since these were outside of Newark City Police Department's jurisdiction. Table 19 contains the summary of the YellowBook data by included and excluded areas.

Table 19. Comparison of commercial establishments by included (n=2,602) and excluded (n=148)

SIC major group	Number of included area (%)	Number of excluded area (%)	Total (%)
sic54xx: food stores	311 (96.0)	13	324
sic55xx: auto dealers/gas stations	170 (95.0)	9	179
sic58xx: eating/drinking places	534 (95.4)	26	560
finance: banks	89 (95.7)	4	93
sic72xx: personal services	557 (99.8)	1	558
sic73xx: business services	301 (91.8)	27	328
sic75xx: automotive repair, services, parking	362 (90.5)	38	400

As with other point data, the majority of bus stops (93.6 percent) were located within the included study area. None of the subway stations and rapid rail stations was located in the excluded areas. Table 20 shows the mass transit in Newark separated by included and excluded areas.

Table 20. Comparison of mass transit by included and excluded areal units

Mass transit	Number of included area (%)	Number of excluded area (%)	Total(%)
bstops: bus stop	1,982 (93.6)	136	2,118
Subway station	15(100.0)	0	15
Rapid rail station	3(100.0)	0	3

CHAPTER 7. EXPLORATORY DATA ANALYSIS RESULTS

1. Non-spatial Descriptive Statistics

Table 21 shows non-spatial descriptive statistics by areal units. From the examination of the table, there are two important things to notice. First, a big proportion of areal units were without any criminal incidents for year 2007. The most numerous kind of offences among the five offences was motor vehicle theft, about twice more common than theft from motor vehicle. Even for this high frequency crime, about 34 percent of areal units had no crime for 2007. Therefore, it is clear that the data were not normally distributed. Second, for all five types of crime, the Standard Deviations (SD) were always bigger than the mean, signaling a possible overdispersion.

Table 21. Descriptive statistics of criminal offences by areal unit (n=2,602)

Criminal offence	Total number	Number of areal unit with zero count(%)	Mean	SD	Maximum count
Robbery	1,217	1,785 (68.6)	.47	.89	10
Aggravated assault	1,063	1,919 (73.8)	.41	.89	10
Motor vehicle theft	4,227	886 (34.1)	1.63	1.91	
Theft from motor vehicle	2,437	1,370 (52.7)	.94	1.56	
Burglary	1,207	1,825 (70.1)	.46	.90	9
Residential burglary	917	1,979 (76.1)	.35	.76	7

Table 22 shows descriptive statistics on the bus stops and commercial establishments. The spatial concentrations in these variables were even more pronounced than the dependent variables. One of the highest standard deviation was observed in *bstops*. This could be due to the actual concentration of bus stops or the way the bus stops was counted. For the present study, the bus stops were counted based on how many bus lines stop at the particular locations. For this reason, if one bus stop was shared by three different bus lines, then this bus stop was counted three times. This may have inflated the SD of *bstops*.

Table 22. Descriptive statistics of predictors by areal unit (n=2,602)

Bus stop and commercial establishment variable	Total number	Number of units with zero(%)	Mean	SD	Maximum count
bstops: bus stop	2,118	2,115 (81.3)	.762	2.58	47
sic58xx: eating/ drinking place	560	2,238 (86.0)	.205	.61	9
sic55xx: auto dealer/gas station	179	2,465 (94.7)	.065	.32	7
sic54xx: food store	324	2,342 (90.0)	.120	.39	4
sic75xx: automotive repair, services, and parking	400	2,339 (89.9)	.139	.48	6
sic73xx: business services	328	2,394 (92.0)	.116	.58	13
sic72xx: personal services	558	2,251 (86.5)	.214	.65	6
finance: banks	93	2,521 (96.9)	.034	.21	4

In this study, in addition to bus stops and commercial establishments, there is another set of predictors: land use information. Table 23 shows the descriptive statistics on land use information. The highest spatial concentration was observed in *univcoll.* Table 24 shows descriptive statistics on spatial lag calculated using 1,600 feet distance method. These variables will be used in count response estimation methods. For the spatial lags, very few areal units had the value of zero. Also the SDs of the spatial lags were smaller than the mean for all five crime types.

Table 23. Descriptive statistics of land use information by areal unit (n=2,602)

Land use information variable	Number of areal units with zero(%)
mixeduse : mixed land use (y/n)	2,311 (88.8)
pc1fvland: vacant land/building (y/n)	946 (36.4)
phousing : public housing (y/n)	2,332 (89.6)
univcoll : higher education (y/n)	2,547 (97.9)
schools : grade (k-12) (y/n)	2,343 (90.1)
openspace: park and cemetery (y/n)	2,207 (84.8)

Table 24. Descriptive statistics of criminal offence spatial lags at 1,600 feet by areal unit (n=2,602)

Spatial lag by crime type	Number of areal units with zero(%)	Mean	SD	Maximum value
Robbery	13 (.50)	.47	.22	1.56
Aggravated assault	18 (.69)	.41	.26	1.47
Motor vehicle theft	7 (.27)	1.64	.62	8.80
Theft from motor vehicle	7 (.27)	.95	.47	3.85
Burglary	13 (.50)	.47	.32	2.23

2. Exploratory Spatial Data Analysis

2.1. Mean Centers and Standard Directional Ellipse

The shape of the study area resembles somewhat of a crooked cross. Based on the shape of the study areas, it is not surprising to see the mean center located around the center of Newark. Figure 10 shows the mean centers of crime data. Unexpectedly, the mean centers of robbery and motor vehicle theft were very close from each other. However, the information on the mean centers did not reveal too much information on spatial crime patterns in Newark.

The standard directional ellipse corresponds to standard deviation in classical statistics except it provides information on directional bias. Therefore, comparing the standard directional ellipses of different crime types can inform the reader about the relative dispersion of the crime types (Wong and Lee, 2005). Figure 11 shows the standard directional ellipses of criminal offenses.

The standard directional ellipses of all five crime types showed directional bias to north and south, displaying elongated shape. This is not surprising based on the shape of the study area. Also the east side of Newark is surrounded by body of water, therefore, the spatial movements in the east side is mostly limited by several bridges. As with the mean centers, not too much information on spatial patterns was revealed from standard directional ellipses.

Figure 10. Mean centers of criminal offences in Newark

Figure 11. The standard directional ellipses of criminal offenses Newark, NJ

2.2. Cluster Analyses: Fuzzy Mode and NNH Analysis

For each crime type, top 10 locations were identified using the fuzzy mode function with a search radius of 300 feet. In cases where a tie existed (i.e. there are three fuzzy modes all ranked as the 10th fuzzy mode), all of them were included in the calculation. In addition, the NNH clusters were identified by using the distance radius of 660 feet. Use of 300 feet radius for NNH routine frequently produced no clusters. This can be partially explained by how the fuzzy mode is calculated: the fuzzy mode counts the point data more than once whereas NNH routine does not. Also the NNH clusters have the minimum number of points set at the higher threshold than the fuzzy mode. Table 25 summarizes the numbers of fuzzy modes and NNH clusters by crime type.

For robbery, there were total 14 fuzzy modes with minimum number of robberies of eight for 2007. When the NNH routine was performed using 300 feet radius, only one cluster was identified. Consequently, the radius for NNH was expanded to the 660 feet, one eighth of a mile. Using one eighth of a mile, a total of 25 NNH clusters were identified. All of the fuzzy modes were contained within the NNH clusters. This makes sense since the radius used for the fuzzy mode is less than a half of the radius used for NNH cluster analysis. Figure 12 shows the fuzzy modes (300 feet) and NNH clusters (660 feet) of robbery.

Table 25. Number of fuzzy modes and clusters of criminal offences

Crime type	Number of top 10 fuzzy mode(300ft) /minimum points	Number of clusters(300ft) /minimum points	Number of clusters(660ft) /minimum points
Robbery (n=1,217)	14, 8	1 (12)	25(23)
Aggravated assault (n=1,063)	14,10	3 (13)	22(20)
Motor vehicle theft (n=4,227)	148,11	37 (18)	155(46)
Theft from motor vehicle (n=2,437)	91,10	15 (31)	69(48)
Burglary (n= 1,207)	11, 9	0	21(23)

For aggravated assault, there were total 14 fuzzy modes with minimum number of assault of 10 for year 2007 using 300 feet radius. Using 660 feet radius, there were total 22 clusters identified with minimum number of 20 incidents. Figure 13 shows the fuzzy modes (300 feet) and NNH clusters (660 feet) of aggravated assault. It seems that the clusters of robbery and assaults are located in different areas.

For motor vehicle theft and theft from motor vehicle, it is important to remember that their total numbers are more numerous than the other three types of crimes. For this reason, the reader should be mindful of the higher numbers of fuzzy modes and clusters identified. For theft from motor vehicle, a total of 91 fuzzy modes with minimum number of 10 incidents were identified using 300 feet. When using 300 feet for NNH routine, there were total 15 clusters with minimum points of 31. For 600 feet radius of NNH routine, there were 69 clusters identified with the minimum number of point 48. For motor vehicle theft, there were total 148 fuzzy modes with minimum of 11 points. Using 300 feet radius, there were total 37 NNH clusters identified with the minimum point of 18. When using 660 feet, these numbers changed to 155 clusters with the minimum points of 46. Figure 14 and Figure 15 show the fuzzy mode and NNH cluster of motor vehicle theft and theft from motor vehicle respectively using 300 feet radius.

For burglary, a total of 11 fuzzy modes with minimum number of nine burglaries were identified. When using the 300 feet radius, no NNH cluster was identified. When NNH routine was performed using 660 feet radius, there were a total 21 clusters with the minimum number of point 23. Many of the clusters seem to be located in the west side of Newark. This is probably due to the fact that the central business district is located in the east side, perhaps limiting the residences in the areas. Figure 16 shows the fuzzy modes (300 feet) and NNH clusters (660 feet) of burglary.

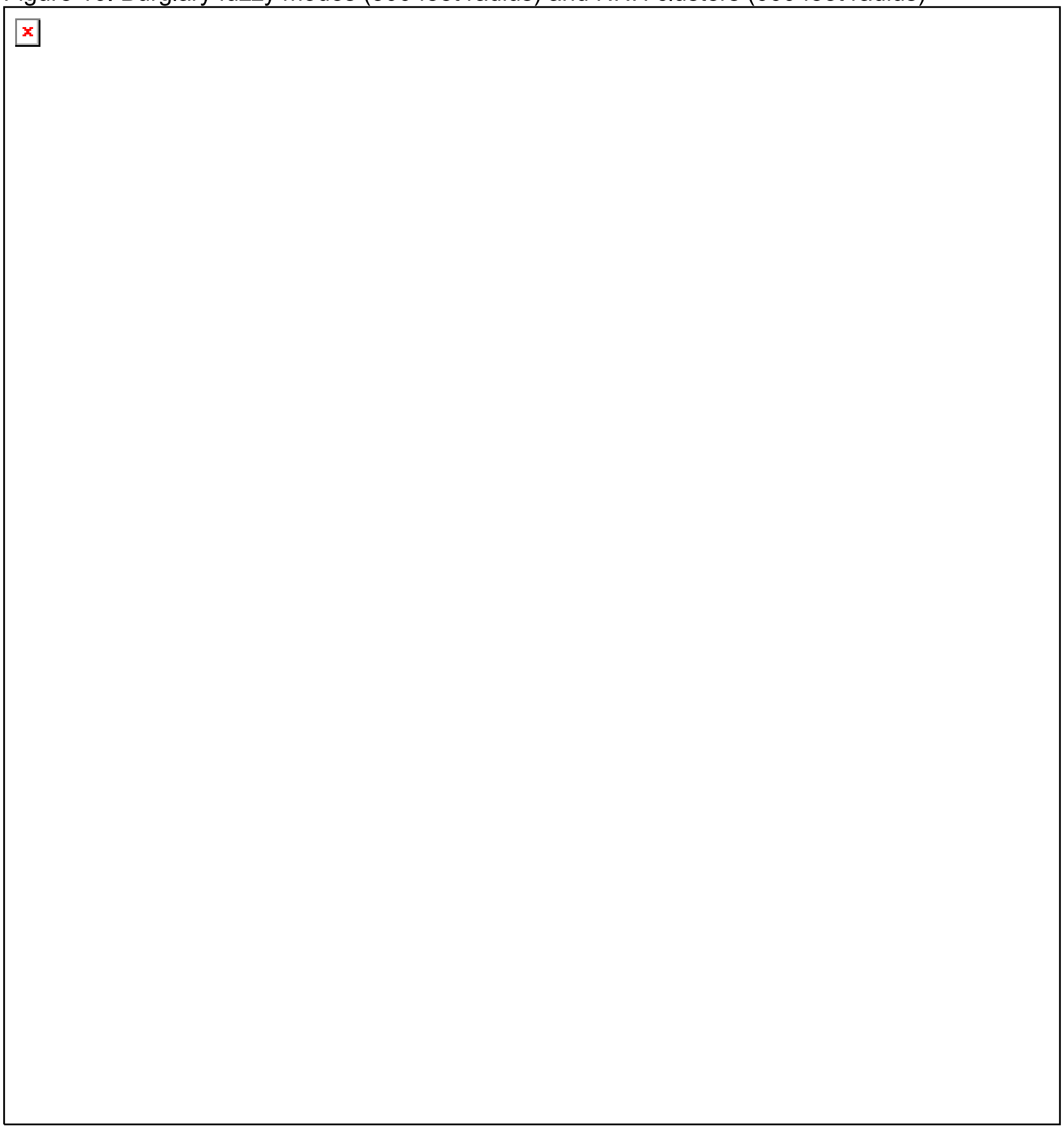
Figure 12. Robbery fuzzy modes (300 feet radius) and NNH clusters (660 feet radius)

Figure 13. Aggravated assault fuzzy modes (300 feet radius) and NNH clusters (660 feet radius)

Figure 14. Motor vehicle theft fuzzy modes and NNH clusters using 300 feet radius

Figure 15. Theft from motor vehicle fuzzy modes and NNH clusters of using 300 feet radius

Figure 16. Burglary fuzzy modes (300 feet radius) and NNH clusters (660 feet radius)



2.3. Average Nearest Neighbor Distance Analysis

Table 26 shows the average distances observed and expected for each crime type using ArcMap 9.3. The expected distances were calculated in Euclidean distance based on the size of the included study area. All five offence types were clustered at statistically significant levels. Generally, the higher the number of crimes, the shorter the expected distances. For burglary, residential burglary displayed tighter distribution than non-residential burglary by looking at the Z values. This seems to suggest that there is a utility in separating burglary into more detail for data analysis. The crime type with the highest Z value was motor vehicle theft which incidentally was the most numerous crimes.

Table 26. Average nearest neighbor distance analysis in feet of criminal offences

Crime type	Observed Distance	Expected Distance	Ratio	Z value
Robbery (n=1,217)	221.4	316.9	0.70	-20.0
Aggravated assault(n=1,063)	226.3	340.0	0.67	-20.7
Theft from motor vehicle (n=2,437)	145.9	224.8	0.64	-32.8
Motor vehicle theft (n=4,227)	104.6	170.7	0.61	-47.6
Burglary (n=1,207)	224.2	320.5	0.70	-19.7
Residential (n=917)	253.1	366.2	0.69	-17.7
Non-residential (n=290)	437.4	664.6	0.65	-10.8

* Two burglary incidents were not classified by premise type.

2.4. Moran's I Statistics

The minimum distance required to ensure every areal unit to have at least one neighbor when calculating spatial weight was found to be 1,555 feet in GeoDa. For this reason, the Moran's I test was performed for each crime type beginning with the 1,600 feet up to 3,300 feet. The Moran's I statistics results are summarized in Table 27.

There are two important things to notice from Table 27. First, all of the examined crime types display positive Moran's I values, signaling the clustering of the data. Second, it can be seen the value of Moran's I statistics decreases as the distance used to define the neighbors increases. The Moran's I statistics show that the distance of 1,600 feet yields to the highest Moran's I values (i.e., clustering) for all five crime types. For this reason, the spatial lag calculated using 1,600 feet will be used for regression estimation methods.

Table 27. Moran's I statistics using distance in feet of criminal offences

Distance in feet	Robbery	Assault	Motor vehicle theft	Theft from motor vehicle	Burglary
1,600	.045	.077	.096	.086	.120
1,800	.040	.070	.088	.078	.115
2,000	.037	.061	.078	.068	.109
2,640	.029	.050	.060	.054	.085
3,300	.026	.040	.047	.048	.075

2.5. LISA and Density Maps

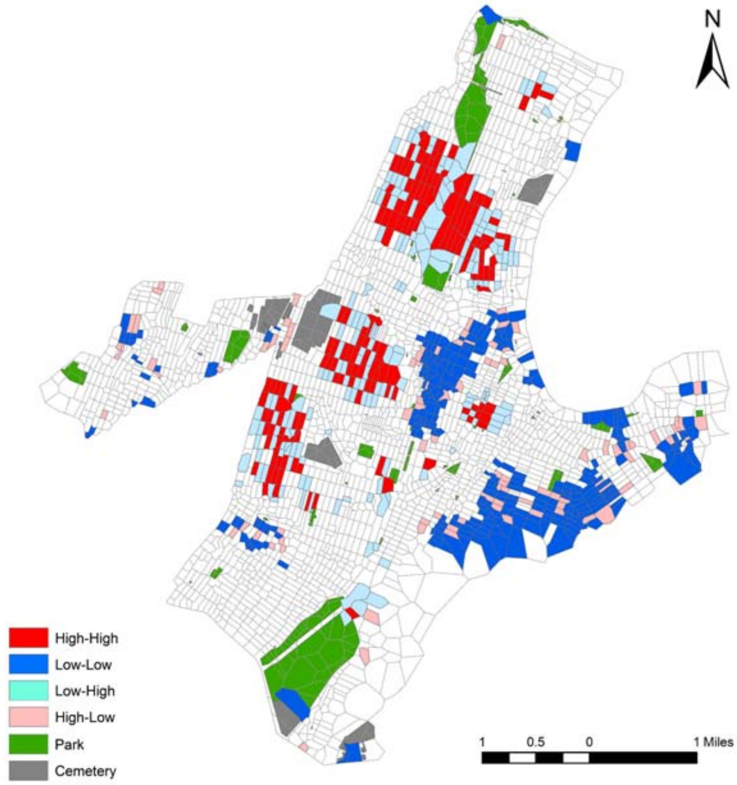
In this section, the spatial patterns of five crime types will be examined. The LISA maps are created by comparing whether the areal unit has the similar or dissimilar values with its neighbors. While the LISA uses areal units to identify one's neighbors, the Kennel density function shows crime concentrations without being restricted by areal units. For both LISA and Kennel density maps, the search radius used is 1,600 feet for all crime types.

The results presented in this section are largely visual and self-explanatory. For this reason, there is little explanation of the maps. Figures 17 and 18 show the LISA and Kennel density of robbery and aggravated assault respectively. These two violence crimes displayed somewhat different spatial patterns although their cores of crime clusters had the common areas: for both crimes, the lower portion of Branch Brook Park and the west side of Newark neighboring Irvington displays crime clusters. In addition, the areas where universities are located were largely immune from the violence offences.

Figure 19 shows LISA and Kennel density of burglary where large clusters were identified in areas adjacent to Irvington. The east side of Newark suffered from very low level of burglary. However, this finding needs to be taken with caution. The east side of Newark is downtown areas where few residences exist. Therefore, the concentration of burglary in the close proximity to Irvington area may simply reflect the high density of residences in the area.

Figures 20 and 21 show the LISA and Kennel density of offences related to motor vehicles. For motor vehicle theft (Figure 20), the Kennel density map shows a wide spread offences in the study area. Again, the University Height areas where the colleges are located suffered from relatively low level of motor vehicle theft. However, this was not the case for theft from motor vehicle.

Figure 17. Robbery spatial pattern using LISA and Kennel density



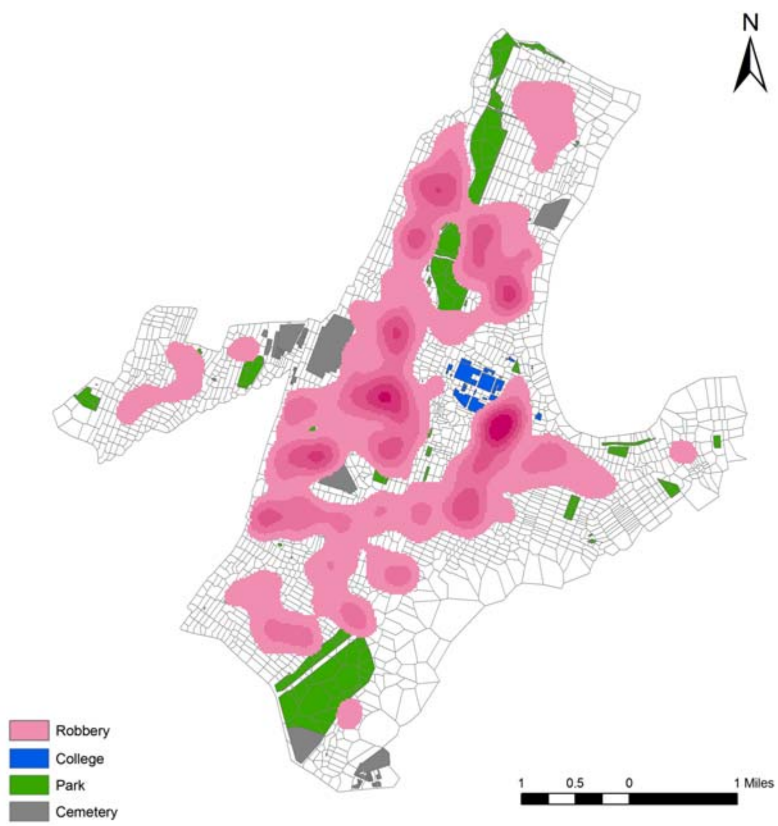
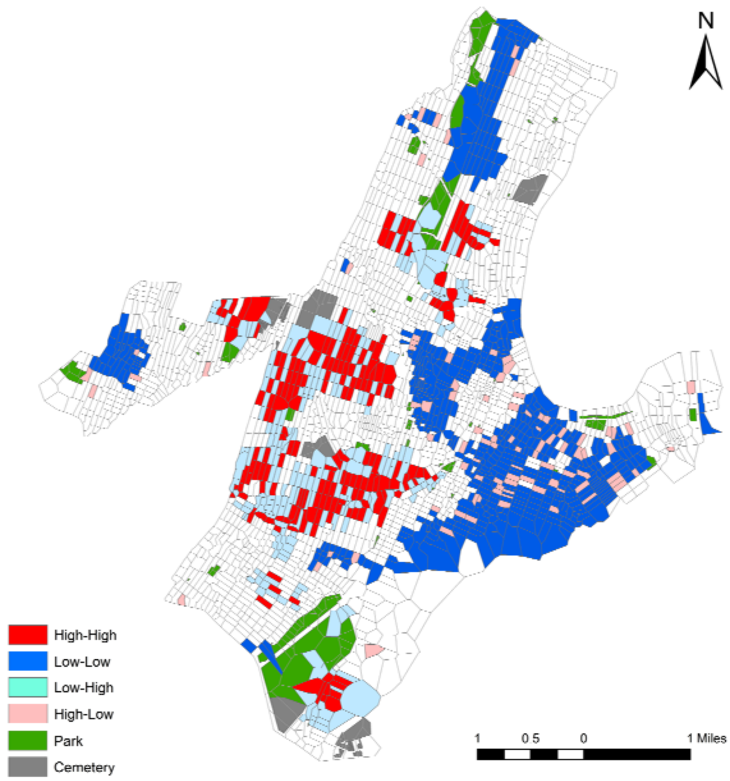


Figure 18. Aggravated assault spatial patterns using LISA and Kennel density



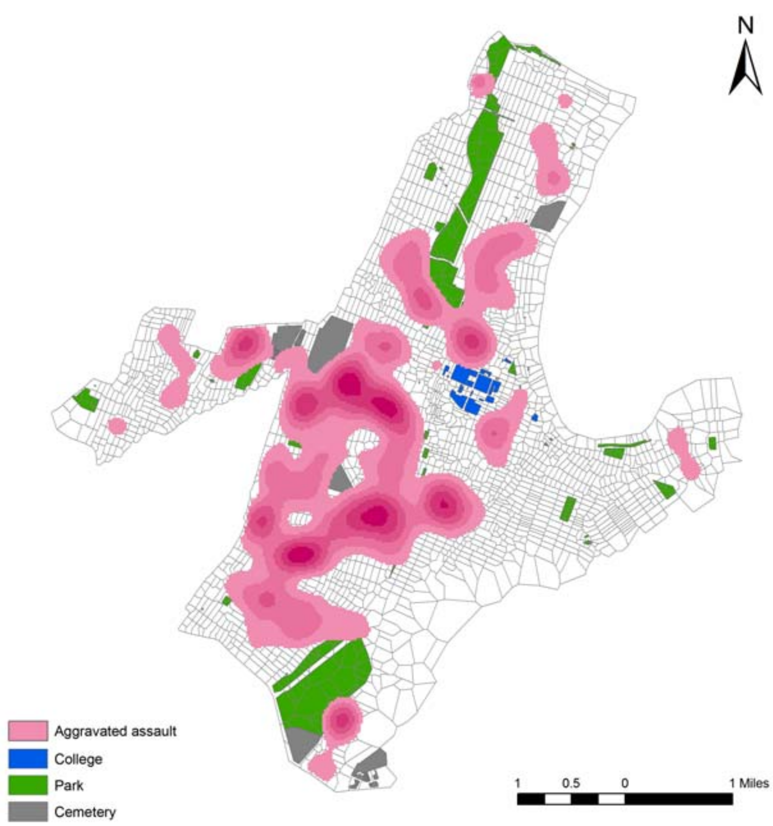
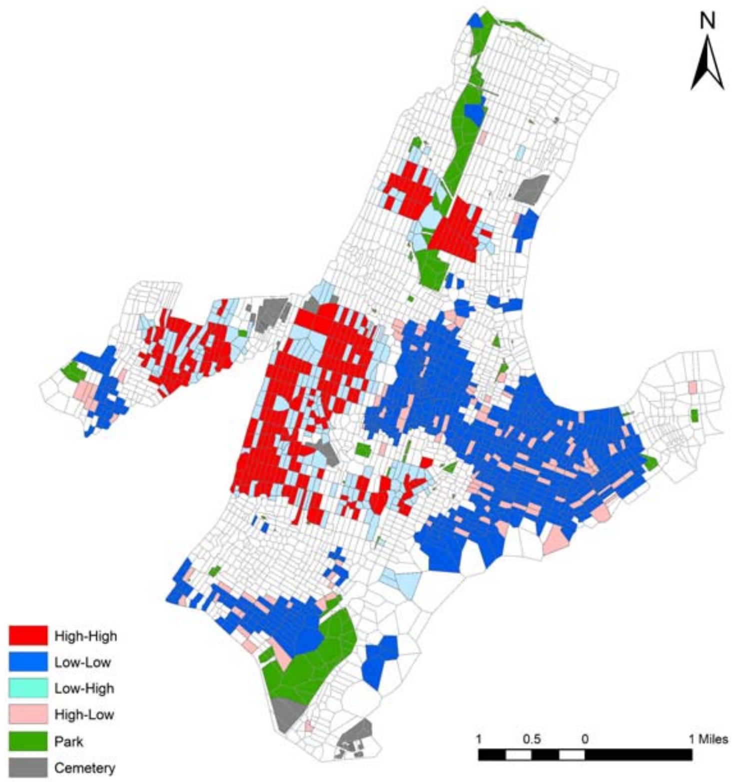


Figure 19. Burglary spatial patterns using LISA and Kennel density



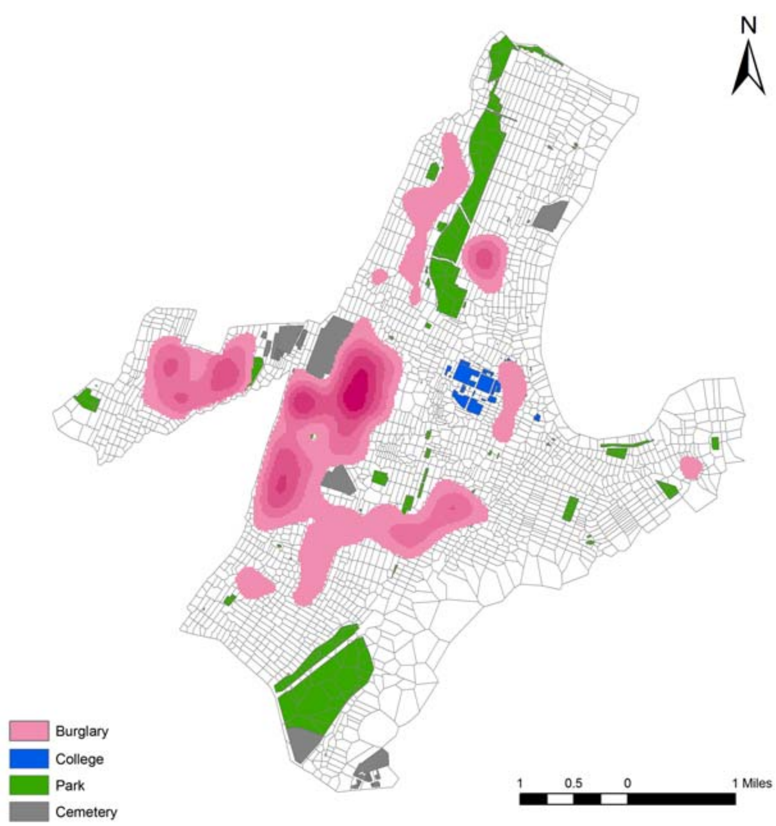
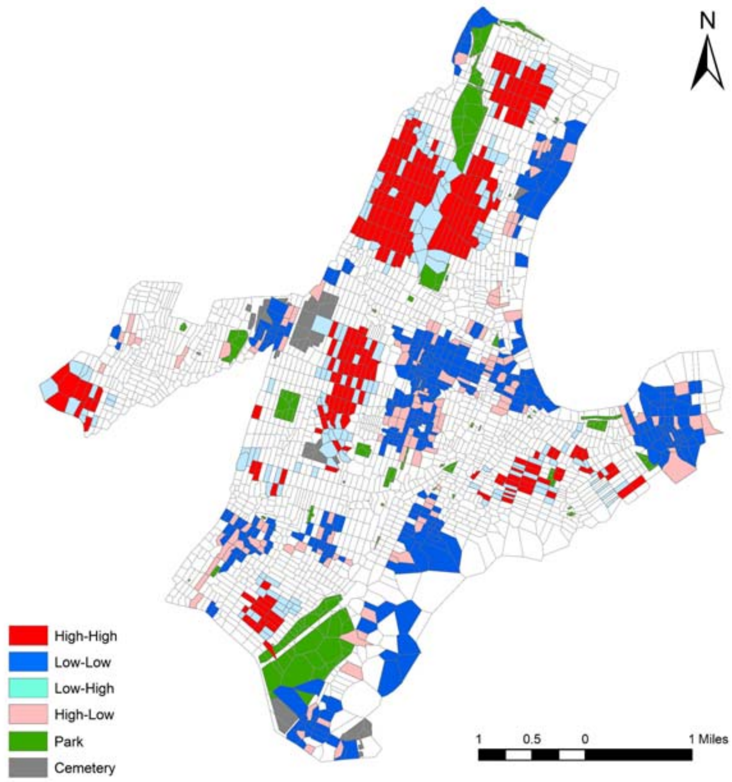


Figure 20. Motor vehicle theft patterns using LISA and Kennel density



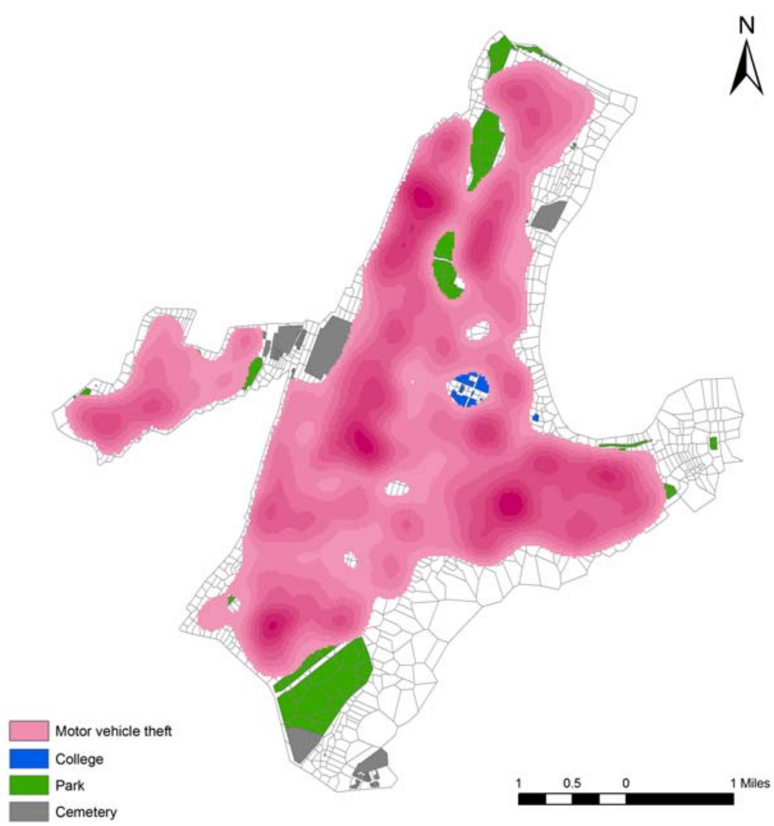
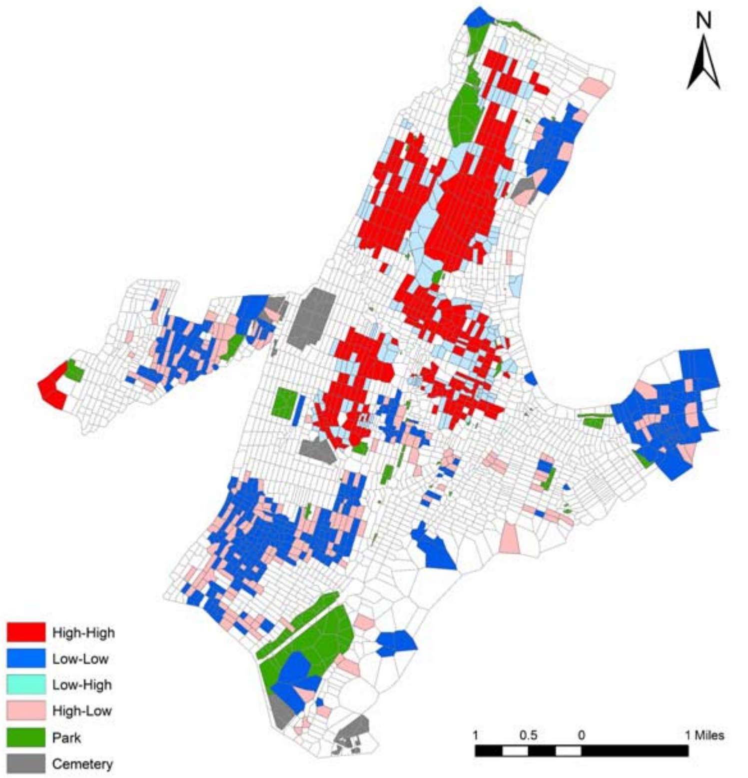
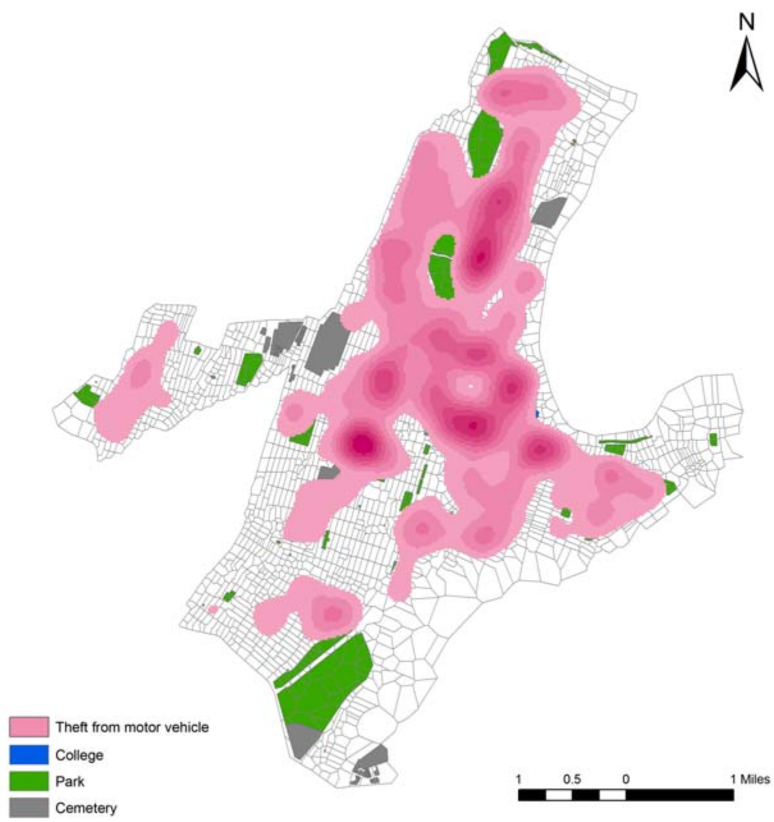


Figure 21. Theft from motor vehicle using LISA and Kennel density





CHAPTER 8. REGRESSION ANALYSIS RESULTS

In this chapter the results of inference statistics will be discussed by each crime type. It will begin with the OLS regression analysis with diagnostics, and then the spatial process models. To address non-normality and over-dispersion of the data, the analysis will include count response modeling methods in Stata.

1. Analysis Results on Robbery

In Chapter 3, it was hypothesized that both bus stops and commercial establishments will increase the risks of robbery in the neighborhoods. In this section, these hypotheses will be tested using the spatial process model and count response model estimations. The detailed regression results and model fit comparisons are available in Appendix A.

1.1. OLS Regression

Table 28 shows the OLS regression results on robbery using GeoDa for all models. The first research question of the present study was related to the relationship between number of bus stops and crime. Using the number of bus stops as the only predictor in Model 1, the number of bus stops was statistically significant to increased robbery. The second research question was concerned with the relationship between robbery and commercial establishment in the area. When examined the influences of business categories on robbery, all three types of retail business were associated with increased robbery. However, only *sic72xx* (personal services) was related to increased robbery among the service industries. The number of banks was not statistically significant. The R-squared of Model 2 was .105, quite a big improvement from Model 1

(R-squared = .028). When both types of predictors (bus stops and commercial establishments) were added to the model (Model 3), no interaction effects between predictors were observed.

Table 28. Robbery OLS regression with spatial lag at 1,600 feet (n=2,602)

	Model 1	Model 2	Model 3	Model 4
R-squared	.028	.105	.115	.128
Log Likelihood	-3339.40	-3231.90	-3216.64	-3197.77
AIC	6682.80	6479.80	6451.29	6425.53
SC	6694.53	6526.71	6504.06	6513.50
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	.424(.018)**	.333(.019)**	.316(.019)**	.242(.031)**
bstops	.057(.007)**		.036(.006)**	.038(.007)**
sic58xx		.243(.031)**	.228(.031)**	.223(.031)**
sic55xx		.142(.052)**	.132(.052)*	.135(.052)**
sic54xx		.276(.045)**	.266(.045)**	.253(.045)**
sic75xx		.055(.035)	.044(.035)	.039(.035)
sic73xx		-.043(.030)	-.049(.029)	-.039(.029)
sic72xx		.174(.029)**	.161(.029)**	.161(.028)**
finance		.068(.081)	.051(.081)	.048(.081)
mixeduse				.150(.052)**
pc1fvland				.054(.034)
phousing				.133(.053)*
univcoll				-.256(.115)*
school				.199(.055)**
openspace				-.031(.046)
Multicollinearity condition number	1.34	2.21	2.28	4.45
Normality of errors: Jarque-Bera	27246.91**	19258.43**	17087.91**	17273.87**
Heteroskedasticity: Breusch-Pagan test	454.47(1)**	971.33(7)**	1058.29(8)**	1171.39(14)**
Moran's I (Z value)	.043(9.87)**	.044(10.06)**	.043(9.97)**	.037(8.84)**
Diagnostic result	None (ns)	SEM	SEM	SEM

* = p<.05, ** = p<.01

Lastly, the land use variables were added to Model 4. Again, no changes in statistical significance on the existing variables were observed. However, the presence of mixed land use, public housing, and grade schools were shown to be related to increased robbery in the areas. The higher educational institutions produced the opposite impacts: they were related to decreased number of robbery. This was not surprising since universities and colleges usually have their own police forces patrolling their vicinities in high frequencies. In addition, these higher educational institutions are clustered together, forming a neighborhood of university, therefore a creation of neighborhood named “University Heights”.

In general, adding more variables led to improved model fits and increased R-squared values. The LL numbers were increased whereas AIC and SC values decreased signaling model improvements. For all models, multicollinearity was not a concern. However, the diagnostics suggested significant problems with normality of errors, heteroskedasticity and spatial dependence.

1.2. Spatial Process Models

The OLS regression diagnostics showed that SEM was suitable over SLM except for Model 1; neither SEM nor SLM was seen to be suitable for Model 1. As previously discussed, spatial weight at 1,600 feet was calculated to take account of spatial dependence. The SEM results are summarized in Table 29.

In Model 1, the number of bus stops was associated with increased robbery. In SEM Model 2, the three retail business and personal business category predictors were statistically significant to increased robbery. In addition, *sic75xx* (automotive related services) displayed statistically significant relationship with number of robbery, a change occurred after taking into account spatial autocorrelation. The statistical significance of *sic75xx* remained intact even after *bstosp* was added to the model. Therefore it seemed that there was no interaction between *bstops* and commercial establishments.

In Model 4, the same as OLS regression result, the presence of mixed land use, public housing, and grade schools were related to increased robbery. The impact of higher educational institutions on robbery disappeared when spatial dependence in the error term was accounted for. This observation seems to support what was observed in

LISA and density maps of robbery (Figure 17). However, other variables including existence of automotive services, business services, banks, vacant land, college, and open spaces were not statistically significant in predicting robbery. In Model 4, the impact of *sic75xx* (automotive related services) seemed to have disappeared. However, the p value of the variable is .072, which could be seen as statistically significant if a higher alpha level was used.

Table 29. Robbery SEM estimation with spatial lag at 1,600 feet (n=2,602)

	Model 1	Model 2	Model 3	Model 4
R-squared	.054	.129	.139	.147
Log Likelihood	-3312.53	-3204.67	-3189.92	-3175.57
AIC	6629.07	6425.35	6397.84	6381.14
Schwarz criterion	6640.79	6472.26	6450.61	6469.11
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	.420(.031)**	.326(.031)**	.309(.031)**	.259(.039)**
LAMBDA(Spatial lag)	.456(.060)**	.453(.060)**	.449(.061)**	.428(.062)**
bstops	.056(.007)**		.036(.007)**	.036(.007)**
sic58xx		.247(.031)**	.234(.030)**	.228(.030)**
sic55xx		.153(.052)**	.140(.052)**	.147(.052)**
sic54xx		.242(.045)**	.232(.044)**	.226(.044)**
sic75xx		.083(.036)*	.071(.035)*	.064(.035)
sic73xx		-.032(.029)	-.034(.029)	-.031(.029)
sic72xx		.178(.028)**	.166(.028)**	.164(.028)**
finance		.103(.080)	.084(.080)	.081(.079)
mixeduse				.157(.053)**
pc1fvland				.026(.035)
phousing				.152(.059)**
univcoll				-.155(.130)
school				.165(.054)**
openspace				-.062(.048)
Heteroskedasticity: Breusch-Pagan(df)	447.81(1)**	1012.59(7)**	1095.87(8)**	1190.40(14)**

* = p<.05, ** = p<.01
Note: An underline was placed when the significance level was changed compared to OLS regression result.

When compared the model fits, there were increases in LL numbers as more predictors were added to the models. For AIC and SC values, they decreased from Model 1 to Model 3. When compared the numbers from Model 3 and Model 4, the value of SC was increased. Therefore, it is not clear whether land use information help explain robbery in Newark in spatial process models. The diagnostic tests revealed that there were significant problems with heteroskedasticity, suggesting a possible misspecification in the model.

1.3. Assessing the Model Fits and Improvements

For spatial process model, it is incorrect to interpret R-squared as OLS regression results. The model fit improvements can be assessed by comparing values of LL, AIC, and SC. For the model fit, the bigger the LL number, the better fit the model is. However, it is the opposite for the AIC and SC as smaller values signal better model fits. The fit statistics of four models are presented in Table 30.

When compared the LL, AIC, and SC values of all four models, it showed that the spatial error model estimation increased the model fits for all models; the LL values increased for all models, whereas the values of AIC and SC decreased relative to OLS regression results. Therefore, it can be concluded that taking account of spatial dependence in the model improved the model fits. In fact, the spatial autoregressive coefficient (LAMBDA, weighted dependent variable) was one of the most significant predictors in predicting robbery.

Checking the orders of Wald test, LR test, and LM statistics on the spatial autoregressive error coefficient of all four models, it shows that the ordering of these three tests violates the expected order. This shows that despite the model improvements made in spatial process modeling, there may be model misspecification violating the asymptotic properties of maximum likelihood estimates. Overall, the diagnostics in the OLS and SEM showed that there were problems related with nonnormality of the data, spatial dependence, and heteroskedasticity. In addition, the study area may not fit the behavioral aspects of robbery, possibly producing the spatial dependence in the error term.

Table 30. Fit statistics on robbery in OLS regression and SEM estimations

	Model 1	Model 2	Model 3	Model 4
OLS regression				
Log Likelihood	-3339.40	-3231.90	-3216.64	-3197.77
Akaike info criterion	6682.80	6479.80	6451.29	6425.53
Schwarz criterion	6694.53	6526.71	6504.06	6513.50
Spatial model				
Log Likelihood	-3312.53	-3204.67	-3189.92	-3175.57
Akaike info criterion	6629.07	6425.35	6397.84	6381.14
Schwarz criterion	6640.79	6472.26	6450.61	6469.11
Wald test	57.68	56.24	54.83	47.68
Likelihood ratio	53.74	54.45	53.45	44.39
Lagrange multiplier	93.98	96.91	94.91	71.15

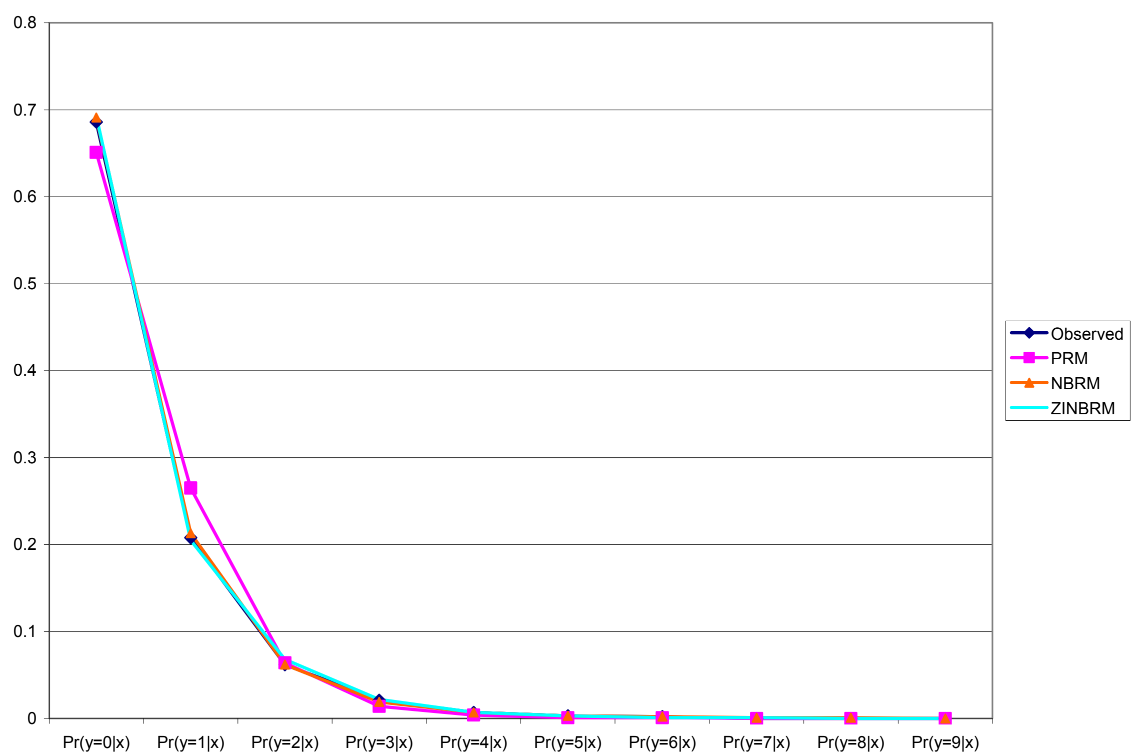
1.4. Negative Bi-nominal Regression

To address non-normal distribution of the data, the next step was to move to the count response regression models. Three types of count data regression methods (PRM, NBRM, and ZINBRM) were performed on all four models. Figure 22 shows the comparison of the robbery outcome distributions by PRM, NBRM and ZINBRM using Model 3. While it is less precise and difficult to draw a statistical conclusion from the graph, it is apparent that the PRM underperforms both NBRM and ZINBRM. The PRM under-predicts zeros and over-predicts count value of one, which is characteristic of count response model not allowing heterogeneity in the sample (Long and Freese, 2006).

The model fit comparisons in Stata showed that NBRM was favored over PRM for all models. However, there was conflicting evidence on whether to perform NBRM or ZINBRM. For BIC test, NBRM was preferred over ZINBRM. This is because BIC penalizes extra predictors more severely than AIC test does (Long and Freese, 2006). For this reason, BIC test showed a preference for the more parsimonious model, in this case, NBRM over ZINBRM. On the other hand, AIC and Vuong test showed preference for ZINBRM over NBRM for all models. The relevant tables are presented in Appendix A.

After reviewing the model fit comparisons, NBRM was chosen as the regression method for a couple of reasons. First, there is no good theoretical basis to expect the areal units without any robbery can be divided into two different latent groups. In this case, while the data may fit the statistical tests, it is possible to over-analyze the data without good theoretical reasons to do so. Second, a parsimonious model is preferred over a complicated one. Since there is some support for NBRM over ZINBRM, it was decided to use NBRM. Table 31 shows the results of NBRM on robbery with the spatial lag at 1,600 feet to account for spatial dependence in the data. The LL test of alpha = 0 shows whether there is a statistically significant difference between the Poisson and NBRM estimations. If this test is not statistically significant, that would mean that there is no statistical difference between PRM and NBRM estimation results. For all models, this value is statistically significant, signaling that there is statistically significant difference between the NBRM and PRM results.

Figure 22. Comparison of robbery distributions by PRM, NBRM, and ZINBRM estimations using Model 3



For Model 1 and Model 2, there was no change observed in the predictors' significance levels in NBRM compared to SEM estimation results. For Model 1, number of bus stops was related to increases in robbery. For Model 2, all three types of retail businesses, and only two out of three types of service industries (*sic75xx:* automotive related services and *sic72xx:* personal services) were associated with increases in robbery. For Model 3, the statistical significance of *sic75xx* (automotive related service) was .051. Therefore, it seems that there was no difference between the SEM result and NBRM result for Model 3. For Model 4, there was no change in the statistical significance level compared to SEM estimation result. The final model showed that the presence of mixed land use, public housing, and school (K-12) was related to increases in robbery. Again, the impact of higher educational institutions on robbery disappeared when spatial lag was added to the model. From this, it can be concluded that the SEM estimation and NBRM estimation led to similar results on robbery.

For all models, spatial lag was highly significant signaling the strong impacts of neighbors. As hypothesized, bus stops and three types of retail categories displayed positive correlations with robbery. Among the services category, only personal services (*sic72xx)* was statistically significant for all models, displaying positive correlation with robbery. Again, the number of banks did not seem to increase robbery occurrences. By adding more variable, there were improvements made with the model fits.

The next step was to find out which model among the four models was the best model. The selection of the final model involved a process of calculating and comparing the fit statistics in NBRM results. The fit statistics are presented at the very bottom of Table 31. More detailed results are available in Appendix A.

Table 31. Robbery NBRM with spatial lag at 1,600 feet (n=2,602)

	Model 1	Model 2	Model 3	Model 4
Pseudo R-squared	.029	.055	.058	.065
LL test of alpha=0	139.80(2)**	260.73(8)**	277.27(9)**	310.45(15)**
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	-1.550(.087)**	-1.664(.087)**	-1.677(.087)**	-1.768(.099)**
Spatial lag	1.356(.152)**	1.249(.148)**	1.223(.147)**	1.149(.149)**
bstops	.090(.013)**		.047(.011)**	.053(.012)**
sic58xx		.242(.050)**	.217(.050)**	.209(.049)**
sic55xx		.248(.094)**	.224(.092)*	.225(.091)*
sic54xx		.368(.077)**	.362(.076)**	.338(.075)**
sic75xx		.133(.064)*	.125(.064)	.112(.064)
sic73xx		-.024(.060)	-.041(.060)	-.017(.060)
sic72xx		.228(.047)**	.203(.047)**	.201(.047)**
finance		.116(.135)	.081(.134)	.082(.133)
mixeduse				.294(.098)**
pc1fvland				.069(.073)
phousing				.292(.105)**
univcoll				-.585(.323)
school				.318(.103)**
openspace				-.130(.099)
Log likelihood	-2381.46	-2257.99	-2249.73	-2233.14
AIC	4644.92	4535.99	4521.45	4500.27
BIC	4668.38	4594.63	4585.96	4599.96
Difference in BIC from previous model	n/a	73.75	8.67	-14.01
Scalar measure model fit compared to previous model		Very strong support for Model 2	Strong support for Model 3	Very strong support for Model 3

* = p<.05, ** = p<.01
Note 1: A variable coefficient is underlined when the significance level is different compared to SEM estimation.
Note 2: While *sic75xx* in Model 3 was shown to be not statistically significant, its p values was .51.

Based on the AIC and BIC numbers, there was a strong support for Model 2 over Model 1 in predicting robbery. In turn, this can be interpreted that number of commercial establishments better predicted robbery than number of bus stops. However, this interpretation needs to be taken with a caution. Model 1 have only two predictors (spatial lag and bus stops) while Model 2 has seven additional predictors. Nevertheless, Model 2 was preferred over Model 1 in predicting robbery using scalar measures.

When comparing Model 2 and Model 3 on robbery, this showed that the combination of bus stops and commercial establishments predicted robbery better than either bus stops or commercial establishments alone. Therefore, it can be concluded that both bus stops and commercial establishments contributed to robbery incidents. The comparison of fit statistics on Model 3 and 4 showed that there was very strong support for Model 3 over Model 4. In conclusion, adding the land use information to the model did not significantly improve the prediction on robbery.

1.5. Comparisons of Model 3 Estimations

The fit statistics in Stata showed that Model 3 may be the best model in explaining robbery incident. In this section, Model 3 results from different regression methods are compared. While it is not appropriate to compare the coefficient from different regression methods, it is suitable to examine the statistical significance levels of the predictors.

Table 32. Model 3 results comparison on robbery with spatial lag at 1,600 feet (n=2,602)

Regression model	OLS regression	SEM	NBRM
R-squared	.115	.139	.058
Log Likelihood	-3216.64	-3189.92	-2249.73
Predictors	Coefficient(SE)	Coefficient(SE)	Coefficient(SE)
Constant	.316(.019)**	.309(.031)**	-1.677(.087)**
LAMBDA (Spatial lag)		.449(.061)**	1.223(.147)**
bstops	.036(.006)**	.036(.007)**	.047(.011)**
sic58xx	.228(.031)**	.234(.030)**	.217(.050)**
sic55xx	.132(.052)*	.140(.052)**	.224(.092)*
sic54xx	.266(.045)**	.232(.044)**	.362(.076)**
sic75xx	.044(.035)	.071(.035)*	.125(.064)
sic73xx	-.049(.029)	-.034(.029)	-.041(.060)
sic72xx	.161(.029)**	.166(.028)**	.203(.047)**
finance	.051(.081)	.084(.080)	.081(.134)

* = p<.05, ** = p<.01

The variable *sic75xx* was not statistically significant in the OLS regression. However, when spatial dependence was accounted for, this variable became statistically significant. While *sic75xx* was shown to be not statistically significant in NBRM model, its probability was .051 whereas the p value of the same variable in SEM model was .045. Therefore, it can be seen that their significance levels was actually the same for SEM and NBRM models. In conclusion, the existence of bus stops, retail trades (*sic58xx:*

eating and drinking place, *sic55xx:* automotive related retail*, and sic54xx:* food store), personal services (*sic72xx),* and automotive related services (*sic75xx)* were found to be contributing to robbery occurrences.

There are two ways to interpret the count response model results; by comparing the expected distribution which was presented in Figure 22. The remaining part is to calculate the expected changes on robbery by predictors which are summarized in Table 33. While the interpretation can be carried out by using the increase in one unit or SD of the predictors, it is difficult to know the unit of spatial lag. Therefore, the SD is used to interpret expected discrete changes on robbery by predictors.

From Table 33, one of the strongest predictor was spatial lag of robbery. For one SD (.22) increase in the spatial lag of robbery, an areal unit's expected robbery increases by 30 percent, holding all other variables constant. When this figure was compared with number of bus stops, one SD (2.58) increase in *bstops* increases about 13 percent of expected robbery count, holding all other variables constant. Using SD, the comparable figures are 14 percent for *sic58xx* (eating and drinking place), 16 percent for *sic54xx* (food store), and 14 percent for *sic72xx* (personal services). The variable *sic55xx* (automotive dealers and gasoline service station) had the smallest impact; one SD (.32) increase in *sic55xx* is expected to increase robbery count in the areal unit by nine percent.

Table 33. Expected percentage change in robbery by predictors

Predictor	Raw coefficient	Expected percentage change in Y by		SD of X
		One unit increase X	SD increase in X	
Spatial lag	1.214**	236.7	30.2	.22
Bstops	.048**	4.9	13.1	2.58
sic58xx	.220**	24.7	14.4	.61
sic55xx	.263**	30.1	8.9	.32
sic54xx	.366**	44.2	15.5	.39
sic72xx	.202**	22.4	14.0	.65

* = p<.05, ** = p<.01

2. Analysis Results on Aggravated Assault

In this section, the influence of bus stops and commercial establishments on aggravated assault will be examined. In Table 2 in Chapter 3, the relationship between bus stops and aggravated assault was not hypothesized citing lack of empirical and theoretical basis. On the other hand, it was hypothesized that commercial establishments would be associated with increased assault. In this section, the regression results and model fit comparisons are presented. The full test results are available in Appendix B.

2.1. OLS Regression

As the first step, the diagnostics OLS regression was performed to find out more about the structure of spatial dependence as well as the effectiveness of the global model. For violence offences, the presence of spatial error seems plausible since their occurrences seem to transcend the administrative boundaries between Irvington and Newark. Table 34 shows the OLS regression results in GeoDa using spatial weight at a distance of 1,600 feet.

From Table 34, the impact of bus stops on aggravated assault was apparent; *bstop*s was statistically significant in all four models. While the hypothesis concerning the relationship between aggravated assault and number of bus stops was not formed, it turned out that *bstop*s was an important variable in predicting aggravated assault.

For the second research question, it was hypothesized that the commercial establishments would increase aggravated assault. In Model 2, the impacts of seven business category predictors on aggravated assault were examined. Only three of the seven variables were statistically significant. The variables of *sic54xx* (food store) and *sic72xx* (personal services) were related to increased aggravated assault whereas *sic75xx* (automotive related services) had the opposite impact: this variable was related with lowered numbers of aggravated assault.

When *bstops* (number of bus stop) was added to the model (Model 3), none of business variables lost their statistical significance, and *bstops* was still statistically associated with the increases in aggravated assault. Therefore, there seems to be no interaction between bus stops and commercial establishments on aggravated assault.

Table 34. Aggravated assault OLS regression with spatial lag at 1,600 feet (n=2,602)

	Model 1	Model 2	Model 3	Model 4
R-squared	.009	.041	.045	.087
Log Likelihood	-3363.24	-3320.77	-3314.90	-3256.82
AIC	6730.48	6657.53	6647.80	6543.64
Schwarz criterion	6742.21	6704.44	6700.58	6631.60
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	.384(.018)**	.352(.019)**	.341(.019)**	.196(.032)**
bstops	.032(.007)**		.023(.007)**	.026(.007)**
sic58xx		.046(.032)	.037(.032)	.035(.031)
sic55xx		.040(.054)	.033(.054)	.041(.053)
sic54xx		.277(.047)**	.271(.047)**	.250(.046)**
sic75xx		-.111(.037)**	-.118(.037)**	-.127(.036)**
sic73xx		-.023(.031)	-.027(.031)	-.010(.030)
sic72xx		.129(.030)**	.121(.030)**	.131(.029)**
finance		.050(.084)	.040(.084)	.059(.083)
mixeduse				.022(.053)
pc1fvland				.130(.035)**
phousing				.520(.055)**
univcoll				-.239(.117)*
school				.131(.056)*
openspace				-.039(.047)
Multicollinearity condition number	1.34	2.21	2.28	4.45
Normality of Errors: Jarque-Bera	43544.64**	45362.21**	44661.80**	42836.89**
Heteroskedasticity: Breusch-Pagan test	70.27(1)**	172.98(7)**	187.62(8)**	608.65(14)**
Moran's I (error)	.080**	.077**	.079**	.055**
Diagnostic result	SEM	SEM	SEM	SLM

* = p<.05, ** = p<.01

For Model 4 where the land use information was added, there was no change in statistical significance on the previous variables from Model 3. The presence of vacant lands or buildings (pc1fvland), public housing, and grade schools were related to increases in aggravated assault whereas the opposite was true for the higher educational institutions (*univcoll*). Overall, the values of R-squared were relatively low; the most complex global model (Model 4) was able to explain no more than nine percent of variance on the dependent variable. The comparable figure for robbery was about 13 percent suggesting that the predictors of bus stops, commercial establishments, and land use information are better predictors for robbery than aggravated assault.

By comparing values of LL, AIC, and SC of four models, it was seen that adding more variables to the models generally improved the model fits: we observe the bigger LL values, and smaller AIC and SC values as we add more predictors. The diagnostics showed that there was no concern related to multicollinearity. However, the statistics related to normality of errors and heteroskedasticity showed statistical significance, suggesting model misspecification problems.

2.2. Spatial Process Models

From the OLS regression diagnostics, SEM was identified as the suitable alternative for Models 1, 2, and 3. This means that the spatial dependence in the context of aggravated assault was seen as statistical nuisance. When land use information was added to the model (Model 4), the SLM was identified as a suitable spatial process model. The results of SEM and SLM estimations are summarized in Table 35.

When the SEM results were compared to the OLS regression results, there were no changes in statistical significance in predictors for Models 1, 2, and 3. The predictors of *bstop*s, *sic54xx* (food store) and *sic72xx* (personal services) were related to increased aggravated assault while *sic75xx* (automotive related services) was related to the decreased aggravated assault. However, for Model 4, the influence of educational setting (*univcoll* and *schools*) on aggravated assault had disappeared, suggesting that their influences were probably due to their neighboring values. This was not surprising for *univcoll* variable: they are located in University Heights area where four higher educational institutions are geographically clustered. The disappearance of *schools*' influences on aggravated assault in the spatial model suggests that the spatial lag can explain the apparent impacts of the schools on aggravated assaults, that there are many violent crimes in the areas surroundings the schools.

Table 35. Aggravated assault SEM and SLM estimations with spatial lag at 1,600 feet (n=2,602)

Model	SEM Model 1	SEM Model 2	SEM model 3	SLM model 4
R-squared	.073	.100	.106	.125
Log Likelihood	-3292.80	-3253.73	-3245.15	-3211.74
AIC	6589.60	6523.46	6508.29	6455.48
Schwarz criterion	6601.33	6570.37	6561.07	6549.30
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	.377(.044)**	.343(.043)**	.329(.044)**	.024(.036)
LAMBDA(Spatial lag)	.620(.048)**	.610(.048)**	.617(.048)**	.522(.051)**
bstops	.037(.007)**		.028(.007)**	.027(.007)**
sic58xx		.055(.043)	.045(.031)	.037(.031)
sic55xx		.061(.052)	.050(.053)	.054(.052)
sic54xx		.241(.045)**	.233(.045)**	.231(.045)**
sic75xx		-.089(.036)*	-.098(.036)**	-.112(.035)**
sic73xx		.005(.030)	.004(.030)	-.001(.029)
sic72xx		.140(.029)**	.131(.029)**	.129(.028)**
finance		.083(.081)	.068(.081)	.069(.081)
mixeduse				.028(.052)
pc1fvland				.086(.034)*
phousing				.414(.054)**
univcoll				-.<u>181(.115)</u>
school				.<u>087(.054)</u>
openspace				-.065(.046)
Heteroskedasticity: Breusch-Pagan(df)	86.13(1)**	182.49(7)**	197.74(8)**	635.39(14)**

* = p<.05, ** = p<.01
Note: A variable coefficient is underlined when the significance level is different compared to SEM estimation.

By comparing the values of LL, AIC, and SC of four models, it was concluded that adding more variables to the models generally improved the model fits: we observe the bigger LL values, and smaller AIC and SC values.

2.3. Assessing the Linear Model Fits and Improvements

The fit statistics of four models are presented in Table 36. When the LL, AIC, and SC values of four models between OLS regression and spatial process model were compared, it showed that the spatial regression model estimations increased the model fits for all models; the LL values increased whereas the values of AIC and SC decreased relative to OLS regression results for all models.

This means that taking account of spatial dependence in the model improved the model fits. In fact, the spatial autoregressive coefficient (LAMBDA) was again one of the highly significant predictors in predicting aggravated assault. Despite the model improvements made, the diagnostic tests revealed that there were significant problems with heteroskedasticity, suggesting a possible model misspecification. There were violations in the expected orders of Wald test, LR test, and LM statistics for all models, possibly suggesting the problems with model specification.

Table 36. Fit statistics on aggravated assault in OLS regression and spatial process models

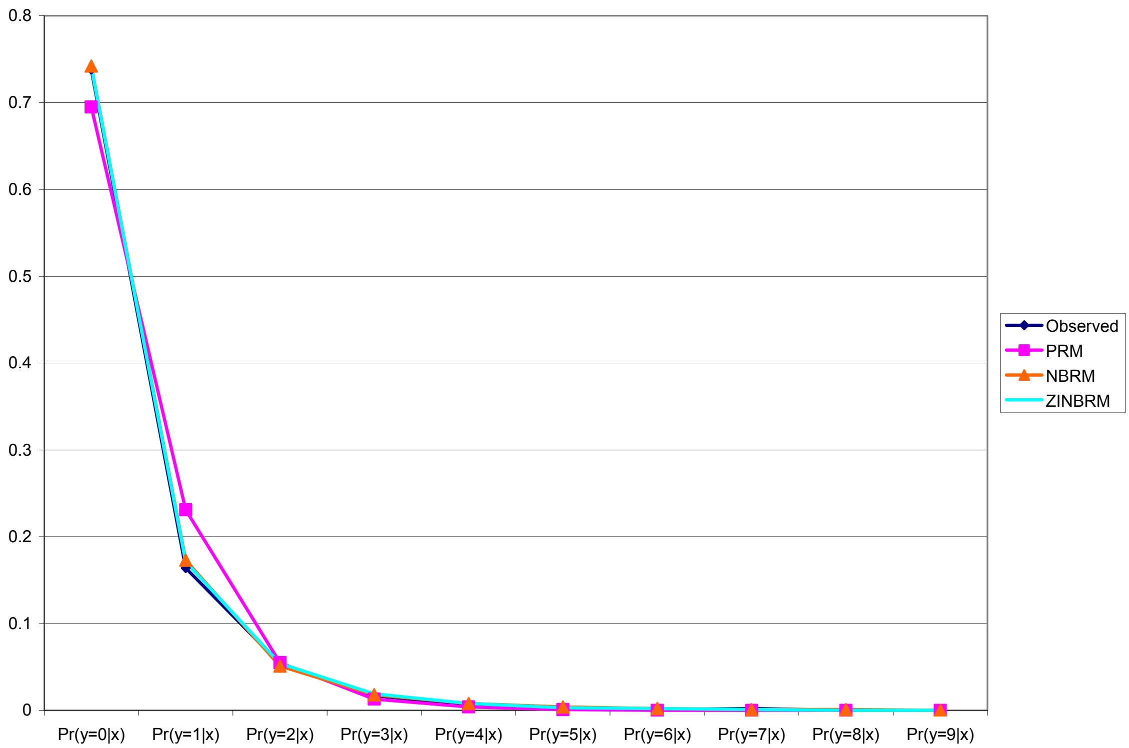
	Model 1	Model 2	Model 3	Model 4
OLS regression				
Log Likelihood	-3363.20	-3320.77	-3314.90	-3256.82
Akaike info criterion	6730.48	6657.53	6647.80	6543.64
Schwarz criterion	6742.21	6704.44	6700.58	6631.60
Spatial model				
Log Likelihood	-3292.80	-3253.73	-3245.15	-3211.74
Akaike info criterion	6589.60	6523.46	6508.29	6455.48
Schwarz criterion	6601.33	6570.37	6561.07	6549.30
Wald test	169.39	158.11	165.84	104.37
Likelihood Ratio	140.88	134.07	139.50	90.16
Lagrange multiplier	326.55	303.36	318.83	166.69

2.4. Negative Bi-nominal Regression

The same as with robbery analysis, three types of count data regression methods (PRM, NBRM, and ZINBRM) were performed on all four models. Figure 23 shows the distributions of PRM, NBRM and ZINBRM on aggravated assault using Model

4. From Figure 23, it is apparent that PRM underperforms both NBRM and ZINBRM by under-predicting zeros and over-predicting count value ones as was the case with robbery prediction.

Figure 23. Comparison of aggravated assault distributions by PRM, NBRM, and ZINBRM estimations of Model 4



The results of fit statistics comparisons on four models were almost identical to that of robbery results. The fit statistics showed that NBRM was favored by PRM for all models. The LL test of alpha = 0 was statistically significant, an evidence that there was a statistically significant difference on estimations between PRM and NBRM. However, this was not the case for NBRM and ZINBRM comparisons. Again for BIC test, NBRM was preferred over ZINBRM. However, AIC and Vuong test showed preference for ZINBRM over NBRM for all models. After reviewing the model fit comparison, NBRM was chosen as the regression method. The full fit statistics comparing PRM, NBRM, and ZINBRM are available in Appendix B.

Table 37 shows the summary results of NBRM on aggravated assault with a spatial lag using distance of 1,600 feet. For all models, spatial lag was highly significant signaling the strong impacts of neighbors. For Models 1 to 3, there was no difference on significance level from the spatial process models. For aggravated assault, spatial lag, *bstops*, *sic54xx* (food store), and *sic72xx* (personal services) all contributed to aggravated assault. Only one variable, *sic75xx* (automotive related services) was shown to be related to the lowered numbers of aggravated assault.

When compared to spatial process model, there were two predictors whose significance levels changed in NBRM estimation in Model 4. The impact of vacant lands or buildings (pc1fvland) lost its statistical significance in relation to aggravated assault. On the other hand, *univcoll* (higher educational institutions) gained its statistical significance after accounting for non-normality and overdispersion of the data.

The next step was to find out which model in NBRM estimations was the best in predicting aggravated assault. The model fit statistics comparisons are presented in the last two rows in Table 37. The fit statistics comparison between Model 1 and Model 2 showed very strong support for Model 2, a model using commercial establishments as the predictors. While this supports the hypothesis, it is also important to remind the reader that Model 2 contained seven variables in addition to the spatial lag while Model 1 contained only one variable in addition to the spatial lag. When compared Model 2 and Model 3, there was positive support for Model 3. Lastly, the fit statistic comparisons showed that there was strong support for Model 4 over Model 3. Therefore, a conclusion can be drawn that land use information is useful in analyzing spatial patterns of aggravated assault. However, a caveat needs to be taken with this finding: the aggravated assault in this study contained violence between intimates, and interpersonal factors may play a bigger role in this type of crime than factors such as bus stops and commercial establishments.

Table 37. Aggravated assault NBRM results with spatial lag at 1,600 feet (n=2,602)

	NBRM Model 1	NBRM Model 2	NBRM Model 3	NBRM Model 4
Pseudo R-squared	.048	.063	.065	.078
LL test of alpha=0	274.00**	238.96**	226.55**	203.45**
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	-1.935(.085)**	-2.003(.087)**	-2.035(.088)**	-2.154(.102)**
Spatial lag	1.995(.147)**	1.979(.145)**	1.989(.144)**	1.703(.148)**
bstops	.080(.015)**		.050(.014)**	.059(.014)**
sic58xx		.092(.064)	.066(.064)	.058(.063)
sic55xx		.154(.122)	.124(.123)	.123(.123)
sic54xx		.404(.090)**	.395(.089)**	.362(.088)**
sic75xx		-.333(.108)**	-.340(.107)**	-.361(.106)**
sic73xx		.039(.070)	.021(.071)	.055(.071)
sic72xx		.249(.057)**	.217(.057)**	.234(.056)**
finance		.126(.172)	.103(.172)	.128(.169)
mixeduse				.064(.121)
pc1fvland				.220(.086)
phousing				.684(.107)**
univcoll				-.849(.421)*
school				.182(.120)
openspace				-.181(.113)
Log Likelihood	-2080.04	-2048.41	-2042.46	-2015.07
AIC	4168.08	4166.83	4106.92	4064.15
BIC	4191.54	4175.47	4171.42	4163.83
Difference in BIC from previous model	n/a	16.07	4.05	7.59
Scalar measure model fit compared to previous model		Very strong support for Model 2	Positive support for Model 3	Strong support for Model 4

* = p<.05, ** = p<.01
Note: A variable coefficient is underlined when the significance level is different compared to spatial process model estimation.

2.5. Comparisons of Model 4 Estimations

Model 4 was identified as the best model in NBRM in Stata using scalar measures. For this reason, Model 4 regression results from the OLS regression, SLM, and NBRM are compared in Table 38. There were two variables whose statistical significance levels changed from OLS regression when spatial dependence was accounted for; in SLM, the influences of higher educations and schools disappeared compared to the OLS regression results. However, in NBRM where non-normality and overdispersion of the data were addressed, *univcoll* became statistically significant to lowered level of aggravated assault. The opposite was true for *phousing* (public housing); it was related to increased number of assault, different from the SLM result. *pc1fvland*(vacant lands or buildings), while significant in both OLS regression and SLM estimations, lost its statistical significance in NBRM.

Table 38. Model 4 comparisons on aggravated assault with spatial lag at 1,600 feet (n=2,602)

Regression method	OLS regression	SLM	NBRM
R-squared	.087	.125	.078
Log Likelihood	-3256.82	-3211.74	-2015.073
Predictor	Coefficient(SE)	Coefficient(SE)	Coefficient(SE)
Constant	.196(.032)**	.024(.036)	-2.154(.102)**
Spatial lag		.522(.051)**	1.703(.148)**
bstops	.026(.007)**	.027(.007)**	.059(.014)**
sic58xx	.035(.031)	.037(.031)	.058(.063)
sic55xx	.041(.053)	.054(.052)	.123(.123)
sic54xx	.250(.046)**	.231(.045)**	.362(.088)**
sic75xx	-.127(.036)**	-.112(.035)**	-.361(.106)**
sic73xx	-.010(.030)	-.001(.029)	.055(.071)
sic72xx	.131(.029)**	.129(.028)**	.234(.056)**
finance	.059(.083)	.069(.081)	.128(.169)
mixeduse	.022(.053)	.028(.052)	.064(.121)
pc1fvland	.130(.035)**	.086(.034)*	.220(.086)
phousing	.520(.055)**	.414(.054)**	.684(.107)**
univcoll	-.239(.117)*	-.181(.115)	-.849(.421)*
school	.131(.056)*	.087(.054)	.182(.120)
openspace	-.039(.047)	-.065(.046)	-.181(.113)

* = p<.05, ** = p<.01

Figure 23 presents the observed and expected distributions of aggravated assault using Model 4. Table 39 shows the expected percentage change in aggravated assault by predictors. Using SD of predictors, one SD change in spatial lag of aggravated assault increases the expected aggravated assault by 58 percent, by far the highest influence among the predictors. The second largest impact was seen from *phousing*. Since *phousing* was a binary variable, it was not suitable to use SD for its interpretation. Therefore, the impact of *phousing* predictor was interpreted using its presence; the presence of public housing increase the aggravated assault by 92 percent, holding all other variables constant.

There were only two predictors related to lower aggravated assault. One unit increase in *sic75xx* (automotive related services) is expected to lower aggravated assault by 14 percent in the areal unit, holding all other variables constant. However, *univcoll* was not statistically significant when only the statistically significant predictors were run in the model, therefore not interpreted. For both *sic54xx* and *sic72xx*, one SD increase is expected to increase aggravated assault by about 18 percent in the areal unit holding all other predictors constant. For *bstops*, one SD increase in *bstops* would increase aggravated assault by 17 percent, the smallest influences among the predictors.

To summarize, it seems that both number of bus stops and commercial establishments seem to influence the occurrences of aggravated assault. Some of the commercial establishments were related to lower number of aggravated assaults, possibly signaling the positive impacts of businesses in reducing crime opportunities.

Table 39. Expected percentage change in aggravated assault by predictors

Predictor	Raw coefficient	Expected percentage change in Y by		SD of X
		Unit increase X	SD increase in X	
Spatial lag	1.760**	480.9	57.9	.26
bstops	.060**	6.2	16.8	2.58
sic54xx	.414**	51.2	17.7	.39
sic75xx	.322**	-27.5	-14.3	.48
sic72xx	.258**	29.4	18.2	.65
phousing	.652**	92.0	22.0	.31
univcoll	-.793	-54.8	-10.8	.14

* = p<.05, ** = p<.01

3. Analysis Results on Motor Vehicle Theft

In Table 2 in Chapter 3, it was hypothesized that both bus stops and commercial establishments would increase the crime opportunity for motor vehicle theft. In this section, the data analysis results on motor vehicle theft will be discussed in detail. The regression and model fit comparison results are available in Appendix C.

3.1. OLS Regression

The OLS regression results on motor vehicle theft in Table 40 shows that number of bus stops contributed to the occurrences in motor vehicle theft. For the impacts of commercial establishments on motor vehicle theft (Model 2), *sic58xx* (eating and drinking place), *sic54xx* (food store), and *sic72xx* (personal services) were statistically significant to increased motor vehicle theft. Somewhat surprisingly, *sic55xx* (motor vehicle related retail) and *sic75xx* (motor vehicle related service) did not show statistical significance with motor vehicle theft. This can be explained by suggesting that people usually do not leave their vehicles unattended for a long time in gas stations, auto dealers, or auto repair shops.

In Model 3, while the statistical significance of *sic58xx* (eating and drinking place) seemed to have disappeared when *bstops* (number of bus stops) was added to the model, its p value was .051. Therefore, there seems to be very little interaction among the predictors of bus stops and commercial establishments. For Model 4, all of the statistically significant variables from Model 3 retained their statistical significance. For land use information variable, four out of six variables were statistically significant to motor vehicle theft; *pc1fvland* (vacant lands or buildings) and *schools* (grade k-12) were related to the increased motor vehicle theft while *phousing* (public housing) and *univcoll* (higher educational institution) contributed to the decreases.

From the diagnostic tests, the multicollinearity was not a concern. However, there were problems suggested with normality of errors and heteroskedasticity as with other types of crimes already discussed.

Table 40. Motor vehicle theft OLS regression with spatial lag at 1,600 feet (n=2,602)

Regression method	OLS model 1	OLS model 2	OLS Model 3	OLS Model 4
R-squared	.012	.034	.039	.055
Log Likelihood	-5364.43	-5334.86	-5327.56	-5305.82
AIC	10732.90	10685.70	10673.10	10641.60
Schwarz criterion	10744.60	10732.60	10725.90	10729.60
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	1.564(.039)**	1.457(.042)**	1.431(.042)**	1.325(.070)**
bstops	.080(.014)**		.056(.015)**	.061(.015)**
sic58xx		.158(.069)*	.135(.069)	.123(.069)
sic55xx		.052(.118)	.037(.117)	.038(.117)
sic54xx		.558(.101)**	.542(.101)**	.509(.101)**
sic75xx		.120(.080)	.103(.079)	.095(.079)
sic73xx		.096(.066)	.087(.066)	.098(.066)
sic72xx		.178(.064)**	.159(.064)*	.143(.064)*
finance		-.033(.183)	-.059(.182)	-.069(.181)
mixeduse				.277(.117)*
pc1fvland				.149(.077)
phousing				-.275(.120)*
univcoll				-.672(.258)**
school				.539(.123)**
openspace				-.159(.103)
Multicollinearity condition number	1.34	2.21	2.28	4.44
Normality of errors: Jarque-Bera	4896.47	5177.61**	4818.96**	4855.26**
Heteroskedasticity: Breusch-Pagan (df)	100.58(1)**	111.33(7)**	207.95(8)**	226.34(14)**
Moran’s I (error)	.099**	.091**	.094**	.087**
Diagnostic result	SEM	SLM	SEM	SLM

* = p<.05, ** = p<.01

3.2. Spatial Process Models

For motor vehicle theft, Models 1 and 3 diagnostics identified SEM as the alternative while the SLM was the choice for Models 2 and 4. The spatial regression estimation results are summarized in Table 41.

For all models, the spatial lag was highly statistically significant, again signaling the influence of neighboring values. *Bstops* was related to increases in motor vehicle theft in Model 1. In Model 2, in addition to spatial lag, three variables were shown to be statistically significant to increased motor vehicle theft; *sic58xx* (eating and drinking place), *sic54xx* (food store), and *sic72xx* (personal services). When *bstops* was added to the model (SEM Model 3), the statistical significance of *sic72xx* disappeared while *sic73xx* (business services) and *sic58xx* (eating and drinking place) became statistically significant. No changes occurred in the statistical significance levels of *bstops* and *sic54xx* compared to the OLS result; both of them were related to the increases in motor vehicle theft.

Table 41. Motor vehicle theft SEM and SLM with spatial lag at 1,600 feet (n=2,602)

	SEM model 1	SLM model 2	SEM Model 3	SLM Model 4
R-squared	.096	.108	.116	.125
Log Likelihood	-5269.60	-5249.71	-5239.07	-5233.66
Akaike info criterion	10543.20	10517.40	10496.10	10479.30
Schwarz criterion	10554.93	10570.20	10548.91	10573.10
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	1.522(.108)**	.420(.081)**	1.401(.104)**	.277(.096)**
LAMBDA(Spatial lag)	.669(.043)**	.639(.045)**	.656(.045)**	.632(.044)**
bstops	.087(.014)**		.068(.014)**	.063(.014)**
sic58xx		.150(.066)*	.<u>145(.067)*</u>	.115(.066)
sic55xx		.088(.113)	.090(.113)	.076(.112)
sic54xx		.472(.097)**	.424(.097)**	.426(.096)**
sic75xx		.118(.076)	.136(.078)	.091(.076)
sic73xx		.120(.064)	.<u>141(.064)*</u>	.119(.063)
sic72xx		.143(.062)*	.<u>111(.062)</u>	.<u>113(.061)</u>
finance		-.011(.175)	-.033(.174)	-.046(.174)
mixeduse				.272(.112)*
pc1fvland				.<u>155(.073)*</u>
phousing				-.<u>107(.116)</u>
univcoll				-.<u>371(.247)</u>
school				.480(.118)**
openspace				-.167(.099)
Heteroskedasticity: Breusch-Pagan test(df)	76.75(1)**	134.93(7)**	194.26(8)**	225.99(14)**

*= p<.05, ** = p<.01

Note: A variable coefficient is underlined when the significance level is different compared to OLS regression result.

For Model 4, among the seven business category variables examined, none of the service industries (sic75xx, *sic73xx*, and *sic72xx*) were statistically significant. For retail businesses, only *sic54xx* (food store) was statistically significant. For land use predictors, the presence of mixed land use, vacant lands or buildings, and schools were related to increased motor vehicle theft. Compared with the OLS regression Model 4, *sic72xx*, *phousing*, and *univcoll* (higher educational institutions) lost their statistical significance when spatial dependence was accounted for. When compared to the statistical significances between the OLS regression and spatial process model results, there were several differences in Model 3 and Model 4. This observation is different from that of robbery and aggravated assault where very few differences across the regression results were observed. This can be interpreted as the varying impacts of spatial dependence in predicting motor vehicle theft compared to robbery or aggravated assault . As with other crime types already discussed, heteroskedasticity was still a concern.

3.3. Assessing the Linear Model Fits and Improvements

The fit statistics of four models are presented in Table 42. The comparisons of LL, AIC, and SC values of four models between OLS regression and spatial process models showed that accounting for spatial dependence increased the model fits for all models. Checking the orders of Wald test, LR test, and LM statistics, again the ordering violated the expected order as with the two previously discussed crime types, suggesting the problems with model specification.

Table 42. Fit statistics on motor vehicle theft in OLS regression and spatial models

	Model 1	Model 2	Model 3	Model 4
OLS regression				
Log Likelihood	-5364.43	-5334.86	-5327.56	-5305.82
Akaike info criterion	10732.90	10685.70	10673.10	10641.60
Schwarz criterion	10744.60	10732.60	10725.90	10729.60
Spatial model				
Log Likelihood	-5269.60	-5249.71	-5239.07	-5233.66
Akaike info criterion	10543.20	10517.40	10496.10	10479.30
Schwarz criterion	10554.93	10570.20	10548.91	10573.10
Wald test	236.76	206.07	216.44	204.86
Likelihood Ratio	189.65	170.30	176.98	164.31
Lagrange multiplier	496.71	430.70	446.82	391.04

3.4. Negative Bi-nominal Regression

The three types of count data regression methods (PRM, NBRM, and ZINBRM) were performed on motor vehicle theft using four models. Figure 24 shows the distributions of PRM, NBRM and ZINBRM on motor vehicle theft using Model 3. From Figure 24, it is again apparent that PRM underperforms both NBRM and ZINBRM.

The results of fit statistics comparisons on four models were almost identical to other crime type results; the NBRM was the favored method over the PRM for all models. For comparison between the NBRM and ZINBRM, the BIC test preferred NBRM to ZINBRM. For AIC and Vuong test, they showed preference for ZINBRM over NBRM for all four models. After reviewing the model fit comparisons, NBRM was chosen as the regression method for the same reasons: preference of parsimonious model and no theoretical basis to think that there are two latent groups in the data. The full fit statistics comparing PRM, NBRM, and ZINBRM are available in Appendix C.

Figure 24. Comparison of motor vehicle theft distributions by PRM, NBRM, and ZINBRM estimations of Model 3

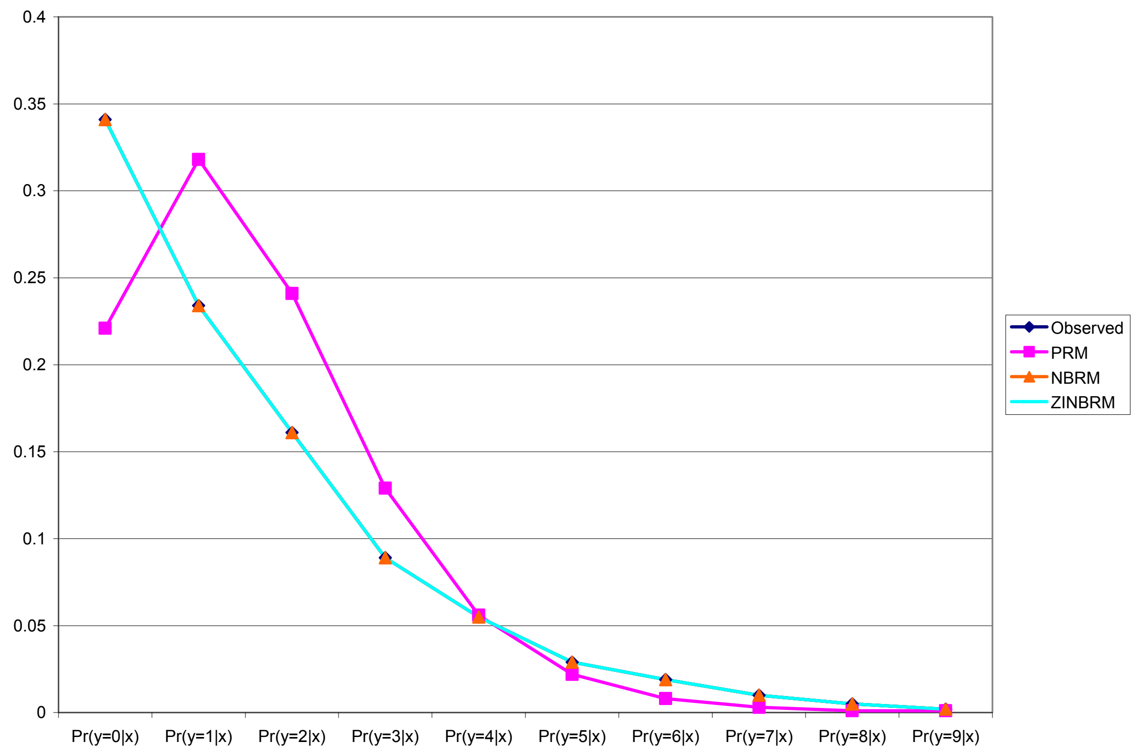


Table 43 shows the results of NBRM on motor vehicle theft with a spatial lag using distance of 1,600 feet. For all models, spatial lag was highly significant signaling the strong impacts of neighbors. By adding more variables, the model fits were improved based on AIC. The LL test of alpha = 0 was statistically significant, evidence that there were statistically significant differences on estimations between the PRM and NBRM.

For Model 1, *bstops* (number of bus stops) was statistically significant to the increased motor vehicle theft. In Model 2, the NBRM estimation produced different result from the SLM estimation; *sic73xx* (business services) was associated with increased motor vehicle theft. As was the case with the spatial process model, *sic72xx* (personal services), *sic58xx* (eating and drinking place), *sic54xx* (food store) were related to increases in motor vehicle theft.

Table 43. Motor vehicle theft NBRM with spatial lag at 1,600 feet (n=2,602)

	Model 1	Model 2	Model 3	Model 4
Pseudo R-squared	.026	.032	.033	.036
LL test of alpha=0	689.86**	686.94**	667.29**	630.70**
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	.427(.065)**	-.479(.065)**	-.490(.065)**	-.561(.075)**
Spatial lag	.503(.035)**	.494(.035)**	.493(.034)**	.481(.035)**
bstops	.039(.008)**		.025(.008)**	.027(.008)**
sic58xx		.090(.038)*	.080(.038)*	.066(.038)
sic55xx		.075(.067)	.063(.067)	.060(.066)
sic54xx		.197(.055)**	.189(.054)**	.177(.054)**
sic75xx		.074(.046)	.066(.046)	.056(.045)
sic73xx		.079(.038)*	.076(.038)*	.083(.037)*
sic72xx		.096(.034)**	.087(.034)*	.087(.034)*
finance		-.019(.099)	-.040(.099)	-.028(.098)
mixeduse				.176(.066)**
pc1fvland				.106(.046)*
phousing				-.042(.075)
univcoll				-.179(.169)**
school				.235(.068)**
openspace				-.111(.062)
Log Likelihood	-4404.74	-4381.61	-4376.50	-4361.07
AIC	8817.48	8783.21	8774.99	8756.15
BIC	8840.94	8841.85	8839.50	8855.84
Difference in BIC from previous model	n/a	-.91	1.44	-16.34
Scalar measure model fit compared to previous model		Weak support for Model 1	Weak support for Model 3	Very strong support for Model 3

* = p<.05, ** = p<.01

Note: A variable coefficient is underlined when the significance level is different compared to spatial process model estimation.

Adding *bstops* to the model did not change statistical significance of the commercial establishment predictors. In SEM Model 3, *sic72xx* (personal services) was not statistically significant but in NBRM, this variable was statistically significant. When NBRM Model 4 was compared to SEM Model 4, both *sic72xx* and sic73xx became statistically significant to increases in motor vehicle theft. In addition, the presence of mixed land use, vacant lands or buildings, and schools all contributed to motor vehicle theft. On the other hand, *univcoll* (higher educational institutions) had the opposite impacts. From Table 43, it can be seen that accounting for non-normality of the data produced regression results quite different from spatial process models.

To find out which model was the best model in predicting motor vehicle theft, comparisons of fit statistics of the four models were carried out. The fit statistics are presented at the very bottom of Table 43. Based on the scalar measure (mainly BIC), there was a weak support for Model 1 over Model 2 in predicting motor vehicle theft. This can be interpreted as that number of bus stops is better predictor of motor vehicle theft than commercial establishments. In substantive sense, the areas where bus stops are located may provide ample spaces either on-street or in parking structure. When compared Model 1 and Model 3, this showed that the combination of bus stops and commercial establishments predicted the occurrences of motor vehicle theft better than the bus stops alone. Therefore, it can be concluded that both bus stops and commercial establishments independently contributes to motor vehicle incidents. The fit statistics on Model 3 and 4 showed that there was very strong support for Model 3 over Model 4. This suggests that bus stops and commercial establishments are suitable measures in predicting motor vehicle theft.

3.5. Comparison of Model 3 Estimations

Based on the scalar measure, Model 3 was identified as the best model in NBRM. For this reason, the Model 3 regression results from OLS regression, SLM, and NBRM are compared in Table 44. Table 44 shows that the global OLS regression produced very low R-squared value. In addition, there were several differences in the statistical significance levels from the OLS regression result compared to the SEM estimation. For instance, *sic58xx* was statistically significant in SEM and NBRM but not in the OLS regression. This was the same for *sic73xx*.

Regardless of the regression method used, *bstops* and *sic54xx* (food store) were always related to increases in motor vehicle theft. Other predictors, however, changed their statistical significance when spatial dependence was accounted for. For both SEM and NBRM, *sic58xx* (eating and drinking place), and *sic73xx* (business services) were associated with the increases in motor vehicle theft. When the non-normality of the data was addressed by NBRM, *sic72xx* (personal services) became statistically significant predictor to increased motor vehicle theft.

Table 44. Model 3 estimation comparison on motor vehicle theft with spatial lag at 1,600 feet (n=2,602)

Regression methods	OLS Model 3	SEM Model 3	NBRM Model 3
R-squared	.039	.116	.033
Log Likelihood	-5327.56	-5239.07	-4376.50
AIC	10673.10	10496.10	8774.99
Schwarz criterion	10725.90	10548.91	8839.50
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	1.431(.042)**	1.401(.104)**	-.490(.065)**
Spatial lag(LAMBDA)		.656(.045)**	.493(.034)**
bstops	.056(.015)**	.068(.014)**	.025(.008)**
sic58xx	.135(.069)	.145(.067)*	.080(.038)*
sic55xx	.037(.117)	.090(.113)	.063(.067)
sic54xx	.542(.101)**	.424(.097)**	.189(.054)**
sic75xx	.103(.079)	.136(.078)	.066(.046)
sic73xx	.087(.066)	.141(.064)*	.076(.038)*
sic72xx	.159(.064)*	.111(.062)	.087(.034)*
finance	-.059(.182)	-.033(.174)	-.040(.099)

* = p<.05, ** = p<.01

Table 45 shows the expected percentage change in motor vehicle theft by predictors. Using SD of predictors, one SD increase in spatial lag of motor vehicle theft is expected to increase motor vehicle theft by 35 percent, by far the highest influence compared to other predictors. For *sic54xx* (food store), one unit increase in SD is expected to increase motor vehicle theft by 8 percent, holding all other predictors constant. For bus stop, the figure is about 7 percent. The smallest influence was observed in *sic72xx* (personal services); one unit increase in SD is expected to cause about 6 percent increase in motor vehicle theft.

Table 45. Expected percentage change in motor vehicle theft by predictors

Predictor	Raw coefficients	Expected percentage change in Y by		SD of X
		Unit increase X	SD increase in X	
Spatial lag	.491**	63.4	35.3	.62
bstops	.026**	2.6	6.9	2.58
sic58xx	.084*	8.8	5.3	.61
sic54xx	.185**	20.3	7.6	.39
sic73xx	.081*	8.5	4.8	.58
sic72xx	.086*	8.9	5.7	.65

* = p<.05, ** = p<.01

4. Analysis Results on Theft From Motor Vehicle

When the hypotheses were formulated regarding theft from motor vehicle, it was speculated that theft from motor vehicle will display the similar patterns as motor vehicle theft. In this section, the data analysis results on theft from motor vehicle will be discussed in detail. All relevant regression results and fit statistic comparisons are presented in Appendix D.

4.1. OLS Regression

Table 46 shows the OLS regression results on theft from motor vehicle using GeoDa for all models. In Model 1, the variable of *bstops* was statistically significant to increased theft from motor vehicle. When examined the influences of commercial establishments on theft from motor vehicle (Model 2), *sic58xx* (eating and drinking place), *sic54xx* (food store), and *sic72xx* (personal services) were statistically significant to increases in theft from motor vehicle. In Model 3, *sic72xx* became not statistically significant, signaling interaction effect with *bstops*. However, there were no other changes in statistical significance of *bstops, sic58xx and sic54xx* from Model 2.

Table 46. Theft from motor vehicle OLS regression with spatial lag at 1,600 feet (n=2,602)

	Model 1	Model 2	Model 3	Model 4
R-squared	.007	.020	.019	.031
Log Likelihood	-4846.62	-4829.54	-4825.74	-4813.80
AIC	9697.24	9675.54	9669.47	9657.59
Schwarz criterion	9708.97	9722.00	9722.25	9745.56
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	.899(.032)**	.825(.034)**	.809(.035)**	.691(.058)**
bstops	.050(.012)**		.033(.012)**	.031(.012)**
sic58xx		.133(.057)*	.120(.057)**	.109(.057)
sic55xx		.126(.097)	.117(.097)	.111(.097)
sic54xx		.266(.083)**	.257(.083)**	.246(.083)**
sic75xx		.107(.065)	.097(.065)	.086(.065)
sic73xx		.033(.055)	.028(.055)	.028(.055)
sic72xx		.110(.053)*	.098(.053)	.100(.053)
finance		.062(.150)	.046(.150)	.033(.150)
mixeduse				.318(.097)**
pc1fvland				.131(.064)*
phousing				-.165(.100)
univcoll				.144(.213)
school				.197(.101)
openspace				.004(.085)
Multicollinearity condition number	1.34	2.21	2.28	4.45
Normality of errors/ Jarque-Bera	333273.5**	322009.9**	323032.2**	333530.9**
Heteroskedasticity: Breusch-Pagan test	8.75(1)**	277.23(7)**	279.28(8)**	333.02(14)**
Moran’s I (error)	.082**	.083**	.080**	.078**
Diagnostic result	SLM	SLM	SLM	SLM

* = p<.05, ** = p<.01

For Model 4, *sic58xx* lost its statistical significance. The only stable predictors across the models were *bstops* and *sic54xx* (food store). For land use predictors, the presence of mixed land use and *pc1fvland* (vacant lands or buildings) contributed to the occurrences of theft from motor vehicle. Overall, the R-squared of the models were very low suggesting the weakness of the models. The OLS regression diagnostics showed that SLM was the preferred method in dealing with spatial dependence in theft from motor vehicle.

4.2. Spatial Process Models

The OLS regression diagnostics on spatial dependence showed that SLM was the suitable estimation method on theft from motor vehicle. For this reason, the SLM estimations were performed on theft from motor vehicle in GeoDa. For theft from motor vehicle, the spatial lag was highly statistically significant for all models, again signaling the influence of neighboring values. In Model 1, *bstops* was statistically significant to increased theft from motor vehicle. In Model 2, only one predictor (sic54xx) was statistically significant in predicting theft from motor vehicle. This is perhaps not surprising since *sic54xx* was the only significant commercial establishment predictor in OLS regression methods across the four models.

Table 47. Theft from motor vehicle SLM with spatial lag at 1,600 feet (n=2,602)

	SLM model 1	SLM model 2	SLM model 3	SLM model 4
R-squared	.078	.089	.091	.098
Log Likelihood	-4769.00	-4752.31	-4750.27	-4739.45
AIC	9544.00	9522.61	9520.54	9510.89
Schwarz criterion	9561.60	9575.39	9579.18	9604.72
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	.295(.051)**	.226(.052)**	.219(.053)**	.093(.068)
Spatial lag	.647(.045)**	.644(.045)**	.638(.045)**	.638(.045)**
bstops	.038(.011)**		.023(.012)*	.024(.012)*
sic58xx		.<u>102(.055)</u>	.092(.055)	.081(.055)
sic55xx		.156(.093)	.149(.093)	.140(.093)
sic54xx		.235(.080)**	.228(.080)**	.211(.080)**
sic75xx		.111(.063)	.104(.063)	.091(.063)
sic73xx		.021(.053)	.018(.053)	.024(.053)
sic72xx		.<u>099(.051)</u>	.090(.051)	.091(.051)
finance		.091(.145)	.080(.145)	.071(.144)
mixeduse				.294(.093)**
pc1fvland				.155(.061)*
phousing				-.083(.096)
univcoll				-.133(.205)
school				.152(.098)
openspace				.015(.082)
Heteroskedasticity: Breusch-Pagan (df)	8.42(1)**	253.55(7)**	256.10(8)**	309.96(14)**

* = p<.05, ** = p<.01

Note: A variable coefficient is underlined when the significance level is different compared to OLS regression result.

For Model 3, only *bstops* and *sic54xx* were statistically significant to theft from motor vehicle. Unexpectedly, the two variables related to automotive *(sic55xx* and *sic75xx*) were not statistically significant in predicting occurrences of theft from motor vehicle even after taking account of spatial dependence. For land use predictors in Model 4, the presence of mixed land use and *pc1fvland* (vacant lands or buildings) contributed to the increases in theft from motor vehicle. All other land use predictors were not statistically significant in predicting theft from motor vehicle.

4.3. Assessing the Linear Model Fits and Improvements

The fit statistics of four models are presented in Table 48. Again, it showed that accounting for spatial dependence increased the model fits for all models. The ordering of Wald test, LR test, and LM statistics violated the expected order, again suggesting possible model misspecification.

Table 48. Fit statistics on theft from motor vehicle in OLS regression and spatial process models

	Model 1	Model 2	Model 3	Model 4
OLS regression				
Log Likelihood	-4846.62	-4829.54	-4825.74	-4813.80
Akaike info criterion	9697.24	9675.54	9669.47	9657.59
Schwarz criterion	9708.97	9722.00	9722.25	9745.56
Spatial process model				
Log Likelihood	-4769.00	-4752.31	-4750.27	-4739.45
Akaike info criterion	9544.00	9522.61	9520.54	9510.89
Schwarz criterion	9561.60	9575.39	9579.18	9604.72
Wald test	205.18	207.02	200.83	205.45
Likelihood Ratio	155.24	154.50	150.93	148.70
Lagrange multiplier	354.92	348.38	336.90	320.08

4.4. Negative Bi-nominal Regression

The three types of count data regression methods (PRM, NBRM, and ZINBRM) were performed using all four models on theft from motor vehicle. The fit statistics comparisons yielded the same results as other crime types examined: the NBRM was preferred to PRM for all models, and there was conflicting support for NBRM over

ZINBRM. The NBRM results are summarized in Table 49. By looking at the coefficients for all models, it can be seen that the spatial lag of theft from motor vehicle is highly statistically significant.

No changes occurred on the statistical significance in NBRM Model 1 compared to spatial process model result. For Model 2, three predictors (*sic55xx*: automotive related retail, *sic54xx*: food store, *sic75xx*: automotive related services) were associated with the increases in theft from motor vehicle. When *bstops* was added to the model (Model 3), the statistical significance of commercial establishment predictors remained the same from Model 2.

Table 49. Theft from motor vehicle NBRM with spatial lag at 1,600 feet (n=2,602)

	NBRM Model 1	NBRM Model 2	NBRM Model 3	NBRM Model 4
Pseudo R-squared	.032	.036	.037	.041
LL test of alpha=0	651.70**	623.58**	623.29**	599.03**
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	-.961(.067)**	-1.025(.069)**	-1.031(.069)**	-1.186(.083)**
Spatial lag	.822(.059)**	.819(.059)**	.811(.059)**	.810(.059)**
bstops	.036(.011)*		.023(.011)*	.023(.011)*
sic58xx		.076(.048)	.067(.048)	.049(.048)
sic55xx		.193(.086)*	.179(.086)*	.172(.085)*
sic54xx		.148(.072)*	.149(.072)*	.136(.072)
sic75xx		.126(.058)*	.123(.058)*	.114(.057)*
sic73xx		.015(.047)	.014(.047)	.022(.047)
sic72xx		.070(.045)	.063(.045)	.072(.045)
finance		.132(.130)	.111(.130)	.103(.128)
mixeduse				.309(.084)**
pc1fvland				.182(.061)**
phousing				-.131(.098)
univcoll				-.128(.187)
school				.142(.091)
openspace				-.011(.081)
Log Likelihood	-3371.29	-3357.87	-3355.57	-3340.67
AIC	6750.58	6735.75	6733.14	6715.33
BIC	6774.04	6794.39	6797.65	6815.02
Difference in BIC from previous model	n/a	-20.35	-23.61	-40.99
Scalar measure model fit compared to previous model		Very strong support for Model 1	Very strong support for Model 1	Very strong support for Model 1

* = p<.05, ** = p<.01

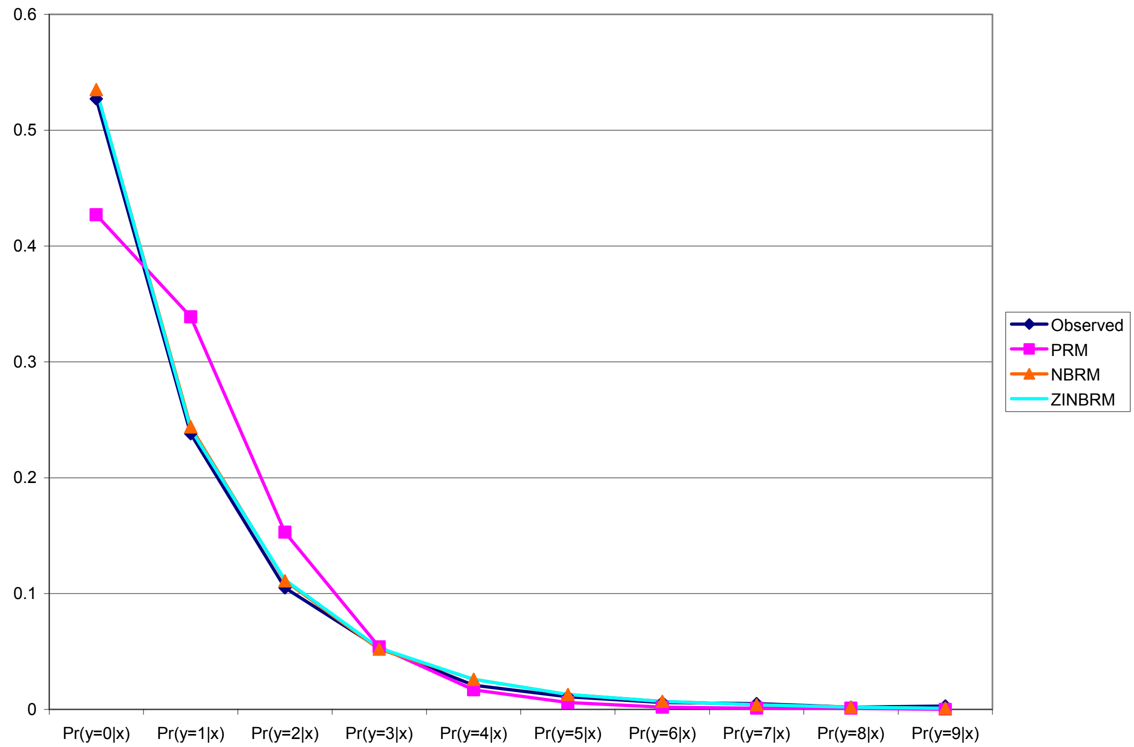
Note: A variable coefficient is underlined when the significance level is different compared to spatial process model estimation.

For Model 4, while *sic54xx* was not statistically significant, its p value was .057, not quite different from spatial process model. Two land use predictors were statistically significant to increased theft from motor vehicle: *mixeduse* and *pc1fvland* (vacant lands or buildings). The fact that *pc1fvland* seemed to increase the theft from motor vehicle can be interpreted as low levels of guardianship may increase the opportunities for theft from motor vehicle. For mixed land use, this area can be described as edges where the distinction between insider and outsider cannot be easily made, therefore, providing support for crime pattern theory.

To find out which model is the best model, the model fit comparisons in NBRM estimations were performed in Stata. The results are presented in the two bottom rows in Table 49. The result showed that Model 1 was the strongly preferred model over the remaining three models. When the coefficients were examined in more detail, it was found that the spatial lag was the most significant predictor than any other variables. Therefore, it was concluded that the identification of Model 1 as the best model was probably due to the undue influence of spatial lag, rather than the strong performance of bus stops in predicting theft from motor vehicle.

Since the analysis of motor vehicle theft did not present similar results, it was unexpected to observe underperformance of the models in predicting theft from motor vehicle. Many predictors were unstable across the border except for the few, it was decided to build a new model using all predictors until the predictors lost their statistical significance. Figure 25 shows the comparison of Poisson, NBRM and ZINBRM on theft from motor vehicle using an ad hoc model. The comparison of fit statistics shows that NBRM is the preferred method to PRM and ZINBRM. Table 50 summarizes the NBRM result and expected percentage change in theft from motor vehicle by predictors.

Figure 25. Comparison of theft from motor vehicle distributions by PRM, NBRM, and ZINBRM estimations using an ad hoc model



The new *ad hoc* model showed that *bstops* (number of bus stops) was related to the increases in theft from motor vehicle. This was not surprising since *bstops* was always statistically significant in predicting theft from motor vehicle. For predictors of commercial establishments, both motor vehicle related businesses (*sic55xx* and *sic75xx*) were statistically significant to increased theft from motor vehicle. In addition, *sic54xx* (food store) was also related to increased risks of theft from motor vehicle. The presence of mixed land use and *pc1fvland* (vacant lands or buildings) also contributed to the increased theft from motor vehicle. The biggest influence was exerted by the spatial lag as expected.

When expected percentage changes in theft from motor vehicle were calculated, the spatial lag had by far the highest influence than any other predictor; one SD increase in spatial lag is expected to increase theft from motor vehicle by 47 percent, holding all other predictors constant. For *bstops*, one unit increase in SD is expected to increase theft from motor vehicle by eight percent, holding all variables constant. The figures for *sic55xx* and *sic75xx* are 6 percent each.

For the two binary variables, the use of SD in their interpretations is not appropriate. For this reason, their interpretations are carried out by using their presence; the presence of mixed land use in the areal unit is expected to increase the theft from motor vehicle by approximately 11 percent holding all other variables constant. The comparable figure for presence of vacant lands or buildings is 19 percent.

Table 50. Theft from motor vehicle NBRM with spatial lag at 1,600 feet (n=2,602)

Predictor	Coefficients (S.E)	Expected percentage change in Y by		SD of X
		Unit increase in X	SD increase in X	
Constant	-1.172(.080)**			
Spatial lag	.823(.058)**	127.8	46.6	.47
bstops	.028(.011)**	2.9	7.6	2.58
sic55xx	.178(.085)*	19.5	5.9	.32
sic54xx	.207(.067)**	23.0	8.5	.39
sic75xx	.116(.057)*	12.3	5.7	.48
mixeduse	.318(.084)**	37.5	10.6	.32
pc1fvland	.171(.060)**	18.6	8.6	.48
LL test of alpha=0	608.17**			
Log Likelihood	-3346.88			
AIC	6711.76			
BIC	6765.54			

* = p<.05, ** = p<.01

5. Analysis Results on Burglary

In this section, analysis results of burglary will be discussed. Earlier in Chapter 3, the hypothesis regarding the relationship between bus stops and burglary was not formulated due to the lack of empirical basis. The same was true for second research question: there was no hypothesis regarding the influences of commercial establishments on burglary. Detailed regression results and fit statistics comparisons are available in Appendix E.

5.1. OLS Regression

Table 51 summarizes the OLS regression results for burglary. For burglary, the OLS regression result on Model 1 shows that the number of bus stops did not contribute crime opportunities for burglary. In fact, for all other remaining models, *bstops* (number of bus stops) was not significant. In Model 2, only *sic54xx* (food store) was shown to be statistically significant to burglary opportunities. This was the same in Model 3 when number of bus stop was added to the model.

For the last model (Model 4), the existence of *pc1fvland (*vacant lands or buildings), public housings, and schools were statistically significant to increased burglary. The presence of *univcoll* (higher educational institutions) did not seem to have any influence on burglary. This is somewhat expected since in this areas, the majority of the residences are on campus housing. Any incidents on the campus housing would have been handled by the university police, rather than by the city police. Based on the Rutgers University Police statistics, there were a total of seven burglaries occurred in year 2007. Overall, the R-squared for all models were very low; less than five percent in the final model.

For the diagnostic section, there was no concern for multicollinearity. When the values of LL, AIC, and SC were compared among OLS regression results, Model 2 showed some improvements from Model 1 mainly because one of the predictors was statistically significant in predicting burglary. Since *bstops* was not statistically significant, it did not improve the model fit for Model 3 over Model 2. There were three statistically significant land use variables in Model 4; *pc1fvland*, *phousing*, and *schools*. Therefore, the presence of public housings, vacant lands or buildings, and grade schools were related to increased burglary.

Table 51. Burglary OLS regression with spatial lag at 1,600 feet (n=2,602)

	Model 1	Model 2	Model 3	Model 4
R-squared	.001	.029	.029	.047
Log Likelihood	-3417.9	-3381.67	-3381.25	-3356.49
AIC	6839.81	6779.34	6780.51	6742.98
Schwarz criterion	6851.54	6826.25	6833.28	6830.94
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	.454(.018)**	.417(.020)**	.414(.020)**	.281(.033)**
bstops	.013(.007)		.006(.007)	.009(.007)
sic58xx		.028(.033)	.025(.033)	.026(.033)
sic55xx		.005(.056)	.003(.056)	.006(.055)
sic54xx		.355(.048)**	.353(.048)**	.331(.048)**
sic75xx		-.043(.038)	-.045(.038)	-.051(.037)
sic73xx		.011(.031)	.010(.031)	.018(.031)
sic72xx		.033(.030)	.030(.030)	.034(.030)
finance		-.106(.086)	-.109(.086)	-.084(.086)
mixeduse				-.086(.055)
pc1fvland				.198(.036)**
phousing				.162(.057)**
univcoll				-.206(.122)
school				.120(.058)*
openspace				-.066(.049)
Multicollinearity condition number	1.34	2.21	2.28	4.45
Normality of Errors: Jarque-Bera	20229.85**	16945.04**	16975.97**	16157.74**
Heteroskedasticity: Breusch-Pagan test	1.35	287.69(7)**	288.60(8)**	432.16(14)**
Moran's I (error)	.122(n/a)	.115(n/a)	.116(n/a)	.103(n/a)
Diagnostic result	SEM	SLM	SLM	SLM

* = p<.05, ** = p<.01

5.2. Spatial Process Models

The diagnostics results in OLS regression showed that there were problems with normality of errors and heteroskedasticity. For burglary, SLM estimation was identified as suitable method to address the spatial dependence for Models 2, 3 and 4. For Model 1, SEM was identified as an alternative to the OLS regression estimation.

In this section, the spatial lag was added to the estimation to address spatial dependence. Table 52 shows two things: the spatial lag and LAMBDA were highly statistically significant. In fact, its coefficient was by far the highest. Second, among the predictors of bus stops and commercial establishments, *sic54xx* (food store) was the only statistically significant predictor in burglary. This may signal the fact that food stores are located in close proximity or in residential areas. In Model 3, *bstops* was not statistically significant while *sic54xx* was still related to increased burglary. In Model 4, the presence of *pc1fvland* (vacant lands or buildings) and public housing were shown to be related to increased burglary incidents. The presence of schools did not seem to influence burglary when spatial dependence was accounted for. The diagnostic showed that there was a problem with heteroskedasticity.

Table 52. Burglary SEM and SLM with spatial lag at 1,600 feet (n=2,602)

	SEM Model 1	SLM Model 2	SLM model 3	SLM model 4
R-squared	.109	.129	.129	.137
Log Likelihood	-3293.82	-3262.96	-3262.14	-3249.48
AIC	6591.64	6543.92	6544.28	6530.96
Schwarz criterion	6603.37	6596.69	6602.92	6624.79
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	.448(.056)**	.092(.026)**	.088(.026)**	.020(.036)
LAMBDA(Spatial lag)	.704(.040)**	.691(.041)**	.692(.041)**	.668(.042)**
bstops	.018(.007)**		.008(.007)	.011(.007)
sic58xx		.035(.031)	.031(.031)	.031(.031)
sic55xx		.031(.052)	.029(.053)	.030(.052)
sic54xx		.302(.045)**	.300(.045)**	.286(.045)**
sic75xx		-.025(.035)	-.028(.036)	-.034(.035)
sic73xx		.037(.030)	.035(.030)	.040(.030)
sic72xx		.029(.029)	.026(.029)	.028(.029)
finance		-.052(.082)	-.056(.082)	-.043(.081)
mixeduse				-.044(.053)
pc1fvland				.124(.034)**
phousing				.126(.054)*
univcoll				-.108(.116)
school				.068(.055)
openspace				-.088(.046)
Heteroskedasticity: Breusch-Pagan(df)	2.72	286.11(7)**	287.27(8)**	421.57(14)**

* = p<.05, ** = p<.01
Note: A variable coefficient is underlined when the significance level is different compared to OLS regression result.

5.3. Assessing the Model Fits and Improvements

As were the cases for other crime types examined, adding spatial lag to the model improved the model fits compared to OLS regression estimations. Table 53 summarizes the model fit statistics. The tests statistics of Wald, LR, and LM do not conform to the expected order, again suggesting that there may be problems related to model misspecification.

Table 53. Fit statistics on burglary in OLS regression and spatial process models

	Model 1	Model 2	Model 3	Model 4
OLS regression				
Log Likelihood	-3417.90	-3381.67	-3381.25	-3356.49
Akaike info criterion	6839.81	6779.34	6780.51	6742.98
Schwarz criterion	6851.54	6826.25	6833.28	6830.94
Spatial process model				
Log Likelihood	-3293.82	-3262.96	-3262.14	-3249.48
Akaike info criterion	6591.64	6543.92	6544.28	6530.96
Schwarz criterion	6603.37	6596.69	6602.92	6624.79
Wald test	302.80	287.24	290.33	252.08
Likelihood Ratio	248.17	237.42	238.23	214.01
Lagrange multiplier	753.65	702.20	703.77	589.38

5.4. Count Response Regression Models

As were the cases with previously discussed criminal offences, PRM, NBRM and ZINBRM were performed on all four models. The model fit statistics showed preference for NBRM over PRM. For NBRM and ZINBRM, it was not clear which regression method was a better choice. As a preliminary data analysis, the NBRM was chosen as the regression method on burglary and presented in Table 54.

In Model 1, *bstops* was statistically significant to increased burglary. However, the influence of spatial lag on burglary was in much greater magnitude. For Model 2, *sic54xx* (food store) was the only significant predictor associated with increased burglary. In Model 3, *bstops* was not statistically significant while *sic54xx* was significant, the same as the OLS regression result. For Model 4, *bstops* became significant as with *pc1fvland* (vacant lands or buildings), and *phousing* (public housing). The review of the NBRM results showed that number of bus stops and commercial establishments may not be the best predictors for burglary. In this research study, there were no hypotheses formed regarding the influences of bus stops and commercial establishments on burglary citing the lack of empirical basis. Perhaps due to this reason, it was not surprising to see that the models underperformed in predicting burglary.

Table 54. Burglary NBRM with spatial lag at 1,600 feet (n=2,602)

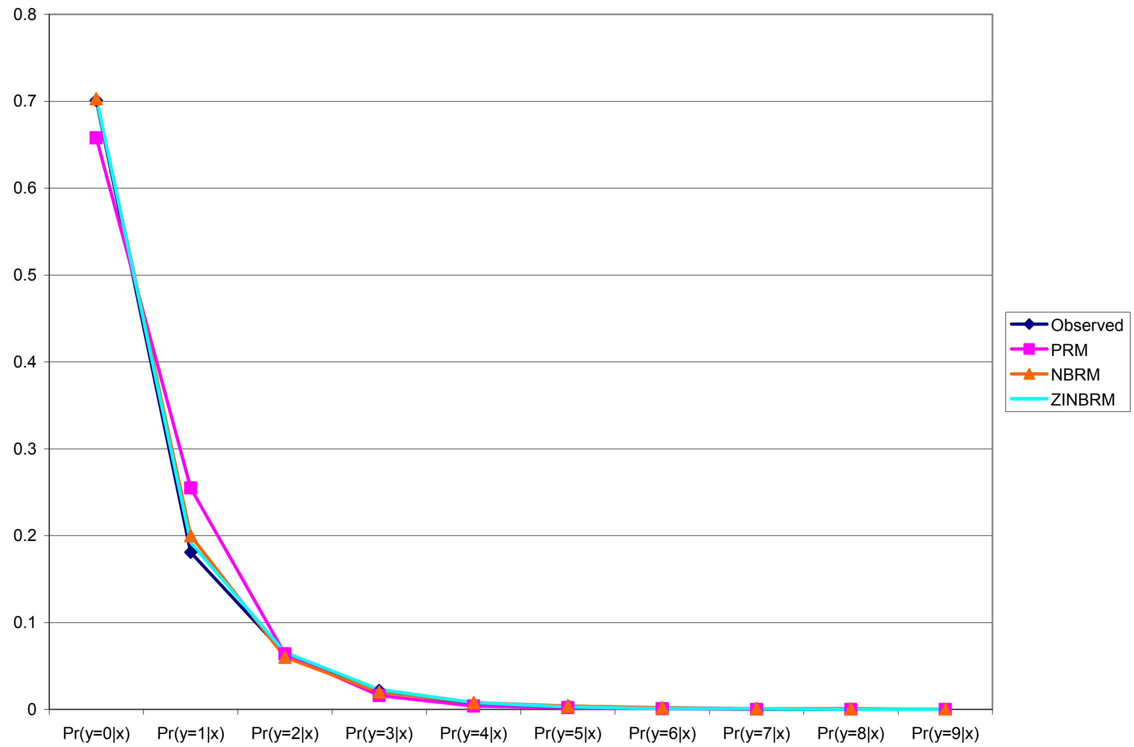
	Model 1	Model 2	Model 3	Model 4
Pseudo R-squared	.054	.062	.063	.067
LL test of alpha=0	185.05**	159.44**	158.14**	145.03**
Predictor	Coeff(SE)	Coeff(SE)	Coeff(SE)	Coeff(SE)
Constant	-1.677(.069)**	-1.728(.071)**	-1.742(.071)**	-1.868(.088)**
LAMBDA(Spatial lag)	1.569(.099)**	1.526(.098)**	1.531(.098)**	1.457(.099)**
bstops	.038(.013)**		.022(.014)	.028(.014)*
sic58xx		.063(.061)	.050(.062)	.045(.062)
sic55xx		.140(.107)	.133(.108)	.134(.107)
sic54xx		.334(.078)**	.331(.078)**	.307(.078)**
sic75xx		-.064(.084)	-.071(.084)	-.083(.083)
sic73xx		.087(.057)	.085(.057)	.102(.057)
sic72xx		.091(.055)	.081(.056)	.084(.056)
finance		-.024(.174)	-.036(.175)	-.019(.174)
mixeduse				-.015(.115)
pc1fvland				.238(.080)**
phousing				.276(.106)**
univcoll				-.385(.322)
school				.010(.114)
openspace				-.176(.015)
LL test of alpha=0	185.05**	159.44**	158.14**	145.03**
Log likelihood	-2241.09	-2222.83	-2221.59	-2210.97
AIC	4490.18	4465.64	4465.19	4455.93
BIC	4513.64	4524.28	4529.69	4555.62
Difference in BIC from previous model	n/a	-10.65	-16.06	-41.98
Scalar measure model fit compared to previous model		Very strong support for Model 1	Very strong support for Model 1	Very strong support for Model 1

* = p<.05, ** = p<.01

Note: A variable coefficient is underlined when the significance level is different compared to spatial process model estimation.

Yet to decide the best model among the four models, the scalar measures of fit statistics were computed and compared. The comparisons of four models showed that Model 1 was very strongly favored over all other models. The same as the case with theft from motor vehicle, this result was due to the undue influence of spatial lag rather than the strong performance of bus stops in predicting burglary. For this reason, Model 4 was run repeatedly until all the predictors remained statistically significant. Figure 26 shows the distributions of PRM, NBRM, and ZINBRM of burglary based on an ad hoc model.

Figure 26. Comparison of burglary distributions by PRM, NBRM, and ZINBRM estimations using an ad hoc model



For the ad hoc model, the NBRM was the favored regression method over PRM. The fit statistics comparison of ZINBRM and NBRM estimations showed conflicting results: BIC test favored NBRM whereas Voung and AC tests favored ZINBRM. Because there was no hypothesis regarding what contributed to the occurrences of burglary, it was decided to examine the burglary using both NBRM and ZINBRM. Table 55 shows the result of NBRM and ZINBRM.

According to NBRM result, *bstops*, *sic54xx*, *pc1fvland*, *phousing* were associated with increased burglary in addition to the spatial lag. Spatial lag of burglary was one of the most significant variables, highlighting the importance of neighboring values in explaining burglary.

Table 55. Burglary NBRM and ZINBRM with spatial lag at 1,600 feet (n=2,602)

	NBRM	ZINBRM regression model	
LL test of alpha=0	151.26**		
Log Likelihood	-2218.19	-2203.46	
AIC	4450.37	4432.91	
BIC	4491.42	4509.14	
Predictor	Coeff (SE)	Count Coeff(SE)	Inflate Coeff(SE)
Constant	-1.864(.083)**	-1.399(.150)**	.693 (.442)
Spatial lag	1.443(.098)**	.984(.130)**	-5.323(1.345)**
bstops	.031(.013)*	.057(.019)**	.059 (.037)
sic54xx	.392(.072)**	.345(.079)**	-.277 (.473)
pc1fvland	.239(.079)**	.260(.108)*	.062 (.425)
phousing	.258(.107)*	.181(.130)	-.264 (.721)

* = p<.05, ** = p<.01

In ZINBRM, it is assumed that there are two distributions; always zero group (inflate column) and not always zero groups (count column). Always zero group values are thought to be due to structural reasons while not always zero group values represent the areal units whose opportunity for burglary is not possible. The ZINBRM regression estimation showed that spatial lag, *bstops, sic54xx*, and *pc1fvland* were statistically significant to increased burglary.

The binary distribution (inflate) contains coefficients for the factor change in the odds of belonging to the always zero group compared to the odds of belonging to the not always zero group. For inflate column (always zero group), the increases in spatial lag were related to reduced odds of having zero burglary, meaning that the spatial lag in inflate column displayed the opposite sign compared to the count column, producing the same results.

Table 56. Expected percentage change in burglary by predictors in NBRM

Predictor	Raw coefficient	Expected percentage change in Y by		SD of X
		Unit increase X	SD increase in X	
Spatial lag	1.443**	323.4	59.1	.32
bstops	.031**	3.2	8.4	2.58
sic54xx	.392**	48.0	16.7	.39
pc1fvland	.239**	27.0	12.2	.48
phousing	.258*	29.4	8.2	.31

* = p<.05, ** = p<.01

The next step was to calculate the expected percentage change on burglary by predictor. For easier interpretation of the result, the expected changes in burglary by predictors using NBRM estimation were calculated and summarized in Table 56. The expected percentage change in burglary by spatial lag was by far the greatest; one SD increase in spatial lag is expected to increase burglary by 59 percent, holding all other variables constant. The second biggest impact was seen by *sic54xx* (food store); one SD increase is expected to increase burglary by 17 percent. For the two binary variables, the presence of *pc1fvland* (vacant lands or buildings) and *phousing* (public housing) would increase expected burglary by 12 and 8 percentages respectively.

Table 57. Expected percentage change in burglary by predictors in ZINBRM

Predictor	Raw coefficient	Expected percentage change in Y by		SD of X
		Unit increase X	SD increase in X	
Count				
Spatial lag	.984**	167.5	37.2	.32
bstops	.057**	5.9	15.9	2.58
sic54xx	.344**	41.1	14.5	.39
phousing	.260*	29.7	13.3	.48
Inflate				
Spatial lag	-5.323**	-99.5	-82.0	.32

* = p<.05, ** = p<.01

Table 57 shows the expected percentage change in burglary in ZINBRM estimation. For those with possibility of burglary (not always zero group), having public housing in the unit increases the expected burglary by 30 percent, holding all other factors constant. The strongest impact was again from the spatial lag; one SD increase in spatial lag is expected to increase burglary by 37 percent, holding all other predictors constant. Both *bstops* and *sic54xx* have about the same impacts on the expected burglary; about 15 and 16 percentages each.

For binary distribution (the odds of having zero burglary in the units), one SD increase in spatial lag of burglary is expected to decrease the chance of having zero burglary (or being in the always zero group) by 82 percent, holding all other variables constant. All other remaining predictors were not statistically significant in the binary column.

6. Summary of Data Analysis Results

The present research study analyzed five types of criminal offences using different visual and regression methods. The aim of the study was to answer two research questions; whether bus stops or commercial establishments increase crime opportunities in the vicinities. The theoretical frameworks were based on three opportunity theories; routine activity theory, crime pattern theory, and rational choice theory. It was hypothesized that bus stops would increase crime opportunities for robbery, motor vehicle theft, and theft from motor vehicles. It was also hypothesized that commercial establishments would increase crime opportunities for robbery, aggravated assault, motor vehicle theft, and theft from motor vehicle. Not hypothesized were the influences of bus stops and commercial establishments on burglary.

The first step of data analysis was non-spatial descriptive statistics. In this step, it was revealed that the data were over-dispersed and highly skewed; the majority of the areal units contained zero values. The second step of data analysis was visual examinations of crime patterns in Newark. The visual presentations of the exploratory spatial pattern analyses using LISA and density maps showed that spatial patterns among the five offences seemed to be different albeit the core of clusters may be the same; that there were crime clusters in areas adjacent to Irvington and East Orange for all crime types examined. Another apparent spatial pattern was relatively low crime occurrences in University Heights area. Also downtown Newark was relatively crime free, perhaps due to close proximity to the higher educational institutions with their own police forces patrolling the area in addition to the city police.

To examine the impacts of bus stops and crime in relation to commercial establishments, four regression models were built using variables of bus stops, commercial establishments, and land use information. The first model contained number of bus stops as the predictor. The second model contained the seven types of commercial establishments as the predictors. The third model combined Model 1 and Model 2. In the final model, the land use information was added to Model 3. For spatial process models and count response models, spatial lag was added to each model to account for spatial dependence.

The next step was hypothesis testing by employing several regression methods. The diagnostics from the OLS regression in GeoDa provided information regarding whether the spatial dependence was observed in dependent variable or in error term. The LM results from the OLS regression in GeoDa were used to decide whether to perform SLM or SEM on data analysis. The results of spatial process models were compared withthe results of the OLS regression to examine the impacts of spatial dependence on the models. The comparisons of LL, AIC, and SC showed that accounting for spatial dependence increased the model fits. In addition, adding more predictors to the models generally improved the model fits. The model fits improvements, however, were not observed at times from Model 3 to Model 4, suggesting that land use information may not add new information when the presences of bus stops and commercial establishments were taken account for. The diagnostics in spatial process model suggested problems with non-normality of the data and possible model misspecifications.

Despite the improved model fits from the global OLS regression to spatial process model, the spatial process models are still liner models and do not address non-normality of the data. For this reason, it was decided to perform count response modeling methods which do not assume normality of the data. The three types of count response regressions were performed: PRM, NBRM, and ZINBRM. For all five crime types examined, the PRM always underperformed NBRM by under-predicting zeros and over-predicting some of count values. When the fit statistics from NBRM and ZINBRM were compared, it was not clear which regression model was the better choice. The BIC which punishes severely for extra parameters always preferred NBRM to ZINBRM. However, Voung and AC tests preferred ZINBRM to NBRM. In the end, the NBRM was preferred over ZINBRM for two reasons. First, a preference was given to a parsimonious model over complicated one such as ZINBRM. Since there was some support for the NBRM over ZINBRM, the method of choice appeared to be a judgment call. Second, there was no basis to speculate that the zero counts were caused by unobserved latent grouping variable. For burglary, both the NBRM and ZINBRM results were presented and discussed. This was because there was no hypothesis formed for burglary in relation to bus stops and commercial establishments in this study.

For all five crime types examined, number of bus stops was always associated with increased crimes. This supports theoretical framework of crime opportunity theories. While an examination of temporal patterns will help understand dynamics of crime opportunities further, it is clear that bus stops attract both offenders and victims, increasing crime opportunities in the local neighborhoods.

The impacts of businesses on crime were examined by using seven types of business predictors. Regression results showed that some of these predictors displayed more salient impacts than other business types. Overall, the retail categories exerted more influences than services categories in crime occurrences. Food store *(sic54xx)* was associated with increased crime opportunities for all five crime types. This was the only variable other than bus stops that affected all five crime types. The regression results showed that *sic58xx* (eating and drinking place) was statistically significant to increases in robbery and motor vehicle theft. Unexpectedly, this variable was not associated with increased aggravated assault. The fact that two retail businesses (eating and drinking places and automotive related retail) were not associated with aggravated assault while associated with increased robbery may be explained by what aggravated assault consists of in Newark. This crime type included intimate partner violence as well as violence between strangers or acquaintances. The spatial patterns of violence between intimates may be quite different from that of violence between strangers or acquaintances. Also the fact that presence of public housing was associated with increased aggravated assault suggests that some of these violent crimes may occur between family members. For *sic55xx* (automotive dealers and gasoline stations), it was statistically significant to robbery and theft from motor vehicle but not for motor vehicle theft. This can be explained by hypothesizing that motor vehicles may be left unattended for a short time in these areas but not long enough time to allow theft of motor vehicle to occur.

For the service industries, *sic75xx* (automotive related services) was related to lower aggravated assaults. This predictor was also related to increases in theft from motor vehicle offences but not statistically significant in relation to motor vehicle theft. As with automotive related retail businesses, this can be again explained by hypothesizing that motor vehicles in these areas may be left unattended for a short time period for theft from it to occur. The variable *sic73xx* (business services) was related to increased motor vehicle theft. For *sic72xx* (personal services), it was related to increased robbery and motor vehicle theft. Therefore, it seemed that commercial establishments, retail businesses in particular, attracted more targets in the areas, possibly increasing the risks of predatory crimes.

For land use information, mixed land use was related to increases in robbery, motor vehicle theft, and theft from motor vehicle. This would support a perspective in crime pattern theory. The mixed landuse areas often called edges would experience higher number of criminal incidents because offenders would have easier time blending in. The presence of vacant lands or buildings was related to increases in motor vehicle theft, theft from motor vehicle, and burglary. The areas where vacant land or buildings were located would represent areas with lower levels of guardianship, therefore, facilitating commission of crime. The presence of public housing was significant for three crime types: robbery, aggravated assault, and burglary. For aggravated assault and burglary, this may simply represent increased target densities. There was only one other variable in addition to automotive related services that seemed to lower the crime in the area: the higher educational institutions for aggravated assault and motor vehicle theft. For these two crime types, it is not clear whether this finding was due to a measurement error; that crimes on the campus are not usually reported to Newark Police Department. The presence of grade schools seems to have increased crime opportunities for robbery and motor vehicle theft.

There were two predictors which had never displayed statistical significance regardless of crime types examined or regression methods used; finance (banks), and *openspace* (park and cemetery). For *openspace*, it is not clear how much influence the methods in recording addresses and geocoding had on this finding. For instance, the police may record crime occurred in a park to a nearest intersection, and crimes with a park address may not be geocoded. Also crime occurred in the cemetery may have the addresses of cemetery rather than actual locations of crime within the cemetery.

Since there were four models, the last step was to decide which model was the best model. The model fits in count response models were discussed using two methods. First, the comparisons of observed and expected distributions were presented in graphs for all crime types. The graphs showed expected distributions in PRM, NBRM, and ZINBRM as well as the observed distribution. It was clear from the graph that the PRM underperformed NBRM and ZINBRM for all crime types examined. The second method in assessing the model fit was to examine the expected changes on crime by predictors in NBRM or ZINBRM. The general theme was the overwhelmingly strong impacts of spatial lag on the dependent variables than any other predictors for all five crime types.

Table 58. Predictors shown to be statistically significant to crime opportunities in NBRM

	Robbery	Assault	Motor vehicle theft	Theft from motor vehicle	Burglary
Best model	3	4	3	n/a (M1)	n/a (M1)
Spatial lag	sig	sig	sig	sig	sig
bstops	sig	sig	sig	sig	sig
sic58xx	sig		sig		
sic55xx	sig			sig	
sic54xx	sig	sig	sig	sig	sig
sic75xx		sig(-)		sig	
sic73xx			sig		
sic72xx	sig	sig	sig		
finance					
mixeduse	sig		sig	sig	
pc1fvland			sig	sig	sig
phousing	sig	sig			sig
univcoll		sig(-)	sig(-)		
school	sig		sig		
openspace					

Using scalar measures of fit, Model 3 was chosen as the best model for robbery and motor vehicle theft. This means that both bus stops and commercial establishments were good predictors in explaining crime patterns for robbery and motor vehicle theft. For these two crimes, land use information was not as useful predictors. For aggravated assault, Model 4 was chosen as the best model. The presence of public housing was associated with the increases whereas the higher educational institutions were associated with the decreases.

For theft from motor vehicle and burglary, Model 1 using only two predictors (spatial lag and bus stops) was chosen as the best model. Closer examinations of these two models showed that this was due to the over-performance of spatial lag rather than the strong impacts of bus stops. For this reason, ad hoc models were created to find out which predictors were statistically significant to these two types of crime. For burglary, only one type of commercial establishments (*sic54xx*: food store) was associated with increased burglary. In addition, *bstops*, *pc1fvand* (vacant building or land) or *phousing* (public housing) contributed to increases in burglary. For theft from motor vehicle, three types of businesses were associated with the increases; *sic55xx* (automotive related retail), *sic54xx* (food store), and *sic75xx* (automotive related services). In addition, the presence of mixed land use and *pc1fvand* (vacant lands or buildings) were also related to increased theft from motor vehicle. Lastly, the regression results showed that there were very few interactions among the predictors. Therefore, it appears that both bus stops and commercial establishments influence crime opportunities. Table 58 contains the summary of the NBRM regression results of five crime types.

CHAPTER 9. CONCLUSIONS AND DISCUSSIONS

1. Limitations and Suggestions for Future Research

There are problems related to research in general ranging from measurement problems which plague the field of criminal justice as shown by spousal assault results[9], or questions regarding whether empirical research can aid in answering normative questions (Moore, 2002). Setting these limitations aside, there are several limitations related to the present study. In this section, the study limitations and suggestions for future research will be discussed.

1.1. Absence of Data from Other Police Departments

The crime data used in this study were provided by Newark City Police Department. The data from PATH Police and NJ Transit Police were not available for the present study. The areas served by the PATH Police were excluded from the analysis so the absence of PATH Police crime data is not likely a concern. However, the NJ Transit Police is legally responsible for maintaining the order and protecting the properties in the NJ transit system. Since the buses, light rails, and commuter rails are used by the public regularly, what goes on inside the mass transit may prove to be important in the context of the present study. For this reason, crime data on mass transit ingrained within the city should be utilized whenever possible to provide a fuller picture on crime patterns and crime opportunity dynamics in the city.

[9] For instance, Maxwell and his colleagues (2002) found that when using follow-up interviews rather than official crime records, the observed relationship between minority and spousal assault disappeared. Indeed, the white suspects had higher recidivism rates than their ethnic minority counterparts in spousal assault cases using follow-up interviews.

While the use of crime data on mass transit is desired, there is a reason to suspect that mass transit crimes are seriously underreported (Levine and Wachs, 1986) (Crime Concern, 2004). A recent household survey in Britain showed that only 20 percent of the victims on the public transport reported their victimization. Only a quarter of these reports went to the transit related personnel or the transit police (Crime Concern, 2004). Some of the citied reasons for not reporting their victimization to the transit police were difficulties in reporting; the victims could not locate a telephone or transit officer to report the crime (Levine and Wachs, 1986). In addition, some may not know of the existence of separate transit police, and many citizens simply call the city police rather than transit police as a matter of convenience (Liggett et al., 2001).

Table 59. Crimes reported to University Police Departments in Newark

Crime type	Newark 07	Rutgers 07	UMDNJ 06	NJIT 06	ECC 06
Robbery	1,217	5	14	15	14
Aggravated assault	1,063	0	20	4	20
motor vehicle theft	4,227	23	23	38	15
Attempted motor vehicle theft	Included in motor vehicle theft	n/a	14	n/a	n/a
Burglary	1,207	7	6	8	6

Note: The offences include incidents happened on campus, non-campus building or property, and on adjacent public property including public streets.
1. Rutgers University crime statistics are published by Department of Public Safety, and available on http://nwkpolice.rutgers.edu/stats.html
2. UMDNJ crime statistics are published by Department of Public Safety, and available on http://umdnjcaprod.umdnj.edu/ops/pdf/2007_Annual_Security_Report.pdf
3. NJIT crime statistics are published by Department of Public Safety, and available on http://www.njit.edu/publicsafety/pdf/crime-stats-04-06.pdf
4. Essex Count College (ECC) crime statistics are published by Department of Public Safety, and available on http://www.essex.edu/police/clery.report.html

In addition to the NJ Transit Police crime data, there are at least four separate university and college police departments in charge of maintaining public safety in and around their institutions. Table 59 summarizes the crime statistics from three universities and one college in the University Heights area. The examinations of the crime statistics on and around the campus showed two things. First, there are relatively small numbers of crimes in University Heights. In Newark, a total of 55 areal units included higher educational institutions in their unit. The sum of these areal units was .31 square mile, which is about 1.8 percent of total area size. While the crime types examined are more than 1.8 percent of total crime, it is not clear whether University Heights area is more dangerous considering the high ambient population in the area. Unfortunately, the crime locations were not collected in detail, so it is impossible to pinpoint where these incidents occurred when the premises identified as an “adjacent public street” or on campus.

1.2. Absence of Temporal Analyses: Short-Term and Long-Term

The reason for spatial pattern analysis is to explain the spatial process that form the distributions or the causes of changes in the patterns (Wong and Lee, 2005). It is suggested that the changes in spatial patterns at different times need to be analyzed to fully understand the dynamics and anatomy of behaviors related to geography (Wong and Lee, 2005). In fact, the spatial patterns observed in the present study are the result of the full year’s process which was summed up at the end of the year. However, the observed patterns may not be stable throughout the year. In the future, analyzing spatial patterns of more than one year can detect spatial changes in the long term. This information will be useful in understanding the relationships between bus stops and crime in conjunction with commercial establishments.

In addition, there is evidence that crime concentrates in time as well as in space. In what is referred to as a stationary fallacy, there is a widespread assumption that crime hot spots are 24-7 high crime areas despite the temporal variation during the day. In fact, temporal variation of crime in a day has been demonstrated in robbery (Felson and

Poulsen, 2003), motor vehicle theft (Rengert, 1997), bus shelter damage (Newton, 2004), assault against bus crews (Department for Transport, 2002), burglary (Ratcliffe, 2004), and crimes nearby bus stops (Loukaitou-Sideris, 1999; Liggett et al., 2001). Indeed, bus service frequency and ridership level vary significantly depending on the time of the day even on the same bus lines. Frequencies of bus services are higher in the peak hours than off-peak hours. There are decreased frequencies in bus services in evening hours when most of the businesses are closed for the day and commuters are back at their residences. Since the main research interests of the study was to examine the influences of bus stops and commercial establishments on crime patterns, it may be appropriate to examine crime patterns by selecting the criminal incidents occurred during the peak hours. Another alternative would be to analyze criminal incidents which occurred during the normal business hours or exclude crime between late night and early morning hours when buses are not in the operation. If crime were analyzed using different timeframes, it is possible that different spatial crime patterns or regression results may have been observed. (Ratcliffe, 2002)

There are some practical difficulties in performing temporal pattern analysis. First, for crime types without contact between offender and victim, it is often unknown when exactly the crime occurred. For crimes such as burglary, motor vehicle theft, and theft from motor vehicle, the possible timeframe for the crime occurrence may be longer than several hours. While there are several methods suggested by Ratcliffe (2002) such as dividing the probability of crime occurrence equally for the given time period, this may be a quite cumbersome manual process. Second, most often the ridership information on the buses is not gathered hourly or even daily. Thus, use of ridership information which does not reflect hourly fluctuation may not be useful. Third, ridership levels vary by geographic locations even on the same bus line or by the time of day. Depending on the time of the day, more people will get on and get off at different locations. For instance, buses will pick up passengers in the residential areas in the morning and unload them in business or commercial areas. In the late afternoon hours, this pattern would be reversed. Due to the numerous nature of the bus stops, the ridership information is not often gathered by bus lines by hourly and geographically. Fourth, even more complicating factor is that it is possible that there are different crime patterns not only by daily, but also weekly, monthly, or even seasonally. While not performed in this study, investigation of temporal aspects of crime seems to hold intriguing if not promising prospective. Undoubtedly spatial pattern analysis combined with temporal pattern analysis will shed more light on the influences of bus stops and commercial establishments on criminal offences.

1.3. Questions on Reliability of the YellowBook Data

The present research study employed the Verizon's YellowBook data available for purchase. The businesses can advertise with Verizon's YellowBook services for free (appears without background highlighted on the screen) or pay anywhere from $22 up to $100 a month (appears in background highlighted in different colors and in several mediums). While the coverage (i.e., the estimated percentage of businesses subscribing to YellowBook services) of the businesses is high, this means that there are different needs and demands depending on the business types and sizes. Incidentally, my question regarding the coverage was not answered, and this appears to be due to lack of knowledge rather than a need for secrecy.

One of the remaining questions on using the YellowBook data was whether the YellowBook data would be a good measure of commercial establishments. One way to assess the reliability and representativeness of the YellowBook data was to compare them with the government license data on the liquor license. To assess the YellowBook data's reliability, the Alcoholic Beverage Control (ABC) license data in Newark were purchased from Department of Law and Public Safety, Division of Alcoholic Beverage Control in May 2008. The NJ State ABC data showed that there were total 390 valid liquor licenses in Newark. Very few of them were club licenses (n=22) and over 70 percent (n=275) of them were licensed to bars or restaurants. Only 23.8 percent (n=93) were liquor store licenses. The ABC data were geocoded using the same process as the Newark City Police crime data. Over 98 percent of ABC data were successfully geocoded. Table 60 shows the geocoding result of ABC license data in Newark.

Table 60. Geocoding results of ABC license types in Newark

License type	Matched(%)	Unmatched	Total
Club	21 (95.5)	1	22
Plenary retail with broad package	270 (98.2)	5	27
Liquor store	93(100.0)	0	93
Total	384 (98.5)	6	390

To assess the accuracy of the YellowBook data as a measure of commercial activities, the comparisons of this data with the government data were carried out. The underlying assumption was that the government data were more accurate therefore any discrepancies between the government and the YellowBook data would beinterpreted as inadequacies of the YellowBook data.

The government data came from two sources: the ABC data from NJ State government, and land parcel data from Newark municipality. Among the parcel data, the parcel type that was also listed in the YellowBook database was chosen: religious organization. However, these data were gathered and completed in different timeframe. For the ABC data, its completion date was May 2008, and for land parcel data between 2006 and 2007. In addition, the YellowBook data were compiled in October 2007. Since the data were gathered at different time frames, a comparison at point level may not be necessary or appropriate.

The first problem in comparing the YellowBook data with government data was the differences in the total number of cases. The ABC data listed 93 liquor store licenses while the YellowBook data listed 65. Also, for the religious organizations, the YellowBook data listed only 452 while the city records showed 996 for 2007.

The comparison of the government records with the YellowBook data are summarized in Table 61. The first column (matched-none exist) contains number of areal units where both the YellowBook data and government records show none of the chosen category existed. The next column (matched-exist) shows number of areal units where both data sources showed that the facility in our interest existed. The other two columns are self-explanatory.

For liquor store licenses, 97.7 percent of the areal unit provided the same information on the YellowBook data and the ABC data. To find out whether this was a good match, total number of liquor stores was considered. In Newark, there were 93 liquor stores according to the ABC data, and 65 according to the YellowBook data. Even if we consider none of the records from both dataset matches, and no areal unit has more than one liquor stores, only 158 out of 2,602 areal units will have a liquor store in its unit, yielding 94 percent of areal units providing the same information; none-exist. Also while 51 areal units have liquor store according to both datasets, 62 areal units have conflicting information. For this reason, there seems to be poor a matching rate between the YellowBook data and the ABC data.

The comparison of religious organization also showed a poor matching rate; when counting the areal units with religious organizations reported, only 59 areal units displayed the same information whereas 454 areal units displayed conflicting information. About 80 percent of the areal unit did not house religious organization. However, it was not clear whether religious organizations would aggressively advertise their locations. Maybe these religious organizations may not feel a need to advertise their locations in YellowBook service since they mainly rely on their neighbors in close proximity.[10]

Table 61. Comparison of the YellowBook and government data in Newark (n=2,602)

	Matched-none exist	Matched-exist	Exist in government data only	Exist in YellowBook only
Liquor store	2,489 (95.7)	51 (2.0)	46 (1.8)	16 (0.6)
Religious organization	2,089 (80.3)	59 (2.3)	429(16.5)	25(10.0)

Unfortunately, the comparison of the YellowBook data and government data did not help resolve the concerns rose. One possibility in obtaining a better measure of commercial establishment than relying on the YellowBook data is to use case-control design particularly for the high crime areas. The case-control design would require an environmental survey of the areas and comparing the results with the YellowBook data to assess the reliability and accuracy of the datasets. Due to the time constraint, this step was not taken for the present research study.

1.4. Crime Classification Method: Is there a need to further separate them?

It is noted that the current crime taxonomy based on legal definition may not serve the best interests of crime analysts (Felson, 2006). Under the current crime

[10] Due to the relatively high property tax in NJ, there is a suggestion that many of these registered religious organizations may not be indeed a place of worship. Rather some of the religious organizations may be only in name without performing services to obtain exemption of property taxes. However, it is important to clarify that there is no evidence provided in this study that this happens in Newark.

classification, one crime type may include many diverse behaviors. An initial thought of the study design was to differentiate crimes within the same categories to examine patterns of sub-crime types for robbery and burglary. This reasoning was based on the premises that crimes against different targets will display different spatial dynamics.

This plan was abandoned for different reasons. For robbery, it was not clear what kind of criteria needs to be used to separate the offences. There are several ways to classify robbery; it can be differentiated using the type of premises where it occurred; commercial or street. Or it can be differentiated by using the modus operandi; stealth for pick-pocket, confrontation, or snatch for speed. Or it can be differentiated using the targets; cash, jewelry, drugs, and motor vehicles. To separate robbery for spatial pattern analysis, factors influencing spatial patterns need to be used in order to be a meaningful classification method. For the purpose of this research, it was not clear what factors within the robbery category influenced spatial patterns.

For burglary, the category was already classified into three types: residential, quasi-public, and commercial. Due to the land use designation which restricts where commercial establishments or residences can be located, this classification method is not only relevant but also desirable for spatial pattern analysis. However, the majority of the burglary was classified as residential. For this reason, the results of data analysis using the burglary as a whole and the residential burglary were virtually identical. For commercial and quasi-public categories, simply there were not enough numbers of cases to carry out separate data analysis. Nevertheless, it is possible to separate burglary into several categories using more than one year of data to have a large enough number of cases.

Another problematic crime classification in this study was aggravated assault. The category of aggravated assault included shooting and beating between strangers as well as intimates. Since there was no information on offender and victim relationship, it was difficult to know the proportion of intimate partner violence in aggravated assault which may be significantly influenced by land use patterns.

1.5. Spatial Effects: Spatial Dependence and Heteroskedacity

The present research study discussed two aspects of spatial effects; spatial dependence and spatial heterogeneity. Mainly dealt with in this research is the aspect of spatial dependence. The spatial dependence and spatial interaction are multidirectional. This means that the actions are often dependent on the structure and position within the society, interaction with other individuals, and the attributes of the individuals responsible for the actions (Ward and Gleditsch, 2008). For this reason, the predictors examined in this research are also dependent on other variables. For instance, it was suggested that where the buses began and ended their routes will influence bus stop related crimes. Also the behaviors such as robbery and assault may not adhere to the administrative or spatial boundaries used. This will cause measurement errors that may manifest in the form of spatial dependence, spatial heterogeneity, and heteroskedacity. If spatial dependence transcends the study area, there may be methodological problems affecting statistical inferences in spatial process which is called edge effects (Anselin, 1988).

In the context of the present research study, it is probable that administrative boundaries of city may not have any geographical impacts on crime patterns. This is particularly true when considering Newark with its eight adjacent neighbors. A portion of Newark includes a protruding region surrounded by Irvington, East Orange, and Maplewood (see Figure 1). Particularly for violence offences, a high level of spatial interactions is assumed in this part of Newark. This is reflected in Operation Ceasefire Initiative which began in May 2005 and concluded in December 2008. The target area of Operation Ceasefire included a two mile radius area in Irvington and Newark. This was due to the conclusion that violence offences such as shooting incidents were problematic in this area. Therefore, it is not be surprising that the OLS diagnostics identified the SEM as the structure of spatial dependence from the OLS regression diagnostics tests. In the future, a way to address possible edge effect may need to be considered by either including areas with high spatial interaction or excluding a protruding portion of the city from the analysis. A stringent adherence to administrative boundaries may not serve the best research design when examining spatial patterns.

1.6. Areal Units and Measures of Predictors

The diagnostics in GeoDa suggested possible model misspecifications for all crime types examined. There are several causes for this finding. First, the use of 1,600 feet distance based spatial lag may not be suitable. When calculating spatial lag in GeoDa, the minimum distance necessary to ensure every unit to have at least one neighbor was calculated to be 1,555 feet. This rather long distance was due to the fact that there were several big areal units despite the fact that the majority of the areal units in the study area were quite small. These big areal units were often located nearby the excluded areas or included a park or cemetery. These big areal spatial units often did not include long street segments, and may in fact have very little to no crime. For this reason, if this minimum distance was ignored, and the spatial lags were calculated using shorter distance, different regression results may have been obtained.

Indeed, there is a good reason to suspect that the use of 1,600 feet may not be appropriate in the context of bus stops. Bus stops are often located in every other street block, and they may be as little as 500 to 800 feet away from each other. For this reason, within the 1,600 feet distance, there may be two to three bus stops belonging to the same bus line. It is also possible that land use patterns may change within the distance of 1,600 feet. Considering the fact that land use zoning districts were often small within the city, it is possible that 1,600 feet distance may have included non-homogeneous areas possibly with different zoning district codes. Therefore, in the future use of shorter distances in constructing weight matrix may prove to be useful.

Another possible reason for model misspecification is the omission of other important predictors in the models. It is possible that there are other important predictors influencing crime opportunities other than the bus stops, commercial establishments, and land use information. These variables may include population density or temporal variables. Alternatively, addition of social disorganization variables or use of more detailed business category may further enhance our understanding of crime opportunities. In addition, it may be useful to think about how to better measure or reflect the predictors in the model estimations. As with the dependent variables, the spatial dependence on predictors may be reflected as heteroskedacity. For this reason, it may be necessary to use different measures of independent variable in the models. For instance, it may be useful or even appropriate to use spatially lagged predictors to account for spatial dependence among the predictors.

Another possible reason for model misspecification is use of inappropriate areal units. The areal unit used for the study was created using street intersection. Since the size of street blocks varied greatly, some areal units were quite large despite the fact that the majority of areal units were quite small, approximately a size of street blocks. These large areal units often included parks or cemeteries which were not likely to have geocoded criminal incidents. For this reason, in the future, use of more uniform areal unit or shorter distances in constructing weight matrix may prove to be more useful.

2. Discussion and Conclusion

2.1. Ambiguities in preventing crime at bus stops

The effort to prevent crimes at light rail or train stations can be clearly targeted at mass transit facilities. However, this is not the case for crimes at the bus stop environments. The ambiguities in crime prevention efforts at bus stops are related to the cause, responsible agencies, and scope of the crime prevention initiatives. For the cause, crime in the vicinities of bus stops brings a doubt as to whether and how much the existence of bus stops contributes to criminal incidents.

For the responsible agencies, there are disagreements regarding which agencies are responsible for keeping the order around bus stop environments. Depending on who you talk to and the stakeholders, it could be the transit police, city police, or even the transit company if privately owned or operated (Kooi, 2007). Even if it can be agreed that bus services and stops play a significant role in causing crime, and the party responsible for preventing crime is identified, the scope of such initiatives are not easily defined.

The scope of crime prevention initiatives are concerned with geographic and temporal boundaries. For crimes in the vicinity of transit stops and stations, it is difficult to know how far the crime prevention efforts needs to be extended (Newton et al., 2004). For instance, it is suggested that the security of mass transit passengers needs to be evaluated from when passengers begin their travel till the very end when they arrive at their destinations (DETR, 1999). Therefore, crime prevention efforts in whole-journey approach could easily mean city-wide crime prevention efforts. Perhaps due to this reason, there have been very few crime prevention efforts to prevent crime in the vicinity of mass transit. A recent crime prevention initiative in and around mass transit showed promising and encouraging results. The four week crime prevention initiative Operation Bream targeted one high crime bus route in Merseyside, UK (Newton et al., 2004). This safer travel initiative employed high visibility intensive policing targeting crimes both on and around the targeted bus route. The evaluation of the initiative concluded that the operation appeared to have reduced calls for police service yet increased the arrests in the target area. In addition, diffusion of benefits was reported as well; for assault, the crime prevention effects were extended up to about 1,300 feet from the bus route and lasted even after the operation was concluded.

There are certain types of built environments that incite fear of crime to the public. These built environments include enclosed places such as mass transit stations or bus stops located in isolated and deserted areas (Loukaitou-Sideris and Eck, 2007). It is often argued that the loss of ridership may lead to the migration of those who can afford other means of travel rendering the transit systems as the domains for those without alternatives (Nelson, 1997). This population group, often called transit captives, tends to be women and people with low income (Maier and Boyd, 1998). In turn, fear and decreased desirability of the services can lower ridership and lead to a further deterioration of the system (Nelson, 1997).

In the whole-journey approach, the factors such as bus stop locations, the passages to the bus stops, bus stop maintenance, and the characteristics of the surrounding areas all affect the perception of safety (Tucker, 2003). For this reason, it is argued that crime prevention initiatives aimed at reducing crime and increasing ridership should consider the transit vehicles, bus stops, surrounding areas, and the passage to the bus stops (Vogel and Pettinari, 2002).

In relation to the comfort and quality of bus travel, attention should be focused on the traveling environments, and the transit vehicles. Several studies suggest use of more aesthetic vehicles with big clear windows, appropriate lighting, and brightly colored bus seats to improve the quality of bus travel experiences (Lusk, 2001).

Since the majority of the crime in relation to bus services occur at the bus stop, the crime prevention efforts needs to focus on the bus stops themselves. Some of the crime prevention measures that will likely to increase both real and perceived security levels at the bus stop is CCTV surveillance (Crime Concern, 2004) (Tucker, 2003). Installation of CCTV in high crime bus stops would increase the feelings of safety as well as actual safety levels. In addition, unobstructed view of the bus stops from the roads and other businesses could actually increase the feelings of security as well as actual risks. Several studies suggest use of see-through bus shelters to reduce fear of crime (Lusk, 2001) or relocation of a bus stop if located in an isolated area (Tucker, 2003). The logic behind placing a bus stop in an isolated place may never be understood, yet it is not clear whether it can be moved. The bus stops are often located in every other street block; therefore, moving a bus stop even for a short distance would create very closely located bus stops or no bus stops for a relatively long distance. In addition, it is important to maintain the physical environments and social civilities in the area. Considering the high number of bus stops, this may be an expensive task to undertake. However, it was shown that programs relying on volunteers or residents to maintain the bus stops are considered to be effective and inexpensive methods. For instance, Adopt-A-Stop program involving the residents to maintain the bus stops costs very little while keeping the bus stop environments clean and pleasant (Tucker, 2003).

Despite the fact that the majority of the reported crime occurred at the bus stop rather than on the vehicles (Loukaitou-Sideris, 1999), people usually do not experience fear while waiting for and traveling on the bus during the daytime. The results of household-based survey showed that less than five percent of the respondents reported feeling unsafe while waiting at a bus stop during the day time (Crime Concern, 2004). The fear levels increased dramatically after the dark; 49 percent of women and 32 percent of men reported feeling unsafe waiting at bus stops after the dark. The comparable figures for traveling on the bus after the dark were 40 percent for women and 18 percent for men. For this reason, installation of street lights at and around bus stops should be given a consideration to enhance the feelings of personal security (Crime Concern, 2004).

It is often mentioned that the safe pedestrian environments will increase the ridership level of public transportation (DETR, 1999). Making the passage to the bus stops safer and enjoyable may involve widening the sidewalks, trimming the tree branches, cleaning up the streets promptly, or installing and maintaining the street lights. While valuable in their own rights these tasks often belong to the city, not to the police department or the transit companies. In addition, the impacts of these actions may be very subtle or require continued efforts. For this reason, coordinating these initiatives for crime prevention purposes may be difficult. The lack of clear and immediate causal impact may be an obstacle in persuading the stakeholders to participate in the city beautification projects.

Lastly, there are policy implications based on the research findings in relation to crime preventions. One of the most significant variables in relation to crime opportunities was spatial lag, the value of one’s neighbors. In fact, the influence of spatial lag was the most salient and strong regardless of crime types examined or regression methods used.

The clustering of crime hot spots and cold spots can be observed from maps of LISA as well as density. This suggests that crime prevention methods based on geography will probably lead to bigger impacts than commonly thought. It is likely that decreases in crime opportunities in one area will likely to spread, leading to diffusion of benefits.

2.2. Concluding Remarks

While the environmental focus on bus stop related crime is legitimate and desirable, the first step in addressing the fear of crime in relation to mass transit is to find out actual crime patterns. The collection of data and analysis is a foundation of effective crime prevention strategies; the research findings can be used to direct the policing resources and disseminate information to accurately portray the risks of crime at the bus stops (Tucker, 2003).

Expanding mass transit usually faces two very different reactions from the stakeholders in the areas. When the expansion of mass transit efforts are placed in affluent areas, city planners often face resident oppositions stemming from the fear that the new mass transit may be used by criminals to expand their criminal activity spaces. When mass transit expansion efforts are placed in economically depressed inner-city areas, mass transit receives hopes of revitalizing the community by increasing transit revenues, employment rates, social interaction, and environmental benefits (Loukaitou-Sideris, 2000; Loukaitou-Sideris and Banerjee, 2000). Considering the fact that transportation infrastructure is one of the strongest factors shaping urban structure (Button et al., 2004), it is not surprising that mass transit has received more than its fair share of hopes and concerns.

However, the examination of a project which aimed at improving inner-city by using mass transit as an engine (a TOD project) did not live up to its hope; despite the high level of ridership, the LA Blue Line had not produced any economic growth in the intended areas for nine years (Loukaitou-Sideris, 2000). The authors attributed this disappointing result to the absence of antecedents for neighborhood developments. Some of these identified missing antecedents were growing regional economy and positive physical characteristics surrounding the transit station. Also another identified antecedent was resident stability; the areas where the Blue Line passes through were characterized as being populated by ethnic minorities and immigrants with high turn-over rates. Furthermore, the existence of mass transit did not improve employment opportunities when the levels of education and car ownership were controlled (Cervero et al., 2002).

The existing literature on crime also shows that mass transit does not create new crime hot spots, change crime spatial patterns, or lengthen offenders' journey to crime. It also shows that impacts of mass transit are different depending on the area characteristics where it is located (Poister, 1996). In one study, the opening of new light rail station in the commercial area was followed by the increases in larceny and vagrancy. However, the new light rail station in residential area did not show the same increases that can be attributed to the new light rail system. It is hypothesized that due to many available legal and illegal activities around the mass transit, the commercial establishments in close proximity to mass transit may suffer from higher number of criminal incidents (Block and Block, 1995). For instance, in one study the high crime liquor-licensed places were about three times more likely to be located within one block radius of transit stations (Block and Block, 1995).

The idea that mass transit will have noticeable if not measurable impacts on crimes seems to parallel to TODs which aims to bring about economic growth in the areas adjacent to mass transit. In parallel to the ideas of TODs, it seems that there need to be antecedents for relationships between mass transit and crime to develop. In some cases, this could be volume of ambient population or number and type of commercial establishments (Block and Block, 2000; Liggett et al., 2001), poor visibility, lack of natural surveillance such as vacant or abandoned lots (Loukaitou-Sideris, 1999), or neighborhood characteristics (Loukaitou-Sideris *et al.*, 2002; Newton, 2004). In this research study, the predictors related to opportunity theories were used to examine the factors influencing crime occurrences in Newark. Not examined however were the variables related to social disorganization.

In this research study, it was shown that both bus stops and commercial establishments influence crime opportunities in the area. While there are many choices in selecting commercial establishments, this is not the case for mass transit services. The use of bus service is most often out of necessity rather than the pleasure, consequently, users of mass transit have no choice but to go back to the same locations and services repeatedly regardless how unpleasant it might be. For this reason, many bus riders may tolerate the perceived danger and hassles when using public transportation.

The first research question was related to the relationship between number of bus stops and criminal offences. For this reason, Model 1 using bus stop as the predictor was run for all five types of crime. The data analysis result showed that *bstops* (number of bus stops) was almost always statistically significant with the increased number of criminal offences for all crime types. Except for burglary, *bstops* was always statistically significant regardless of the regression methods used.

The second research question was related to number of commercial establishments and criminal offences. To answer this question, the second model was built using seven types of commercial establishments as predictors. In general, retail business categories seem to have more influence on crime opportunities than service categories examined. For robbery, all three types of retail categories were statistically significant to increased opportunities. Among the commercial establishments, food stores displayed consistent influences in increasing crime opportunities than other business categories examined.

The regression models using both bus stops and commercial establishments showed very little interaction effects among the predictors. Since there were very little interaction effects between bus stops and crime, it can be concluded that crime at bus stops were probably not due to commercial activities in the area. To further examine what may cause high criminal incidents in the high crime areas will require environmental observations of the areas.

Based on the research findings of the study, the future research studies need to focus on two aspects. First, environments of high and low crime bus stops areas need to be compared to find out the factors influencing crime opportunities not measured in this research study. These factors may include social and physical incivilities including social disorganization variables. Second, rather than focusing on the street blocks where bus stops are located in, the influences of independent variables located within two to three blocks could be taken into account. For instance, one of the reasons that there were very little interactions between bus stops and commercial establishments may be due to the fact that the areal units in the high activity areas were too small to include predictors located in close proximity but not in the same areal unit. If the weighted measures of predictors were used (i.e., the predictors around the areal units were incorporated into the estimations), it is possible that high levels of interaction could have been observed.

In conclusion, the present research study provided three main findings. First, the existence of bus stops generally increases crime occurrences in the area. Therefore, the concern for personal security in bus stop environments seems legitimate. This means that crime prevention efforts need to be directed at bus stops environments. Second, existence of commercial establishments also seems to increase crime opportunities. In addition, there were very little interaction effects between bus stops and commercial establishments. From this finding, it is inferred that the levels of ambient population related to commercial activities influence crime occurrences in the areas. Since there was no uniform impact of commercial establishments other than food store on crime, it is suggested to carry out custom-tailored data analysis to enhance our understanding of crime patterns in relation to commercial establishments and bus stops. Such study should incorporate environmental survey of the areas to better understand the context of bus stops in relation to crimes. Third, the regression results showed that the location was the most salient factor in explaining spatial patterns of crime. For this reason, it seems promising to devise crime prevention initiative centered on geography to yield the best results.

Often encountered concerns with crime prevention efforts are fears of displacement. However, three theories employed in this study suggest that displacement is not likely since both offenders and victims are constrained by their environments and social constraints. In fact, review of research studies show that diffusion of benefits is often greater than displacement even if displacement of crime occurs (Eck, 1997; Brown, 2004; Clarke and Goldstein, 2004). In conclusion, the future research should focus on environments in relation to temporal aspects of crime. Also crime prevention initiatives need to be devised based on geography to have greater impacts on crime reductions.

Reference List

Andersen, M. A. (2006). "A Spatial Analysis of Crime in Vancouver, British Columbia: A Synthesis of Social Disorganization and Routine Activity Theory." Canadian Geographer **50**(4): pp. 487-502.

Angel, S. (1968). Discouraging Crime Through City Planning. Berkeley, CA, Institute of Urban Crime Through City Planning, University of California**:** 37 pages.

Anselin, L. (1988). Spatial Econometrics: Methods and Models. Dordrecht, The Netherlands, Kluwer Academic Publishers.

Anselin, L. (1998). Spatial Econometrics: Methods and Models. Dordrecht, The Netherlands, Kluwer Academic Publishers.

Anselin, L. (2003). GeoDa 0.9 User's Guide. IL, Spaital Analysis Laboratory, Department of Geography, University of Illinois, Urbana-Champaign

Center for Spatially Intergrated Social Science**:** 226 pg.

Anselin, L. (2005). Exploring Spatial Data with GeoDA: A workbook. IL, Spaital Analysis Laboratory, Department of Geography, University of Illinois, Urbana-Champaign

Center for Spatially Intergrated Social Science**:** 226 pg.

Anselin, L., J. Cohen, D. Cook, W. Gorr and G. Tita (2000). "Spatial Analysis of Crime." Criminal Justice **4**: pp.213-262.

Bailey, L. (2004). Aging Americans: Stranded Without Options. Washington, DC, Surface Transportation Policy Project, American Public Transportation Association.

Barclay, P., J. Buckley, P. J. Brantingham, P. L. Brantingham and T. Whinn-Yates (1996). Preventing Auto Theft in Suburban Vancouver Commuter Lots: Effects of a Bike Patrol. Preventing Mass Transit Crime. Vol. 6. R. V. Clarke. Monsey, NY, Criminal Justice Press.

Belanger, M. E. (1999). Criminal Mobility and Public Transport: The Case of the New York City Subway. School of Criminal Justice. Newark, New Jersey, Rutgers University.

Beller, A., S. Garelik and S. Cooper (1980). "Sex Crimes in the Subway." Criminology **18**(1): pp. 35-52.

Benson, M. L., G. L. Fox, A. DeMaris and J. V. Wyk (2003). "Neighborhood Disadvantage and Violence Against Women in Intimate Relationships." Journal of Quantitative Criminology **19**(3).

Block, R. and C. R. Block (2000). The Bronx and Chicago: Street Robbery in the Environs of Rapid Transit Stations. In Analyzing Crime Patterns: Frontiers in

Practice. V. Goldsmith, P. G. McGuire, J. H. Mollenkopf and T. A. Ross. Thousand Oaks, CA, Sage Publications.

Block, R. and R. Block (1995). Space, Place, and Crime: Hot Spot Areas and Hot Places of Liquor-Related Crime. Crime and Place. J. E. Eck, Willow Tree Press.

Block, R. and S. Davis (1996). The Environs of Rapid Transit Stations: A Focus for Street Crime or Just Another Risky Place? Preventing Mass Transit Crime. R. V. Clarke. Monsey, NY, Criminal Justice Press.

Boyd, S. J., K. M. Armstrong, L. J. Fang, D. R. Medoff, L. B. Dixon and D. A. Gorelick (2007). "Use of "Microecologic Technique" to Study Crime Around Substance Abuse Treatment Centers." Social Science Computer Review **25**(2): pp. 163-173.

Brantingham, P. J. and P. J. Brantingham (1993). "Paths, Nodes and Edges: Considerations on the Complexity and Crime and the Physical Environment." Journal of Environmental Psychology **13**: pp.3-28.

Brantingham, P. J. and P. L. Brantingham (1984). Patterns in Crime. US, Macmillan.

Brantingham, P. J. and P. L. Brantingham (1997). Understanding and Controlling Crime and Fear of Crime: Conflicts and Trade-Offs in Crime Prevention Planning. Crime Prevention at a Crossroads. S. P. Lab, Anderson Publishing Co.

Brantingham, P. J. and P. L. Brantingham (1998). "Environmental Criminology: From Theory to Urban Planning Practice." Studies on Crime and Crime Prevention **7**(1): pp. 31-60.

Brantingham, P. J., P. L. Brantingham and P. S. Wong (1991). "How Public Transit Feeds Private Crime: Notes on the Vancouver 'Skytrain' Experience." Security Journal **2**(2): pp. 91-95.

Brantingham, P. L. and P. J. Brantingham (1995). "Criminality of Place: Crime Generators and Crime Attractors." European Journal on Criminal Policy and Research **3**(3): pp. 5-26.

Brown, R. (2004). The Effectiveness of Electronic Immobilization: Changing Patterns of Temporary and Permanent Vehicle Theft. Understanding and Preventing Car Theft. R. V. Clarke. NY, Criminal Justice Press. **17**.

Budd, T. (2001). Violence at Work: New Findings from the 2000 British Crime Survey. London, Joint publication with Home Office and Health and Safety Executive.

Burrows, J., S. Anderson, J. Bamfield, M. Hopkins and D. Ingram (1999). Crimes Against Business in Scotland, The Scottish Executive Central Research Unit.

Button, K. J., P. Nijkamp and P. Rietveld (2004). Land-use, Transportation and Urban Development. Urban Dynamics and Growth: Advances in Urban Economics. R. Capello and P. Nijkamp. CA, USA, Elsevier Inc.**:** pp. 153 - 179.

Cahill, M. and G. Mulligan (2007). "Using Geographically Weighted Regression to Explore Local Crime Patterns." Social Science Computer Review **25**(2): pp. 174-193.

Canter, D. and P. Larkin (1993). "The Environmental Range of Serial Rapists." Journal of Environmental Psychology **13**: pp. 63-69.

Cervero, R., O. Sandoval and J. Landis (2002). "Transportation as a Stimulus of Welfar-to-Work: Private versus Public Mobility." Journal of Planning Education and Research **22**(50): pp. 50-63.

City-of-Newark, D. o. E. H. D. and I. Philips Preiss Shapiro Associates (2006). Land Use Element of the Master Plan for the City of Newark. City of Newark, NJ, City of Newark, Department of Economic & Housing Development.

Clarke, R. V. (1999). Hot Products: understanding, anticipating and reducing demand for stolen goods, Home Office.

Clarke, R. V., M. E. Belanger and J. A. Eastman (1996). Where Angel Fears to Tread: A Test in the New York City Subway of the Robbery/Density Hypothesis. Preventing Mass Transit Crime. Vol. 6. R. V. Clarke. Monsey, NY, Criminal Justice Press.

Clarke, R. V. and H. Goldstein (2003). Theft from Cars in Center-City Parking Facilities: A Case Study in Implementing Problem-Oriented Policing. Problem_Oriented Policing: From Innovation to Mainstream. J. Knutsson. NY, Criminal Justice Press. **15**.

Clarke, R. V. and H. Goldstein (2004). Theft from Cars in Center-City Parking Facilities: A Case Study in Implementing Problem-Oriented Policing. Problem_Oriented Policing: From Innovation to Mainstream. J. Knutsson. NY, Criminal Justice Press. **15**.

Cohen, L. E. and M. Felson (1979). "Social Change and Crime Rate Trends: A Routine Activity Approach." American Sociological Review **44**(1): pp.588-608.

Cornish, D. and R. V. Clarke (1986). Introduction. The Reasoning Criminal: Rational Choice Perspective on Offending. R. V. Clarke. NY, Springer-Verlag Inc.

Cornish, D. B. and R. V. Clarke (2003). Opportunities, Precipitators and Criminal Decisions: A Reply to Wortley's Critique of Situational Crime Prevention. Theory for Practice in Situational Crime Prevention. M. J. Smith and D. B. Cornish. NY, Willow Tree Press. **Crime Prevention Studies. Vol. 16:** pp. 41-96.

Crime-Concern (2004). People's Perceptions of Personal Security and Their Concerns About Crime on Public Transport: Literature Review. London, UK, Crime Concern, Department for Transport.

Cunradi, C. B., R. Caetano, C. Clark and J. Schafer (2000). "Neighbourhood Poverty as a Predictor of Spousal Assault Among White, Black, and Hispanic Couples in the

United States: A Multilevel Analysis." Analysis of Epidemiology **10**(5): pp. 297-308.

Department for Transport (2002). A Bulletin of Public Transport Statistics: Great Britain 2002 Edition. London, UK, Department for Public Transport.

Department for Transport (2006). Public Transport Statistics Bulletin GB:2006 Edition. London, UK, Department for Public Transport.

Desroches, F. J. (2002). Force and Fear: Robbery in Canada. Toronto, Canada, Scholars' Press Inc.

DETR (1999). Personal Security Issues in Pedestrian Journeys. London, Department of the Environment, Transport and the Regions.

Eck, J. E. (1997). Preventing Crime at Places. Preventing Crime: What Works, What Doesn't, What's Promising. L. W. Sherman, D. Gottfredson, D. MacKenzieet al, National Institute of Justice.

Eck, J. E., R. V. Clarke and R. T. Guerette (2007). Risky Facilities: Crime Concentration in Homogeneous Sets of Establishments and Facilities. Imagination for Crime Prevention: Essays in Honour of Ken Pease. G. Farrell, K. J. Bowers, S. D. Johnson and M. Townsley. Devon, UK, Willan Publishing. **21**.

Eck, J. E. and D. Weisburd (1995). Crime Places in Crime Theory. Crime and Place. J. E. Eck and D. Weisburd. Monsey, NY, Willow Tree Press. **4:** pp.1-33.

Eck, J. E. and D. Weisburd (1995). Crime Places in Crime Theory. Crime and Place. Crime Prevention Studies. D. Weisburd, Willow Tree Press. **4:** pp.1-33.

Farrell, G. and K. Pease (2001). Editors' Introduction: Why Repeat Victimization Matters. Repeat Victimization: Crime Prevention Studies. K. Pease, Criminal Justice Press. **12:** pp.1-4.

Farrell, G., C. Phillips and K. Pease (1995). "Like Taking Candy: Why Does Repeat Victimization Occur?" British Journal of Criminology **35**(3): pp.384-399.

Farrell, G. and W. Sousa (2001). Repeat Victimization and Hot Spots: The Overlap and Its Implications for Crime Control and Problem-oriented Policing. Repeat Victimization: Crime Prevention Studies. K. Please, Criminal Justice Press. **12**.

Felson, M. (1995). Those Who Discourage Crime. Crime and Place. J. E. Eck and D. Weisburd. Monsey, NY, Willow Tree Press. **4**.

Felson, M. (2006). Crime and Nature. CA, Sage Publications Inc.

Felson, M., M. E. Belanger, G. M. Bichler, C. D. Bruzinski, G. S. Campbell, C. L. Fried, K. C. Grofik, I. S. Mazur, A. B. O'Regan, P. J. Sweeney, A. L. Ullman and L. M. Williams (1996). Redesigning Hell: Preventing Crime and Disorder at the Port Authority Bus Terminal. Preventing Mass Transit Crime. Vol. 6. R. V. Clarke. Monsey, NY, Criminal Justice Press.

Felson, M. and R. V. Clarke (1998). Opportunity Makes the Thief: Practical Theory for Crime Prevention. Police Research Series, Paper #98. London, UK, Research, Development and Statistics Directorate, U.K. Home Office.

Felson, M. and E. Poulsen (2003). "Simple Indicators of Crime by Time of Day." International Journal of Forecasting **19**: pp595-601.

Graaff, T. d., R. J. G. M. Florax and P. Nijkamp (2001). "A General Misspecificaiton Test for Spatial Regression Models: Dependence, Heterogeneity, and Nonlinearity." Journal of Regional Science **4**(2): pp. 255-276.

Hilbe, J. M. (2007). Negative Binominal Regression. Cambridge, UK, Cambridge University Press.

Hunter, R. D. (1999). "Convenience Store Robbery Revisited: A Review of Prevention Results." Journal of Security Administration **22**(1): pp. 1-13.

Kenney, D. J. (1987). Crime, Fear, and the New York City Subways: The Role of Citizen Action. NY, Praeger Publishers.

Kent, J., M. Leither and A. Curtis (2005). "Evaluating the usefulness of functional distance measures when calibrating journey-to-crime distance decay functions." Computers, Environment and Urban Systems **In Press**.

Kooi, B. R. (2007). Policing Public Transportation: An Environmental and Procedural Evaluations of Bus Stops. NY, LFB Scholarly Publishing LLC.

Laukkanen, M. and P. Santtila (2005). "Predicting the residential location of a serial commerical robber." Forensic Science International **in Press**.

LaVigne, N. G. (1996). Safe Transport: Security by Design on the Washington Metro. Preventing Mass Transit Crime. Vol. 6. R. V. Clarke. Monsey, NY, Criminal Justice Press.

LeSage, J. P. (1998). Spatial Econometrics.

Levine, N. and Associates (2004). CrimeStat III: A Spatial Statistics Program for Analysis of Crime Incident Locations. Washington, DC, The National Institute of Justice.

Levine, N. and M. Wachs (1986). "Bus Crime in Los Angeles: I- Measuring The Incidence." Transportation Research **20**(4): pp. 273-284.

Liggett, R., A. Loukaitou-Sideris and H. Iseki (2001). "Bus stop-environment connection: Do characteristics of the built environment correlate with bus stop crime?" Transportation Research Record(No. 1760): pp. 20 - 27.

Liggett, R., A. Loukaitou-Sideris and H. Iseki (2003). "Journey to Crime: Assessing the Effects of a Light Rail Line on Crime in the Neighborhoods." Journal of Public Transportation **6**(3): pp. 85-115.

Lockwood, D. (2007). "Mapping Crime in Savannah: Social Disadvantage, Land Use, and Violent Crimes Reported to the Police." Social Science Computer Review **25**(2): pp. 194-209.

Long, J. S. and J. Freese (2006). Regression Models for Categorical Dependent Variables Using Stata. College Station, Texas, A Stata Press Publication.

Loukaitou-Sideris, A. (1999). "Hot Spots of Bus Stop Crime: The Importance of Environmental Attributes." Journal of American Planning Association **65**(4): pp. 395 - 411.

Loukaitou-Sideris, A. (2000). "Transit-Oriented Development in the Inner City: A Delphi Survey." Journal of Public Transportation **3**(2): pp. 75 - 98.

Loukaitou-Sideris, A. and T. Banerjee (2000). "Blue Line Blues: Why the Vision of Transit Village May Not Materialize Despite Impressive Growth in Transit Ridership." Journal of Urban Design **5**(2): pp. 101- 125.

Loukaitou-Sideris, A. and J. E. Eck (2007). "Crime Prevention and Active Living." American Journal of Health Promotion **21**(4. Supp.): pp. 380-389.

Loukaitou-Sideris, A., R. Liggett and H. Iseki (2002). "The Geography of Transit Crime: Documentation and Evaluation of Crime Incidence on and Around the Green Line Stations in Los Angeles." Journal of Planning Education and Research **22**: pp. 131-151.

Loukaitou-Sideris, A., R. Liggett, H. Iseki and W. Thurlow (2001). "Measuring the Effects of Built Environment on Bus Stop Crime." Environment and Planning B: Planning and Design **28**(2): pp. 255-280.

Lu, Y. (2003). "Getting Away with the Stolen Vehicle: An Investigation of Journey-after-Crime." The Professional Geographer **55**(4): pp. 422-433.

Lusk, A. (2001). Bus and Bus Stop Designs Related to Perceptions of Crime. Washington, D.C., Federal Transit Administration, Office of Research, Demonstration and Innovation.

Maier, P. and A. Boyd (1998). "Security Issues Affecting Women in the Transit Environment." Transit Policing **8**(2): pp. 1, 11-12.

Maxwell, C. D., J. H. Garner and J. A. Fagan (2002). "The Preventive Effects of Arrest on Intimate Partner Violence: Research, Policy and Theory." Criminology & Public Policy **2**(1): pp. 51-80.

Mayhew, P. and G. Braun (2004). Parking Lot Security. Understanding and Preventing Car Theft. R. V. Clarke. NY, Criminal Justice Press. **17**.

McQuestion, M. J. (2000). "Endogenous Social Effects on Spousal Assault in Colombia." Social Science Research **32**(2): pp. 335-345.

Miles-Doan, R. (1998). "Violence Between Spouses and Intimates: Does Neighborhood Context Matter?" Social Forces **72**(2): pp. 623-645.

Mitchell, A. (1999). The ESRI Guide to GIS Analysis, Volume 1: Geographic Patterns & Relationships. Redlands, CA, ESRI Press.

Mitchell, A. (2005). The ESRI Guide to GIS Analysis, Volume 2: Spatial Measurements & Statistics. Redlands, CA, ESRI Press.

Moore, M. H. (2002). "The Limits of Social Science in Guiding Policy." Criminology & Public Policy **2**(1): pp.33-42.

Morgan, R. and D. B. Cornish (2006). Introduction: Crime and Disorder on Public Transport. Secure and Tranquil Travel: Preventing Crime and Disorder on Public Transport. M. J. Smith and D. B. Cornish. London, UCL Jill Dando Institute of Crime Science.

Nelson, K. R. (1997). "Policing Mass Transit." FBI Law Enforcement Bulletin **66**(1): pp. 1-7.

Newton, A. D. (2004). Crime and Disorder on Buses: Toward an Evidence Base for Effective Crime Prevention, Ph.D. thesis, University of Liverpool.

Newton, A. D. (2004). "Crime on Public Transport: 'Static' and 'Non-Static' (Moving) Crime Events." Western Criminology Review **5**(3): pp. 25-42.

Newton, A. D., S. D. Johnson and K. J. Bowers (2004). "Crime on Bus Routes: An Evaluation of a Safer Travel Initiative." Policing: An International Journal of Police Strategies & Management **27**(3): pp. 302-319.

Paulsen, D. J. and M. B. Robinson (2004). Spatial Aspects of Crime: Theory and Practice. Boston, Pearson Education, Inc.

Pease, K., Ed. (1998). Repeat Victimisation: Taking Stock. Crime Detection and Prevention Series Paper 90. Home Office Police Research Group.

Plouffe, N. and R. Sampson (2004). Auto Theft and Theft from Autos in Parking Lots in Chula Vista, CA: Crime Analysis for Local and Regional Action. Understanding and Preventing Car Theft. R. V. Clarke. NY, Criminal Justice Press. **17**.

Poister, T. H. (1996). "Transit-Related Crime in Suburban Areas." Journal of Urban Affairs **18**(1): pp. 63-75.

Ratcliffe, J. H. (2002). "Aoristic Signatures and the Spatio-Temporal Analysis of High Volume Crime Patterns." Journal of Quantitative Criminology **18**(1): pp. 23-43.

Ratcliffe, J. H. (2003). Suburb Boundaries and Residential Burglars. Trends & Issues in Crime and Criminal Justice, Australian Institute of Criminology. **No. 246**.

Ratcliffe, J. H. (2004). "The Hotspot Matrix: A Framework for the Spatio-Temporal Targeting of Crime Reduction." Police Practice and Research **5**(1): pp. 05-23.

Ratcliffe, J. H. and T. A. Taniguchi (2008). "Is Crime Higher Around Drug-gang Street Corners? Two Spatial Approaches to the Relationship Between Gang Set Space and Local Crime Levels." Crime Pattern and Analysis **1**(1).

Rengert, G. F. (1997). Auto Theft in Central Philadelphia. Policing for Prevention: Reducing Crime, Public Intoxication and Injury. R. Homel. NY, Criminal Justice Press. **7:** pp.199 -.

Sedelmaier, C. M. (2003). Railroaded: The Effects of a New Public Transportation System upon Local Crime Patterns. Ph.D. dissertation, Rutgers University.

Shaw, C. R. and H. McKay (1942). Juvenile Delinquency and Urban Areas. Chicago, IL., The University of Chicago Press.

Shaw, M. and K. Pease (2000). Research on Repeat Victimisation in Scotland: Final Report, The Scottish Executive Centeral Unit.

Sloan-Howitt, M. and G. L. Kelling (1990). "Subway Graffiti in New York City: "Gettin Up" vs. "Meanin It and Cleanin It"." Security Journal **1**(3): pp. 131-136.

Smith, M. J. and R. V. Clarke (2000). "Crime and Public Transport." Crime and Justice: A Review of Research **27**: pp. 169-233.

State-of-New-Jersey, D. o. S. P. (2007). Uniform Crime Report: State of New Jersey 2006. West Trenton, NJ, Office of Attorney General, State of New Jersey.

Taylor, N. and P. Mayhew (2000 March). "Patterns of Victimisation Among Small Retail Businesses." Trends and Issues in Crime and Criminal Justice **No. 221**.

Tilley, N., J. Smith, S. Finer, R. Erol, C. Charles and J. Dobby (2004). Problem-solving street crime: practical lessons from the Street Crime Initiative.

Tseloni, A. and K. Pease (2003). "Repeat Personal Victimization: 'Boost' or 'Flag'?" British Journal of Criminology **43**: pp.196-212.

Tucker, L. E. (2003). Safer Stops for Vulnerable Customers. Springfield, VA, National Technical Information Service**:** 87pp.

Valdez, A., C. D. Kaplan and R. L. J. Curtis (2007). "Aggressive Crime, Alcohol and Drug Use, and Concentrated Poverty in 24 U.S. Urban Areas." The American Journal of Drug and Alcohol Abuse **33**: pp. 595-603.

Vogel, M. and J. Pettinari (2002). Personal Safety and Transit: Paths, Environments, Stops and Stations: Final Report. Minneapolis, MN, Center for Transportation Studies.

Walker, A., C. Kershaw and S. Nicholas (2006). Home Office Statistical Bulletin: Crime in England and Wales 2005/06. London, National Statistics. Home Office. UK.

Wang, F. and W. W. Minor (2002). "Where the Jobs Are: Employment Access and Crime Patterns in Cleveland." Annals of the Association of American Geographers **9**(3): pp. 435-450.

Ward, M. D. and K. S. Gleditsch (2008). Spatial Regression Models. Thousand Oaks, CA, Sage Publications, Inc.

Weidner, R. R. (1996). Target Hardening at a New York City Subway Station: Decreased Fare Evasion--at What Price? Preventing Mass Transit Crime. Vol. 6. R. V. Clarke. Monsey, NY, Criminal Justice Press.

Wikstrom, P.-O. H. (1995). "Preventing City-Center Street Crimes." Crime and Justice: A Review of Research **19**(3): pp. 429 - 468.

Wiles, P. and A. Costello (2000). Home Office Research Study 207: The 'road to nowhere': the evidence for traveling criminals. London, Home Office.

Wong, D. W. S. and J. Lee (2005). Statistical Analysis of Geographic Information with ArcView GIS and ArcGIS. Hoboken, NJ, John Wiley & Sons, Inc.,.

Wright, R. T. and S. H. Decker (1997). Armed Robbers in Action. Boston, USA, Northeastern University Press.

Yu, S.-s. V. (2004). Spatial Pattern Analysis of Robbery and Spousal Assault in Vancouver Between 1989 and 2000 Utilizing Geographic Information Systems. School of Criminology. Burnaby, Simon Fraser University.

Appendix A. Data Analysis Results on Robbery

Table A- 1. OLS regression Model 1 output in GeoDa on robbery

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :     ROB2007      Number of Observations: 2602
Mean dependent var  :    0.467717      Number of Variables   :    2
S.D. dependent var  :    0.885616      Degrees of Freedom    : 2600

R-squared           :    0.027745      F-statistic           :     74.1947
Adjusted R-squared  :    0.027371      Prob(F-statistic)     :1.20661e-017
Sum squared residual:     1984.17      Log Likelihood        :     -3339.4
Sigma-square        :    0.763141      Akaike info criterion :      6682.8
S.E. of regression  :     0.87358      Schwarz criterion     :     6694.53
Sigma-square ML     :    0.762555
S.E of regression ML:    0.873244
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                  0.4241953    0.01785552       23.75709    0.0000000
Bus stops                 0.05713615   0.00663322        8.61364    0.0000000
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   1.337648
TEST ON NORMALITY OF ERRORS
TEST                     DF          VALUE            PROB
Jarque-Bera               2           27246.91        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF          VALUE            PROB
Breusch-Pagan test        1           454.4714        0.0000000
Koenker-Bassett test      1           54.30984        0.0000000
SPECIFICATION ROBUST TEST
TEST                     DF          VALUE            PROB
White                     2           100.4148        0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                             MI/DF      VALUE          PROB
Moran's I (error)              0.042916     9.8715004      0.0000000
Lagrange Multiplier (lag)          1       91.9264144      0.0000000
Robust LM (lag)                    1        0.3485476      0.5549364
Lagrange Multiplier (error)        1       93.9789587      0.0000000
Robust LM (error)                  1        2.4010919      0.1212506
Lagrange Multiplier (SARMA)        2       94.3275063      0.0000000
========================= END OF REPORT ======================================
```

Table A- 2. OLS regression Model 2 output in GeoDa on robbery

```
============================== BEGINNING OF REPORT============================
Data set            : TPAll07DataZCNov10
Dependent Variable  :     ROB2007      Number of Observations: 2602
Mean dependent var  :    0.467717      Number of Variables   :    8
S.D. dependent var  :    0.885616      Degrees of Freedom    : 2594

R-squared           :    0.104853      F-statistic           :      43.407
Adjusted R-squared  :    0.102438      Prob(F-statistic)     :           0
Sum squared residual:      1826.8      Log Likelihood        :     -3231.9
Sigma-square        :    0.704242      Akaike info criterion :      6479.8
S.E. of regression  :    0.839192      Schwarz criterion     :     6526.71
Sigma-square ML     :    0.702077
S.E of regression ML:      0.8379
------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
------------------------------------------------------------------------------
Constant                 0.3333371     0.01861698      17.905        0.0000000
Eating/drinking places   0.2433432     0.03080657       7.89907      0.0000000
Automotive retail        0.1421449     0.05240526       2.712416     0.0067232
Food store retail        0.2762376     0.04505933       6.130531     0.0000000
Automotive service       0.05484035    0.03543164       1.547779     0.1217957
Business service        -0.04305708    0.02960539      -1.454366     0.1459691
Personal service         0.173651      0.02857585       6.076843     0.0000000
Banks                    0.0676574     0.08141672       0.8310013    0.4060581
------------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   2.207984
TEST ON NORMALITY OF ERRORS
TEST                   DF          VALUE            PROB
Jarque-Bera             2           19258.43        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                   DF          VALUE            PROB
Breusch-Pagan test      7            971.329        0.0000000
Koenker-Bassett test    7           136.1701        0.0000000
SPECIFICATION ROBUST TEST
TEST                   DF          VALUE            PROB
White                  35           270.9256        0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                            MI/DF       VALUE          PROB
Moran's I (error)             0.043581    10.0607950      0.0000000
Lagrange Multiplier (lag)         1       76.3313621      0.0000000
Robust LM (lag)                   1        0.7028176      0.4018386
Lagrange Multiplier (error)       1       96.9103648      0.0000000
Robust LM (error)                 1       21.2818203      0.0000040
Lagrange Multiplier (SARMA)       2       97.6131824      0.0000000
========================= END OF REPORT ==============================
```

Table A- 3. OLS regression Model 3 output in GeoDa on robbery

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :     ROB2007      Number of Observations: 2602
Mean dependent var  :    0.467717      Number of Variables   :    9
S.D. dependent var  :    0.885616      Degrees of Freedom    : 2593

R-squared           :    0.115289      F-statistic           :     42.2375
Adjusted R-squared  :    0.112559      Prob(F-statistic)     :           0
Sum squared residual:     1805.51      Log Likelihood        :    -3216.64
Sigma-square        :    0.696301      Akaike info criterion :     6451.29
S.E. of regression  :    0.834446      Schwarz criterion     :     6504.06
Sigma-square ML     :    0.693892
S.E of regression ML:    0.833002
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                 0.3163831     0.01876383      16.86133     0.0000000
Bus stops                0.03593552    0.006497768      5.53044     0.0000000
Eating/drinking places   0.2283105     0.03075274       7.424069    0.0000000
Automotive retail        0.1322925     0.05213939       2.537286    0.0112295
Food store retail        0.2662683     0.04484079       5.938082    0.0000000
Automotive service       0.04410935    0.03528469       1.250099    0.2113805
Business service        -0.04892829    0.02945713      -1.661       0.0968323
Personal service         0.1608071     0.02850902       5.640568    0.0000000
Banks                    0.05079077    0.08101378       0.6269399   0.5307335
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   2.283673
TEST ON NORMALITY OF ERRORS
TEST                     DF          VALUE            PROB
Jarque-Bera               2           17087.91        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF          VALUE            PROB
Breusch-Pagan test        8           1058.292        0.0000000
Koenker-Bassett test      8           156.7985        0.0000000
SPECIFICATION ROBUST TEST
TEST                     DF          VALUE            PROB
White                    44           467.4347        0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                             MI/DF       VALUE           PROB
Moran's I (error)              0.043129     9.9705135       0.0000000
Lagrange Multiplier (lag)          1       71.4215490       0.0000000
Robust LM (lag)                    1        0.9298603       0.3348992
Lagrange Multiplier (error)        1       94.9124433       0.0000000
Robust LM (error)                  1       24.4207546       0.0000008
Lagrange Multiplier (SARMA)        2       95.8423036       0.0000000
========================= END OF REPORT ==============================
```

Table A- 4. OLS regression Model 4 output in GeoDa on robbery

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :     ROB2007      Number of Observations: 2602
Mean dependent var  :    0.467717      Number of Variables   :   15
S.D. dependent var  :    0.885616      Degrees of Freedom    : 2587

R-squared           :    0.128032      F-statistic           :     27.1323
Adjusted R-squared  :    0.123313      Prob(F-statistic)     :           0
Sum squared residual:      1779.5      Log Likelihood        :    -3197.77
Sigma-square        :    0.687863      Akaike info criterion :     6425.53
S.E. of regression  :    0.829375      Schwarz criterion     :      6513.5
Sigma-square ML     :    0.683898
S.E of regression ML:    0.826981
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                 0.2423574     0.03102783      7.810969    0.0000000
Bus stops                0.03782092    0.006541581     5.781618    0.0000000
Eating/drinking places   0.2229947     0.03062501      7.281455    0.0000000
Automotive retail        0.1353083     0.05194033      2.605073    0.0092378
Food store retail        0.2528819     0.04471557      5.655344    0.0000000
Automotive service       0.03886394    0.03514485      1.105822    0.2688953
Business service        -0.03916671    0.029364       -1.333834    0.1823753
Personal service         0.1606148     0.02843515      5.648459    0.0000000
Banks                    0.04827766    0.0806628       0.5985121   0.5495323
Mixed landuse            0.1502301     0.0521105       2.882914    0.0039727
Vacant land              0.05403458    0.03417521      1.581105    0.1139726
Public housing           0.133339      0.05354728      2.490117    0.0128318
Colleges                -0.2562042     0.1146697      -2.23428     0.0255494
Grade K-12               0.1993391     0.05450758      3.65709     0.0002602
Parks and cemeteries    -0.03057978    0.04591735     -0.6659744   0.5054747
-------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   4.446725
TEST ON NORMALITY OF ERRORS
TEST                    DF          VALUE            PROB
Jarque-Bera              2           17273.87        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                    DF          VALUE            PROB
Breusch-Pagan test      14           1171.391        0.0000000
Koenker-Bassett test    14           172.3146        0.0000000
SPECIFICATION ROBUST TEST
TEST                    DF          VALUE            PROB
White                   119           N/A             N/A

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                           MI/DF       VALUE          PROB
Moran's I (error)            0.037341     8.8350198      0.0000000
Lagrange Multiplier (lag)       1        60.9879579      0.0000000
Robust LM (lag)                 1         0.2122980      0.6449722
Lagrange Multiplier (error)     1        71.1469502      0.0000000
Robust LM (error)               1        10.3712903      0.0012799
Lagrange Multiplier (SARMA)     2        71.3592482      0.0000000
========================= END OF REPORT ==============================
```

Table A- 5. SEM Model 1 output in GeoDa on robbery

```
============================== BEGINNING OF REPORT============================
Spatial Weight       :    TPSW1600ft.GWT
Dependent Variable   :    ROB2007        Number of Observations: 2602
Mean dependent var   :    0.467717       Number of Variables   :    2
S.D. dependent var   :    0.885616       Degree of Freedom     : 2600
Lag coeff. (Lambda)  :    0.456334

R-squared            :    0.053659       R-squared (BUSE)      : -
Sq. Correlation      : -                 Log Likelihood        :-3312.533060
Sigma-square         :    0.742229       Akaike info criterion :     6629.07
S.E of regression    :    0.861527       Schwarz criterion     : 6640.794190
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                0.4202321      0.03146671      13.35482     0.0000000
Bus stops               0.05586549     0.006688015      8.353075    0.0000000
LAMBDA                  0.4563343      0.06008149       7.595256    0.0000000
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF     VALUE         PROB
Breusch-Pagan test                       1        447.813     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF     VALUE         PROB
Likelihood Ratio Test                    1       53.73698     0.0000000
========================= END OF REPORT =======================================
```

Table A- 6. SEM Model 2 output in GeoDa on robbery

```
============================== BEGINNING OF REPORT============================
Spatial Weight       : TPSW1600ft.GWT
Dependent Variable   :     ROB2007       Number of Observations: 2602
Mean dependent var   :    0.467717       Number of Variables   :    8
S.D. dependent var   :    0.885616       Degree of Freedom     : 2594
Lag coeff. (Lambda)  :    0.452533

R-squared            :    0.128847       R-squared (BUSE)      : -
Sq. Correlation      : -                 Log Likelihood        :-3204.672781
Sigma-square         :    0.683258       Akaike info criterion :     6425.35
S.E of regression    :    0.826594       Schwarz criterion     : 6472.257846
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                0.3259575     0.03086349     10.56126     0.0000000
Eating/drinking places  0.2470046     0.03057659      8.078227    0.0000000
Automotive retail       0.1534878     0.05194658      2.954724    0.0031296
Food store retail       0.2417109     0.04454229      5.426547    0.0000001
Automotive service      0.08262932    0.03563487      2.318777    0.0204071
Business service       -0.03183132    0.02937659     -1.083561    0.2785596
Personal service        0.1779599     0.02837971      6.270674    0.0000000
Banks                   0.1034202     0.07996736      1.29328     0.1959143
LAMBDA                  0.452533      0.06034877      7.498628    0.0000000
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
```

```
RANDOM COEFFICIENTS
TEST                                     DF     VALUE         PROB
Breusch-Pagan test                       7       1012.587     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF     VALUE         PROB
Likelihood Ratio Test                    1       54.45229     0.0000000
========================= END OF REPORT ======================================
```

Table A- 7. SEM Model 3 output in GeoDa on robbery

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :     ROB2007       Number of Observations: 2602
Mean dependent var  :    0.467717       Number of Variables   :    9
S.D. dependent var  :    0.885616       Degree of Freedom     : 2593
Lag coeff. (Lambda) :    0.448827

R-squared           :    0.138571       R-squared (BUSE)      : -
Sq. Correlation     : -                 Log Likelihood        :-3189.918128
Sigma-square        :    0.675632       Akaike info criterion :     6397.84
S.E of regression   :    0.821968       Schwarz criterion     : 6450.612578
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                 0.3089868      0.03066097      10.07753     0.0000000
Bus stops                0.03551289     0.00651843       5.448074    0.0000001
Eating/drinking places   0.2342876      0.03049401       7.68307     0.0000000
Automotive retail        0.1402427      0.05171213       2.711989    0.0066882
Food store retail        0.2319818      0.0443309        5.232959    0.0000002
Automotive service       0.07106939     0.03549322       2.002337    0.0452483
Business service        -0.03408415     0.02921422      -1.166697    0.2433328
Personal service         0.1658179      0.02830763       5.857708    0.0000000
Banks                    0.08421126     0.0795997        1.057934    0.2900854
LAMBDA                   0.4488266      0.06060852       7.405339    0.0000000
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF     VALUE         PROB
Breusch-Pagan test                       8       1095.873     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF     VALUE         PROB
Likelihood Ratio Test                    1       53.44927     0.0000000
========================= END OF REPORT ======================================
```

Table A- 8. SEM Model 4 output in GeoDa on robbery

```
============================== BEGINNING OF REPORT============================
Spatial Weight       : TPSW1600ft.GWT
Dependent Variable   :     ROB2007      Number of Observations: 2602
Mean dependent var   :    0.467717      Number of Variables   :   15
S.D. dependent var   :    0.885616      Degree of Freedom     : 2587
Lag coeff. (Lambda) :    0.428312

R-squared            :    0.147493      R-squared (BUSE)      : -
Sq. Correlation      : -                Log Likelihood        :-3175.572269
Sigma-square         :    0.668634      Akaike info criterion :     6381.14
S.E of regression    :    0.817701      Schwarz criterion     : 6469.105072
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                 0.2587936      0.03917368      6.606315     0.0000000
Bus stops                0.03646878     0.006542165     5.574421     0.0000000
Eating/drinking places   0.2284665      0.03039154      7.517437     0.0000000
Automotive retail        0.1466694      0.05153153      2.846206     0.0044245
Food store retail        0.225683       0.04419559      5.106459     0.0000003
Automotive service       0.06356108     0.03536091      1.797495     0.0722569
Business service        -0.03073671     0.02908788     -1.056685     0.2906556
Personal service         0.163952       0.0282075       5.812354     0.0000000
Banks                    0.08069603     0.07929943      1.017612     0.3088625
Vacant land              0.02574043     0.03538234      0.7274937    0.4669234
Public housing           0.1519299      0.05891404      2.578841     0.0099133
Colleges                -0.1548428      0.1295292      -1.195428     0.2319200
Grade K-12               0.1650343      0.05451555      3.027288     0.0024677
Parks and cemeteries    -0.06237631     0.0483371      -1.290444     0.1968968
Mixed landuse            0.1569362      0.05308023      2.956585     0.0031108
LAMBDA                   0.4283119      0.06203116      6.904786     0.0000000
-------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF     VALUE          PROB
Breusch-Pagan test                      14       1190.398     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF     VALUE          PROB
Likelihood Ratio Test                    1       44.38993     0.0000000
========================= END OF REPORT ======================================
```

Table A- 9. PRM Model 1 output in Stata on robbery

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(2)      =     195.06
                                                  Prob > chi2     =     0.0000
Log likelihood = -2419.9707                       Pseudo R2       =     0.0387
------------------------------------------------------------------------------
     rob2007 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wrob |   1.354881   .1151167    11.77   0.000     1.129256    1.580505
      bstops |   .0549711   .0054066    10.17   0.000     .0443745    .0655677
       _cons |  -1.510849   .0699959   -21.58   0.000    -1.648039    -1.37366
------------------------------------------------------------------------------
```

Table A- 10. PRM Model 2 output in Stata on robbery

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(8)      =     379.82
                                                  Prob > chi2     =     0.0000
Log likelihood = -2327.5891                       Pseudo R2       =     0.0754
------------------------------------------------------------------------------
     rob2007 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wrob |   1.133509   .1171518     9.68   0.000     .9038956    1.363122
     sic58xx |   .2170135   .0325008     6.68   0.000      .153313     .280714
     sic55xx |   .2275347   .0622709     3.65   0.000     .1054859    .3495834
     sic54xx |   .2685493   .0488832     5.49   0.000     .1727399    .3643587
     sic75xx |   .1400939   .0516264     2.71   0.007      .038908    .2412797
     sic73xx |  -.0796112    .045451    -1.75   0.080    -.1686936    .0094711
     sic72xx |   .1956615   .0341683     5.73   0.000     .1286929    .2626302
     finance |   .1649812   .0906777     1.82   0.069    -.0127439    .3427063
       _cons |  -1.559632   .0699052   -22.31   0.000    -1.696643    -1.42262
------------------------------------------------------------------------------
```

Table A- 11. PRM Model 3 output in Stata on robbery

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(9)      =     408.56
                                                  Prob > chi2     =     0.0000
Log likelihood = -2313.2185                       Pseudo R2       =     0.0811
------------------------------------------------------------------------------
     rob2007 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wrob |   1.108458   .1177836     9.41   0.000     .8776067     1.33931
      bstops |    .041259   .0066015     6.25   0.000     .0283203    .0541976
     sic58xx |   .2125587   .0329394     6.45   0.000     .1479986    .2771188
     sic55xx |   .2178371   .0638732     3.41   0.001      .092648    .3430263
     sic54xx |   .2558593   .0493919     5.18   0.000     .1590529    .3526657
     sic75xx |   .1208952   .0520752     2.32   0.020     .0188297    .2229606
     sic73xx |  -.0879754   .0470132    -1.87   0.061    -.1801196    .0041688
     sic72xx |   .1819737   .0341235     5.33   0.000     .1150928    .2488546
     finance |   .1327858   .0929993     1.43   0.153    -.0494894     .315061
       _cons |  -1.577855   .0702888   -22.45   0.000    -1.715618   -1.440091
------------------------------------------------------------------------------
```

Table A- 12. PRM Model 4 output in Stata on robbery

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(15)     =     452.85
                                                  Prob > chi2     =     0.0000
Log likelihood = -2291.0716                       Pseudo R2       =     0.0899
------------------------------------------------------------------------------
     rob2007 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wrob |   1.083869    .120417     9.00   0.000     .8478562    1.319882
      bstops |   .0436977    .006578     6.64   0.000      .030805    .0565904
     sic58xx |   .2079635   .0334842     6.21   0.000     .1423355    .2735914
     sic55xx |   .2158251   .0654593     3.30   0.001     .0875272     .344123
     sic54xx |   .2308745   .0510469     4.52   0.000     .1308243    .3309246
     sic75xx |   .1090867   .0521323     2.09   0.036     .0069094    .2112641
     sic73xx |  -.0600131   .0459736    -1.31   0.192    -.1501197    .0300936
     sic72xx |   .1848101    .034331     5.38   0.000     .1175225    .2520977
     finance |   .1163194    .096565     1.20   0.228    -.0729446    .3055833
    zoningc8 |   .2894854    .079565     3.64   0.000     .1335409    .4454299
   pc1fvland |   .0695619   .0623015     1.12   0.264    -.0525468    .1916705
    phousing |   .2911929   .0884976     3.29   0.001     .1177407     .464645
    univcoll |  -.5869365   .2932359    -2.00   0.045    -1.161668   -.0122047
     schools |   .2772035   .0838208     3.31   0.001     .1129177    .4414893
   openspace |  -.1346245   .0849956    -1.58   0.113    -.3012128    .0319639
       _cons |  -1.686824   .0828253   -20.37   0.000    -1.849158   -1.524489
------------------------------------------------------------------------------
```

Table A- 13. NBRM Model 1 output in Stata on robbery

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(2)      =     139.80
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2318.4603                       Pseudo R2       =     0.0293
------------------------------------------------------------------------------
     rob2007 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wrob |   1.355627   .1521322     8.91   0.000     1.057453      1.6538
      bstops |   .0895957   .0127383     7.03   0.000     .0646292    .1145623
       _cons |  -1.549659   .0872612   -17.76   0.000    -1.720687    -1.37863
-------------+----------------------------------------------------------------
    /lnalpha |  -.0437291    .110005                     -.2593349    .1718767
-------------+----------------------------------------------------------------
       alpha |   .9572132   .1052982                      .7715646    1.187531
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  203.02 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2388.359    -2318.46      4     4644.921    4668.377
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table A- 14. NBRM Model 2 output in Stata on robbery

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(8)      =     260.73
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2257.9929                       Pseudo R2       =     0.0546
------------------------------------------------------------------------------
     rob2007 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wrob |   1.249438   .1481403     8.43   0.000     .9590879    1.539787
     sic58xx |   .2421956    .050286     4.82   0.000     .1436369    .3407543
     sic55xx |   .2480275   .0937721     2.65   0.008     .0642377    .4318174
     sic54xx |   .3677943   .0765361     4.81   0.000     .2177863    .5178024
     sic75xx |   .1331449   .0643812     2.07   0.039     .0069599    .2593298
     sic73xx |  -.0241309   .0600235    -0.40   0.688    -.1417749     .093513
     sic72xx |    .228216   .0469402     4.86   0.000     .1362149    .3202172
     finance |   .1156393   .1346471     0.86   0.390    -.1482643    .3795428
       _cons |  -1.663789   .0870108   -19.12   0.000    -1.834327   -1.493251
-------------+----------------------------------------------------------------
    /lnalpha |  -.3341333   .1282579                     -.5855142   -.0827524
-------------+----------------------------------------------------------------
       alpha |   .7159583   .0918273                      .5568195    .9205791
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  139.19 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2388.359   -2257.993     10     4535.986    4594.626
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table A- 15. NBRM Model 3 output in Stata on robbery

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(9)      =     277.27
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2249.7251                       Pseudo R2       =     0.0580
------------------------------------------------------------------------------
     rob2007 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wrob |   1.222575   .1474757     8.29   0.000     .9335278    1.511622
      bstops |   .0467676   .0114614     4.08   0.000     .0243037    .0692315
     sic58xx |   .2170191   .0497667     4.36   0.000     .1194783      .31456
     sic55xx |   .2242624   .0922571     2.43   0.015     .0434418     .405083
     sic54xx |   .3618193   .0758407     4.77   0.000     .2131742    .5104644
     sic75xx |   .1252335   .0641242     1.95   0.051    -.0004475    .2509146
     sic73xx |  -.0413798   .0601443    -0.69   0.491    -.1592604    .0765008
     sic72xx |    .202887   .0470018     4.32   0.000     .1107652    .2950088
     finance |    .080832   .1340317     0.60   0.546    -.1818653    .3435293
       _cons |  -1.676592   .0867186   -19.33   0.000    -1.846557   -1.506626
-------------+----------------------------------------------------------------
    /lnalpha |  -.3813014   .1321917                     -.6403924   -.1222105
-------------+----------------------------------------------------------------
       alpha |    .682972   .0902832                      .5270856     .884962
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  126.99 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2388.359   -2249.725     11      4521.45    4585.955
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table A- 16. NBRM Model 4 output in Stata on robbery

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(15)     =     310.45
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2233.1351                       Pseudo R2       =     0.0650
------------------------------------------------------------------------------
     rob2007 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wrob |   1.148516   .1493072     7.69   0.000     .8558798    1.441153
      bstops |   .0531592     .01172     4.54   0.000     .0301885    .0761299
     sic58xx |   .2089132   .0493691     4.23   0.000     .1121515     .305675
     sic55xx |   .2247896   .0914902     2.46   0.014     .0454721    .4041071
     sic54xx |   .3383081    .075431     4.49   0.000     .1904661    .4861502
     sic75xx |   .1121274   .0637464     1.76   0.079    -.0128133    .2370682
     sic73xx |   -.016552   .0597626    -0.28   0.782    -.1336845    .1005806
     sic72xx |   .2007236   .0469164     4.28   0.000     .1087691     .292678
     finance |   .0815668   .1326497     0.61   0.539    -.1784219    .3415556
    zoningc8 |   .2944121   .0978824     3.01   0.003     .1025661    .4862581
   pc1fvland |   .0686168    .073462     0.93   0.350    -.0753662    .2125997
    phousing |   .2923831    .104615     2.79   0.005     .0873414    .4974248
    univcoll |  -.5846308   .3228076    -1.81   0.070    -1.217322    .0480605
     schools |   .3180115   .1033293     3.08   0.002     .1154897    .5205333
   openspace |  -.1297171   .0987983    -1.31   0.189    -.3233583     .063924
       _cons |   -1.76796    .099004   -17.86   0.000    -1.962004   -1.573916
-------------+----------------------------------------------------------------
    /lnalpha |  -.4440061   .1363078                     -.7111645   -.1768477
-------------+----------------------------------------------------------------
       alpha |   .6414615   .0874362                       .491072    .8379074
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  115.87 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2388.359   -2233.135     17      4500.27    4599.959
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table A- 17. Model fit comparison of PRM & NBRM Model 1 on robbery

```
----------------------------------------------------------
                        Variable |    PRM        NBRM
---------------------------------+------------------------
rob2007                          |
                            WROB |     3.876        3.879
                                 |     11.77         8.91
                          BSTOPS |     1.057        1.094
                                 |     10.17         7.03
                        Constant |     0.221        0.212
                                 |    -21.58       -17.76
---------------------------------+------------------------
lnalpha                          |
                        Constant |                  0.957
                                 |                  -0.40
---------------------------------+------------------------
Statistics                       |
                           alpha |                  0.957
                               N |  2602.000     2602.000
                              ll | -2419.971    -2318.460
                             bic |  4863.534     4668.377
                             aic |  4845.941     4644.921
----------------------------------------------------------
```

```
                                                     legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM         -0.070         1      0.014
NBRM        -0.006         0      0.001

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.686       0.638      0.048     9.417
1        0.208       0.278      0.070    45.769
2        0.062       0.068      0.006     1.394
3        0.021       0.013      0.008    13.291
4        0.007       0.002      0.004    18.991
5        0.003       0.001      0.002    23.013
6        0.002       0.000      0.001    25.019
7        0.000       0.000      0.000     2.079
8        0.000       0.000      0.000     4.605
9        0.000       0.000      0.000     9.437
------------------------------------------------
Sum      0.989       1.000      0.141   153.016

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.686       0.692      0.006     0.130
1        0.208       0.208      0.000     0.000
2        0.062       0.065      0.003     0.330
3        0.021       0.021      0.000     0.015
4        0.007       0.008      0.001     0.472
5        0.003       0.003      0.000     0.103
6        0.002       0.001      0.000     0.091
7        0.000       0.001      0.000     0.265
8        0.000       0.000      0.000     0.013
9        0.000       0.000      0.000     0.424
------------------------------------------------
Sum      0.989       0.999      0.012     1.844

Tests and Fit Statistics

PRM             BIC=-15598.687  AIC=     1.862  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM       BIC=-15793.844  dif=   195.157  NBRM    PRM   Very strong
                AIC=     1.785  dif=     0.077  NBRM    PRM
                LRX2=  203.021  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM            BIC=-15793.844  AIC=     1.785  Prefer  Over  Evidence
```

Table A- 18. Model fit comparison of PRM & NBRM Model 2 on robbery

Variable	PRM	NBRM
rob2007		
WROB	3.054	3.469
	9.56	8.40
SIC58XX	1.250	1.283
	6.94	5.00
SIC55XX	1.250	1.275
	3.58	2.59
SIC54XX	1.322	1.459
	5.82	4.98
SIC75XX	1.156	1.146
	2.83	2.12
SIC73XX	0.924	0.979
	-1.75	-0.35
SIC72XX	1.220	1.259
	5.84	4.90
Constant	0.213	0.190
	-22.29	-19.09
lnalpha		
Constant		0.719
		-2.58
Statistics		
alpha		0.719
N	2602.000	2602.000
ll	-2329.091	-2258.360
bic	4721.094	4587.496
aic	4674.182	4534.720

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
PRM	-0.059	1	0.011
NBRM	-0.005	0	0.001

PRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.686	0.648	0.038	5.667
1	0.208	0.267	0.059	34.346
2	0.062	0.065	0.003	0.336
3	0.021	0.014	0.007	10.405
4	0.007	0.003	0.003	7.907
5	0.003	0.001	0.002	6.138
6	0.002	0.000	0.001	6.482
7	0.000	0.000	0.000	0.204
8	0.000	0.000	0.000	0.831
9	0.000	0.000	0.000	1.539
Sum	0.989	1.000	0.114	73.856

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson

```
------------------------------------------------
0        0.686       0.691      0.005     0.092
1        0.208       0.213      0.005     0.262
2        0.062       0.062      0.000     0.005
3        0.021       0.020      0.001     0.283
4        0.007       0.007      0.001     0.119
5        0.003       0.003      0.000     0.085
6        0.002       0.001      0.000     0.005
7        0.000       0.001      0.000     0.634
8        0.000       0.001      0.000     0.093
9        0.000       0.000      0.000     0.009
------------------------------------------------
Sum      0.989       0.999      0.013     1.587

Tests and Fit Statistics
PRM            BIC=-15741.127  AIC=     1.796  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM      BIC=-15874.725  dif=   133.598  NBRM    PRM   Very strong
               AIC=     1.743  dif=     0.054  NBRM    PRM
               LRX2=  141.462  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM           BIC=-15874.725  AIC=     1.743  Prefer  Over  Evidence
```

Table A- 19. Model fit comparison of PRM & NBRM Model 3 on robbery

```
. countfit rob2007 wrob bstops sic58xx sic55xx sic54xx sic75xx sic73xx sic72xx,
nbreg prm

----------------------------------------------------------
                        Variable |    PRM        NBRM
---------------------------------+------------------------
rob2007                          |
                            WROB |     2.987       3.382
                                 |      9.32        8.27
                          BSTOPS |     1.043       1.048
                                 |      6.39        4.12
                         SIC58XX |     1.243       1.248
                                 |      6.71        4.49
                         SIC55XX |     1.239       1.247
                                 |      3.34        2.40
                         SIC54XX |     1.303       1.446
                                 |      5.46        4.91
                         SIC75XX |     1.133       1.136
                                 |      2.40        1.99
                         SIC73XX |     0.916       0.961
                                 |     -1.88       -0.66
                         SIC72XX |     1.202       1.226
                                 |      5.41        4.34
                        Constant |     0.208       0.187
                                 |    -22.45      -19.32
---------------------------------+------------------------
lnalpha                          |
                        Constant |                 0.685
                                 |                 -2.87
---------------------------------+------------------------
Statistics                       |
                           alpha |                 0.685
                               N |  2602.000    2602.000
                              ll | -2314.165   -2249.906
                             bic |  4699.106    4578.452
                             aic |  4646.330    4519.812
```

```
----------------------------------------------------------
                                              legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM        -0.057         1      0.011
NBRM       -0.005         1      0.001

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.686       0.651      0.035     5.036
1        0.208       0.265      0.057    31.762
2        0.062       0.064      0.002     0.216
3        0.021       0.014      0.007     9.300
4        0.007       0.004      0.003     5.669
5        0.003       0.001      0.001     3.753
6        0.002       0.001      0.001     3.911
7        0.000       0.000      0.000     0.038
8        0.000       0.000      0.000     0.497
9        0.000       0.000      0.000     1.349
------------------------------------------------
Sum      0.989       1.000      0.107    61.530

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.686       0.691      0.005     0.080
1        0.208       0.213      0.005     0.354
2        0.062       0.062      0.000     0.000
3        0.021       0.019      0.002     0.369
4        0.007       0.007      0.001     0.104
5        0.003       0.003      0.000     0.101
6        0.002       0.002      0.000     0.000
7        0.000       0.001      0.000     0.716
8        0.000       0.001      0.000     0.131
9        0.000       0.000      0.000     0.001
------------------------------------------------
Sum      0.989       0.999      0.013     1.858

Tests and Fit Statistics

PRM             BIC=-15763.115  AIC=     1.786  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM       BIC=-15883.768  dif=   120.654  NBRM    PRM   Very strong
                AIC=     1.737  dif=     0.049  NBRM    PRM
                LRX2=  128.518  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM            BIC=-15883.768  AIC=     1.737  Prefer  Over  Evidence
```

Table A- 20. Model fit comparison of PRM & NBRM Model 4 on robbery

Variable	PRM	NBRM
rob2007		
WROB	2.921	3.142
	8.93	7.67
BSTOPS	1.045	1.055
	6.73	4.57
SIC58XX	1.237	1.238
	6.43	4.35
SIC55XX	1.236	1.248
	3.23	2.42
SIC54XX	1.270	1.413
	4.75	4.63
SIC75XX	1.119	1.121
	2.16	1.80
SIC73XX	0.941	0.985
	-1.32	-0.25
SIC72XX	1.207	1.224
	5.49	4.30
mixedz== 9.0000	1.347	1.346
	3.76	3.04
PC1FVLAND	1.068	1.067
	1.05	0.89
PHOUSING	1.336	1.340
	3.27	2.80
UNIVCOLL	0.556	0.558
	-2.00	-1.81
SCHOOLS	1.320	1.374
	3.31	3.07
OPENSPACE	0.872	0.880
	-1.61	-1.30
Constant	0.187	0.171
	-20.38	-17.85
lnalpha		
Constant		0.643
		-3.24
Statistics		
alpha		0.643
N	2602.000	2602.000
ll	-2291.754	-2233.323
bic	4701.469	4592.470
aic	4613.509	4498.645

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
PRM	-0.053	1	0.010
NBRM	-0.005	1	0.001

PRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.686	0.653	0.033	4.342
1	0.208	0.261	0.053	28.451

2	0.062	0.064	0.002	0.230
3	0.021	0.014	0.007	8.086
4	0.007	0.004	0.003	4.739
5	0.003	0.001	0.001	2.904
6	0.002	0.001	0.001	2.732
7	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.261
9	0.000	0.000	0.000	1.109
Sum	0.989	1.000	0.101	52.855

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.686	0.691	0.005	0.092
1	0.208	0.213	0.005	0.326
2	0.062	0.062	0.000	0.002
3	0.021	0.019	0.002	0.382
4	0.007	0.007	0.001	0.115
5	0.003	0.003	0.000	0.115
6	0.002	0.002	0.000	0.000
7	0.000	0.001	0.001	0.753
8	0.000	0.001	0.000	0.155
9	0.000	0.000	0.000	0.000
Sum	0.989	0.999	0.014	1.939

Tests and Fit Statistics

PRM	BIC=-15760.751	AIC= 1.773	Prefer	Over	Evidence
vs NBRM	BIC=-15869.751	dif= 109.000	NBRM	PRM	Very strong
	AIC= 1.729	dif= 0.044	NBRM	PRM	
	LRX2= 116.864	prob= 0.000	NBRM	PRM	p=0.000
NBRM	BIC=-15869.751	AIC= 1.729	Prefer	Over	Evidence

Table A- 21. Model fit comparison of NBRM & ZINBRM Model 1 on robbery

Variable	NBRM	ZINBRM
rob2007		
WROB	3.879	2.686
	8.91	4.95
BSTOPS	1.094	1.062
	7.03	4.37
Constant	0.212	0.325
	-17.76	-8.36
lnalpha		
Constant	0.957	0.648
	-0.40	-2.28
inflate		
WROB		0.085
		-1.66
BSTOPS		0.038
		-0.52
Constant		0.952

```
                                 |                  -0.09
---------------------------------+-----------------------
Statistics                       |
                           alpha |     0.957
                               N |  2602.000    2602.000
                              ll | -2318.460   -2310.711
                             bic |  4668.377    4676.470
                             aic |  4644.921    4635.422
---------------------------------------------------------
                                              legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM        -0.006         0      0.001
ZINBRM        -0.008         0      0.003

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.686       0.692      0.006     0.130
1        0.208       0.208      0.000     0.000
2        0.062       0.065      0.003     0.330
3        0.021       0.021      0.000     0.015
4        0.007       0.008      0.001     0.472
5        0.003       0.003      0.000     0.103
6        0.002       0.001      0.000     0.091
7        0.000       0.001      0.000     0.265
8        0.000       0.000      0.000     0.013
9        0.000       0.000      0.000     0.424
------------------------------------------------
Sum      0.989       0.999      0.012     1.844

ZINB: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.686       0.694      0.008     0.244
1        0.208       0.200      0.008     0.750
2        0.062       0.069      0.008     2.113
3        0.021       0.023      0.002     0.526
4        0.007       0.008      0.001     0.646
5        0.003       0.003      0.000     0.014
6        0.002       0.001      0.000     0.540
7        0.000       0.000      0.000     0.021
8        0.000       0.000      0.000     0.390
9        0.000       0.000      0.000     1.732
------------------------------------------------
Sum      0.989       1.000      0.028     6.976

Tests and Fit Statistics

-------------------------------------------------------------------------
NBRM           BIC=-15793.844  AIC=     1.785  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-15785.751  dif=    -8.093  NBRM    ZINBRM  Strong
               AIC=     1.781  dif=     0.004  ZINBRM    NBRM
               Vuong=   2.290  prob=    0.011  ZINBRM    NBRM  p=0.011
```

Table A- 22. Model fit comparison of NBRM & ZINBRM Model 2 on robbery

Variable	NBRM	ZINBRM
rob2007		
WROB	3.488	2.110
	8.43	4.33
SIC58XX	1.274	1.255
	4.82	4.81
SIC55XX	1.281	1.136
	2.65	1.46
SIC54XX	1.445	1.279
	4.81	3.34
SIC75XX	1.142	1.141
	2.07	1.93
SIC73XX	0.976	0.916
	-0.40	-1.52
SIC72XX	1.256	1.174
	4.86	3.59
FINANCE	1.123	1.136
	0.86	0.98
Constant	0.189	0.328
	-19.12	-8.87
lnalpha		
Constant	0.716	0.454
	-2.61	-4.08
inflate		
WROB		0.022
		-3.21
SIC58XX		0.703
		-0.75
SIC55XX		0.020
		-0.84
SIC54XX		0.000
		-0.01
SIC75XX		1.059
		0.21
SIC73XX		0.000
		-0.02
SIC72XX		0.203
		-1.25
FINANCE		1.675
		0.61
Constant		2.311
		1.92
Statistics		
alpha	0.716	
N	2602.000	2602.000
ll	-2257.993	-2241.628
bic	4594.626	4632.672
aic	4535.986	4521.255

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
NBRM	-0.005	0	0.001
ZINBRM	-0.007	0	0.002

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.686	0.691	0.005	0.088
1	0.208	0.213	0.005	0.275
2	0.062	0.062	0.000	0.005
3	0.021	0.020	0.001	0.288
4	0.007	0.007	0.001	0.117
5	0.003	0.003	0.000	0.086
6	0.002	0.001	0.000	0.004
7	0.000	0.001	0.000	0.642
8	0.000	0.001	0.000	0.097
9	0.000	0.000	0.000	0.008
Sum	0.989	0.999	0.013	1.612

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.686	0.693	0.007	0.194
1	0.208	0.203	0.005	0.271
2	0.062	0.069	0.007	1.727
3	0.021	0.022	0.001	0.106
4	0.007	0.007	0.001	0.269
5	0.003	0.003	0.000	0.002
6	0.002	0.001	0.000	0.331
7	0.000	0.001	0.000	0.137
8	0.000	0.000	0.000	0.062
9	0.000	0.000	0.000	0.627
Sum	0.989	1.000	0.021	3.727

Tests and Fit Statistics

```
-------------------------------------------------------------------------
NBRM           BIC=-15867.595  AIC=     1.743  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-15829.549  dif=   -38.046  NBRM    ZINBRM  Very strong
               AIC=     1.738  dif=     0.006  ZINBRM    NBRM
               Vuong=   3.167  prob=    0.001  ZINBRM    NBRM  p=0.001
```

Table A- 23. Model fit comparison of NBRM & ZINBRM Model 3 on robbery

	Variable	NBRM	ZINBRM
rob2007			
	WROB	3.382	2.262
		8.27	4.93
	BSTOPS	1.048	1.033
		4.12	3.19
	SIC58XX	1.248	1.231
		4.49	4.52
	SIC55XX	1.247	1.143
		2.40	1.56
	SIC54XX	1.446	1.308
		4.91	3.82

```
                         SIC75XX |     1.136       1.109   
                                 |      1.99        1.59   
                         SIC73XX |     0.961       0.921   
                                 |     -0.66       -1.45   
                         SIC72XX |     1.226       1.161   
                                 |      4.34        3.43   
                        Constant |     0.187       0.302   
                                 |    -19.32      -11.02   
---------------------------------+-----------------------
lnalpha                          |
                        Constant |     0.685       0.464   
                                 |     -2.87       -4.36   
---------------------------------+-----------------------
inflate                          |
                            WROB |                 0.035   
                                 |                 -2.60   
                          BSTOPS |                 0.000   
                                 |                 -0.01   
                         SIC58XX |                 0.546   
                                 |                 -0.76   
                         SIC55XX |                 0.032   
                                 |                 -0.84   
                         SIC54XX |                 0.000   
                                 |                 -0.02   
                         SIC75XX |                 0.909   
                                 |                 -0.30   
                         SIC73XX |                 0.005   
                                 |                 -0.39   
                         SIC72XX |                 0.000   
                                 |                 -0.00   
                        Constant |                 2.177   
                                 |                  1.58   
---------------------------------+-----------------------
Statistics                       |
                           alpha |     0.685               
                               N |  2602.000    2602.000   
                              ll | -2249.906   -2230.304   
                             bic |  4578.452    4610.025   
                             aic |  4519.812    4498.608   
---------------------------------------------------------
                                             legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.005         1      0.001
ZINBRM       -0.007         0      0.002

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.686       0.691      0.005     0.080
1        0.208       0.213      0.005     0.354
2        0.062       0.062      0.000     0.000
3        0.021       0.019      0.002     0.369
4        0.007       0.007      0.001     0.104
5        0.003       0.003      0.000     0.101
6        0.002       0.002      0.000     0.000
7        0.000       0.001      0.000     0.716
8        0.000       0.001      0.000     0.131
9        0.000       0.000      0.000     0.001
```

Sum	0.989	0.999	0.013	1.858

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.686	0.693	0.007	0.161
1	0.208	0.205	0.003	0.120
2	0.062	0.068	0.006	1.419
3	0.021	0.022	0.001	0.048
4	0.007	0.007	0.001	0.243
5	0.003	0.003	0.000	0.007
6	0.002	0.001	0.000	0.253
7	0.000	0.001	0.000	0.183
8	0.000	0.000	0.000	0.031
9	0.000	0.000	0.000	0.500
Sum	0.989	1.000	0.018	2.965

Tests and Fit Statistics

```
-------------------------------------------------------------------------
NBRM           BIC=-15883.768  AIC=     1.737  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-15852.196  dif=   -31.573  NBRM    ZINBRM  Very strong
               AIC=     1.729  dif=     0.008  ZINBRM    NBRM
               Vuong=   3.478  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table A- 24. Model fit comparison of NBRM & ZINBRM Model 4 on robbery

Variable	NBRM	ZINBRM
rob2007		
WROB	3.142	2.435
	7.67	4.80
BSTOPS	1.055	1.035
	4.57	3.22
SIC58XX	1.238	1.207
	4.35	4.13
SIC55XX	1.248	1.128
	2.42	1.42
SIC54XX	1.413	1.275
	4.63	3.48
SIC75XX	1.121	1.095
	1.80	1.38
SIC73XX	0.985	0.936
	-0.25	-1.16
SIC72XX	1.224	1.164
	4.30	3.48
mixedz== 9.0000	1.346	1.238
	3.04	1.94
PC1FVLAND	1.067	1.000
	0.89	0.00
PHOUSING	1.340	1.219
	2.80	1.51
UNIVCOLL	0.558	0.513
	-1.81	-1.59
SCHOOLS	1.374	1.194
	3.07	1.49
OPENSPACE	0.880	1.078

```
                                 |     -1.30        0.61
                        Constant |     0.171       0.279
                                 |    -17.85       -9.63
---------------------------------+------------------------
lnalpha                          |
                        Constant |     0.643       0.410
                                 |     -3.24       -4.64
---------------------------------+------------------------
inflate                          |
                            WROB |                 0.155
                                 |                 -1.96
                          BSTOPS |                 0.314
                                 |                 -1.45
                         SIC58XX |                 0.280
                                 |                 -1.11
                         SIC55XX |                 0.018
                                 |                 -0.87
                         SIC54XX |                 0.000
                                 |                 -0.13
                         SIC75XX |                 0.946
                                 |                 -0.15
                         SIC73XX |                 0.076
                                 |                 -0.80
                         SIC72XX |                 0.097
                                 |                 -0.77
             mixedz==     9.0000 |                 0.578
                                 |                 -0.77
                       PC1FVLAND |                 0.764
                                 |                 -0.72
                        PHOUSING |                 0.500
                                 |                 -1.02
                        UNIVCOLL |                 0.426
                                 |                 -0.29
                         SCHOOLS |                 0.314
                                 |                 -1.31
                       OPENSPACE |                 2.947
                                 |                  2.69
                        Constant |                 1.578
                                 |                  0.96
---------------------------------+------------------------
Statistics                       |
                           alpha |     0.643
                               N |  2602.000    2602.000
                              ll | -2233.323   -2210.889
                             bic |  4592.470    4665.563
                             aic |  4498.645    4483.778
----------------------------------------------------------
                                             legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.005         1      0.001
ZINBRM       -0.007         0      0.002

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.686       0.691       0.005     0.092
1        0.208       0.213       0.005     0.326
2        0.062       0.062       0.000     0.002
```

3	0.021	0.019	0.002	0.382
4	0.007	0.007	0.001	0.115
5	0.003	0.003	0.000	0.115
6	0.002	0.002	0.000	0.000
7	0.000	0.001	0.001	0.753
8	0.000	0.001	0.000	0.155
9	0.000	0.000	0.000	0.000
Sum	0.989	0.999	0.014	1.939

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.686	0.693	0.007	0.196
1	0.208	0.204	0.004	0.249
2	0.062	0.068	0.006	1.593
3	0.021	0.022	0.001	0.090
4	0.007	0.007	0.001	0.288
5	0.003	0.003	0.000	0.009
6	0.002	0.001	0.000	0.259
7	0.000	0.001	0.000	0.175
8	0.000	0.000	0.000	0.036
9	0.000	0.000	0.000	0.525
Sum	0.989	1.000	0.021	3.420

Tests and Fit Statistics

```
-------------------------------------------------------------------------
NBRM           BIC=-15869.751  AIC=     1.729  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-15796.657  dif=   -73.093  NBRM    ZINBRM  Very strong
               AIC=     1.723  dif=     0.006  ZINBRM    NBRM
               Vuong=   3.683  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table A- 25. Fit comparison of NBRM Model 1 and Model 2 on robbery

Measures of Fit for nbreg of rob2007, Current = M2, Saved = M1

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-2388.359	-2388.359	0.000
Log-Lik Full Model	-2257.993	-2318.460	60.467
D	4515.986(2592)	4636.921(2598)	120.935(6)
LR	260.733(8)	139.798(2)	120.935(6)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.055	0.029	0.025
McFadden's Adj R2	0.050	0.028	0.023
ML (Cox-Snell) R2	0.095	0.052	0.043
Cragg-Uhler(Nagelkerke) R2	0.113	0.062	0.051
AIC	1.743	1.785	-0.042
AIC*n	4535.986	4644.921	-108.935
BIC	-15867.595	-15793.844	-73.751
BIC'	-197.821	-124.070	-73.751
BIC used by Stata	4594.626	4668.377	-73.751
AIC used by Stata	4535.986	4644.921	-108.935

Difference of 73.751 in BIC' provides very strong support for current model.

Note: p-value for difference in LR is only valid if models are nested.

Table A- 26. Fit comparison of NBRM Model 2 and Model 3on robbery

Measures of Fit for nbreg of rob2007, current = M3, saved = M2

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-2388.359	-2388.359	0.000
Log-Lik Full Model	-2249.725	-2257.993	8.268
D	4499.450(2591)	4515.986(2592)	16.536(1)
LR	277.268(9)	260.733(8)	16.536(1)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.058	0.055	0.003
McFadden's Adj R2	0.053	0.050	0.003
ML (Cox-Snell) R2	0.101	0.095	0.006
Cragg-Uhler(Nagelkerke) R2	0.120	0.113	0.007
AIC	1.738	1.743	-0.006
AIC*n	4521.450	4535.986	-14.536
BIC	-15876.266	-15867.595	-8.672
BIC'	-206.492	-197.821	-8.672
BIC used by Stata	4585.955	4594.626	-8.672
AIC used by Stata	4521.450	4535.986	-14.536

Difference of 8.672 in BIC' provides strong support for current model.

Note: p-value for difference in LR is only valid if models are nested.

Table A- 27. Fit comparison of NBRM Model 3 and Model 4 on robbery

Measures of Fit for nbreg of rob2007, Current = M4, Saved = M3.

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-2388.359	-2388.359	0.000
Log-Lik Full Model	-2233.135	-2249.725	16.590
D	4466.270(2585)	4499.450(2591)	33.180(6)
LR	310.448(15)	277.268(9)	33.180(6)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.065	0.058	0.007
McFadden's Adj R2	0.058	0.053	0.004
ML (Cox-Snell) R2	0.112	0.101	0.011
Cragg-Uhler(Nagelkerke) R2	0.134	0.120	0.014
AIC	1.730	1.738	-0.008
AIC*n	4500.270	4521.450	-21.180
BIC	-15862.262	-15876.266	14.004
BIC'	-192.488	-206.492	14.004
BIC used by Stata	4599.959	4585.955	14.004
AIC used by Stata	4500.270	4521.450	-21.180

Difference of 14.004 in BIC' provides very strong support for saved model.

Note: p-value for difference in LR is only valid if models are nested.

Table A- 28. Expected percent change on robbery by predictors in NBRM Model 3

Negative binomial regression		Number of obs	= 2602
		LR chi2(6)	= 272.82
Dispersion	= mean	Prob > chi2	= 0.0000
Log likelihood	= -2251.95	Pseudo R2	= 0.0571

rob2007	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wrob	1.213927	.1475209	8.23	0.000	.9247913	1.503063
bstops	.0476788	.0114572	4.16	0.000	.0252231	.0701346
sic58xx	.2204551	.0486397	4.53	0.000	.125123	.3157871
sic55xx	.2629644	.0894422	2.94	0.003	.0876609	.4382679
sic54xx	.3662963	.0752727	4.87	0.000	.2187645	.5138282
sic72xx	.2018992	.047048	4.29	0.000	.1096868	.2941116
_cons	-1.65893	.0860494	-19.28	0.000	-1.827584	-1.490276
/lnalpha	-.3653867	.1304055			-.6209767	-.1097967
alpha	.6939282	.090492			.5374193	.8960162

Likelihood-ratio test of alpha=0: chibar2(01) = 133.33 Prob>=chibar2 = 0.000

nbreg (N=2602): Percentage Change in Expected Count

Observed SD: .88578597

rob2007	b	z	P>|z|	%	%StdX	SDofX
wrob	1.21393	8.229	0.000	236.7	30.2	0.2173
bstops	0.04768	4.161	0.000	4.9	13.1	2.5823
sic58xx	0.22046	4.532	0.000	24.7	14.4	0.6119
sic55xx	0.26296	2.940	0.003	30.1	8.9	0.3233
sic54xx	0.36630	4.866	0.000	44.2	15.5	0.3935
sic72xx	0.20190	4.291	0.000	22.4	14.0	0.6492
ln alpha	-0.36539					
alpha	0.69393	SE(alpha) = 0.09049				

LR test of alpha=0: 133.33 Prob>=LRX2 = 0.000

b = raw coefficient
z = z-score for test of b=0
P>|z| = p-value for z-test
% = percent change in expected count for unit increase in X
%StdX = percent change in expected count for SD increase in X
SDofX = standard deviation of X

Appendix B. Data Analysis Results on Aggravated Assault

Table B- 1. OLS regression model 1 output in GeoDa on aggravated assault

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :      ALT07      Number of Observations: 2602
Mean dependent var  :    0.408532     Number of Variables   :    2
S.D. dependent var  :    0.885169     Degrees of Freedom    : 2600

R-squared           :    0.008767     F-statistic           :     22.9966
Adjusted R-squared  :    0.008386     Prob(F-statistic)     :1.71442e-006
Sum squared residual:     2020.86     Log Likelihood        :    -3363.24
Sigma-square        :    0.777252     Akaike info criterion :     6730.48
S.E. of regression  :    0.881619     Schwarz criterion     :     6742.21
Sigma-square ML     :    0.776655
S.E of regression ML:     0.88128
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                 0.3840789     0.01801985      21.31421     0.0000000
Bus stops                0.03210224    0.006694268      4.795481    0.0000017
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   1.337648
TEST ON NORMALITY OF ERRORS
TEST                     DF          VALUE            PROB
Jarque-Bera               2           43544.64        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF          VALUE            PROB
Breusch-Pagan test        1           70.26626        0.0000000
Koenker-Bassett test      1           6.763561        0.0093038
SPECIFICATION ROBUST TEST
TEST                     DF          VALUE            PROB
White                     2            22.6621        0.0000120

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                           MI/DF      VALUE           PROB
Moran's I (error)            0.079998    18.3200083      0.0000000
Lagrange Multiplier (lag)        1      308.5266948      0.0000000
Robust LM (lag)                  1        8.6858233      0.0032069
Lagrange Multiplier (error)      1      326.5464933      0.0000000
Robust LM (error)                1       26.7056219      0.0000002
Lagrange Multiplier (SARMA)      2      335.2323167      0.0000000
========================= END OF REPORT ==============================
```

Table B- 2. OLS regression model 2 output in GeoDa on aggravated assault

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :      ALT07      Number of Observations: 2602
Mean dependent var  :   0.408532      Number of Variables   :    8
S.D. dependent var  :   0.885169      Degrees of Freedom    : 2594

R-squared           :   0.040605      F-statistic           :     15.6839
Adjusted R-squared  :   0.038016      Prob(F-statistic)     :2.84571e-020
Sum squared residual:    1955.95      Log Likelihood        :    -3320.77
Sigma-square        :   0.754028      Akaike info criterion :     6657.53
S.E. of regression  :   0.868348      Schwarz criterion     :     6704.44
Sigma-square ML     :   0.751709
S.E of regression ML:   0.867012
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                 0.3521871      0.01926379     18.28233     0.0000000
Eating/drinking places   0.04626363     0.03187689      1.451322    0.1468127
Automotive retail        0.03950574     0.05422598      0.728539    0.4663595
Food store retail        0.2770445      0.04662483      5.941995    0.0000000
Automotive service      -0.1112425      0.03666264     -3.034219    0.0024356
Business service        -0.023462       0.03063398     -0.7658818   0.4438122
Personal service         0.1290315      0.02956867      4.363791    0.0000133
Banks                    0.05043761     0.08424538      0.5986988   0.5494229
-------------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   2.207984
TEST ON NORMALITY OF ERRORS
TEST                  DF          VALUE            PROB
Jarque-Bera            2           45362.21        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF          VALUE            PROB
Breusch-Pagan test     7           172.9815        0.0000000
Koenker-Bassett test   7           16.29935        0.0225179
SPECIFICATION ROBUST TEST
TEST                  DF          VALUE            PROB
White                 35            41.9691        0.1943772

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                           MI/DF       VALUE            PROB
Moran's I (error)            0.077106    17.7154250       0.0000000
Lagrange Multiplier (lag)        1      291.8843074       0.0000000
Robust LM (lag)                  1        0.1658382       0.6838373
Lagrange Multiplier (error)      1      303.3597365       0.0000000
Robust LM (error)                1       11.6412673       0.0006450
Lagrange Multiplier (SARMA)      2      303.5255747       0.0000000
========================= END OF REPORT ==============================
```

Table B- 3. OLS regression model 3 output in GeoDa on aggravated assault

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :      ALT07        Number of Observations: 2602
Mean dependent var  :   0.408532        Number of Variables   :    9
S.D. dependent var  :   0.885169        Degrees of Freedom    : 2593

R-squared           :   0.044921        F-statistic           :     15.2449
Adjusted R-squared  :   0.041975        Prob(F-statistic)     :4.59055e-022
Sum squared residual:    1947.15        Log likelihood        :     -3314.9
Sigma-square        :   0.750925        Akaike info criterion :      6647.8
S.E. of regression  :   0.866559        Schwarz criterion     :     6700.58
Sigma-square ML     :   0.748328
S.E of regression ML:   0.865059
-----------------------------------------------------------------------------
    Variable           Coefficient     Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                0.3412892     0.01948594     17.51464      0.0000000
Bus stops               0.02309896    0.006747828     3.423169     0.0006286
Eating/drinking place   0.03660073    0.03193623      1.146057     0.2518761
Automotive retail       0.03317277    0.05414592      0.612655     0.5401758
Food store retail       0.2706363     0.04656644      5.811832     0.0000000
Automotive service     -0.1181403     0.03664258     -3.224126     0.0012793
Business service       -0.02723595    0.03059076     -0.8903327    0.3733842
Personal service        0.1207756     0.02960616      4.079406     0.0000465
Bank                    0.03959592    0.08413151      0.4706432    0.6379420
------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   2.283673
TEST ON NORMALITY OF ERRORS
TEST                  DF          VALUE            PROB
Jarque-Bera            2            44661.8        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF          VALUE            PROB
Breusch-Pagan test     8           187.6177        0.0000000
Koenker-Bassett test   8           17.80171        0.0227632
SPECIFICATION ROBUST TEST
TEST                  DF          VALUE            PROB
White                 44           67.90035        0.0118565

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                           MI/DF       VALUE           PROB
Moran's I (error)            0.078992    18.1652418      0.0000000
Lagrange Multiplier (lag)        1      295.0970312      0.0000000
Robust LM (lag)                  1        0.4436709      0.5053560
Lagrange Multiplier (error)      1      318.3841153      0.0000000
Robust LM (error)                1       23.7307550      0.0000011
Lagrange Multiplier (SARMA)      2      318.8277862      0.0000000
========================= END OF REPORT ==============================
```

Table B- 4. OLS regression model 4 output in GeoDa on aggravated assault

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :       ALT07      Number of Observations: 2602
Mean dependent var  :    0.408532      Number of Variables   :   15
S.D. dependent var  :    0.885169      Degrees of Freedom    : 2587

R-squared           :    0.086622      F-statistic           :     17.5244
Adjusted R-squared  :    0.081679      Prob(F-statistic)     : 3.7097e-042
Sum squared residual:     1862.13      Log Likelihood        :    -3256.82
Sigma-square        :    0.719804      Akaike info criterion :     6543.64
S.E. of regression  :    0.848413      Schwarz criterion     :      6631.6
Sigma-square ML     :    0.715654
S.E of regression ML:    0.845964
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                 0.1962776      0.03174004      6.18391     0.0000000
Bus stops                0.02633111     0.006691735     3.93487     0.0000854
Eating/drinking places   0.03531742     0.03132798      1.127345    0.2597023
Automotive retail        0.04144416     0.05313256      0.7800144   0.4354429
Food store retail        0.2502315      0.04574197      5.470501    0.0000000
Automotive service      -0.1265795      0.03595156     -3.520835    0.0004376
Business service        -0.009988118    0.03003802     -0.3325158   0.7395298
Personal service         0.1309971      0.02908785      4.5035      0.0000070
Banks                    0.05889631     0.08251432      0.7137708   0.4754572
Mixed landuse            0.02210815     0.05330663      0.4147355   0.6784266
Vacant land              0.1297234      0.03495966      3.710659    0.0002111
Public housing           0.5201075      0.0547764       9.495103    0.0000000
Colleges                -0.239252       0.1173018      -2.039627    0.0414869
Grade K-12               0.1314584      0.05575874      2.35763     0.0184666
Parks and cemeteries    -0.03944753     0.04697133     -0.8398216   0.4010741
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   4.446725
TEST ON NORMALITY OF ERRORS
TEST                     DF           VALUE            PROB
Jarque-Bera               2            42836.89        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF           VALUE            PROB
Breusch-Pagan test       14            608.6502        0.0000000
Koenker-Bassett test     14            58.74085        0.0000002
SPECIFICATION ROBUST TEST
TEST                     DF           VALUE            PROB
White                    119            N/A             N/A

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                             MI/DF       VALUE            PROB
Moran's I (error)              0.054667    12.8358632       0.0000000
Lagrange Multiplier (lag)          1      166.6909750       0.0000000
Robust LM (lag)                    1       15.8987183       0.0000668
Lagrange Multiplier (error)        1      152.4885195       0.0000000
Robust LM (error)                  1        1.6962627       0.1927775
Lagrange Multiplier (SARMA)        2      168.3872378       0.0000000
========================= END OF REPORT ==============================
```

Table B- 5. SEM Model 1 output in GeoDa on aggravated assault

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :       ALT07         Number of Observations: 2602
Mean dependent var  :    0.408532         Number of Variables   :    2
S.D. dependent var  :    0.885169         Degree of Freedom     : 2600
Lag coeff. (Lambda) :    0.620004

R-squared           :    0.073309        R-squared (BUSE)      : -
Sq. Correlation     : -                  Log likelihood        :-3292.798956
Sigma-square        :    0.726085        Akaike info criterion : 6589.6
S.E of regression   :    0.852106        Schwarz criterion     : 6601.325983
-------------------------------------------------------------------------
Variable           Coefficient     Std.Error    z-value      Probability
-------------------------------------------------------------------------
Constant           0.3765116      0.04423091       8.512408    0.0000000
Bus stops          0.03690921     0.006635292      5.56256     0.0000000
LAMBDA             0.6200042      0.04763687      13.01522     0.0000000
-------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                    DF     VALUE         PROB
Breusch-Pagan test                      1       86.13307     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                    DF     VALUE         PROB
Likelihood Ratio Test                   1       140.8793     0.0000000
========================= END OF REPORT ======================================
```

Table B- 6. SEM Model 2 output in GeoDa on aggravated assault

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :       ALT07         Number of Observations: 2602
Mean dependent var  :    0.408532         Number of Variables   :    8
S.D. dependent var  :    0.885169         Degree of Freedom     : 2594
Lag coeff. (Lambda) :    0.609642

R-squared           :    0.100236         R-squared (BUSE)      : -
Sq. Correlation     : -                   Log likelihood        :-3253.730837
Sigma-square        :    0.704987         Akaike info criterion : 6523.46
S.E of regression   :    0.839635         Schwarz criterion     : 6570.373958

--------------------------------------------------------------------------
    Variable          Coefficient     Std.Error    z-value      Probability
--------------------------------------------------------------------------
    CONSTANT         0.3426689       0.04307769      7.95467     0.0000000
     SIC58XX         0.05486197      0.03106909      1.765805    0.0774284
     SIC55XX         0.060795        0.0527379       1.152776    0.2490023
     SIC54XX         0.2413022       0.04521382      5.336912    0.0000001
     SIC75XX        -0.08946704      0.03632091     -2.463238    0.0137689
     SIC73XX         0.004676447     0.0298525       0.1566518   0.8755193
     SIC72XX         0.1401855       0.02884713      4.859602    0.0000012
     FINANCE         0.08331194      0.08104678      1.027949    0.3039739
      LAMBDA         0.6096424       0.04848523     12.57378     0.0000000
---------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
```

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	7	182.4901	0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : **TPSW1600ft.GWT**

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	134.0693	0.0000000

========================= END OF REPORT ======================================

Table B- 7. SEM Model 3 output in GeoDa on aggravated assault

============================== BEGINNING OF REPORT============================

Spatial Weight	: **TPSW1600ft.GWT**		
Dependent Variable	: **ALT07**	Number of Observations	: 2602
Mean dependent var	: 0.408532	Number of Variables	: 9
S.D. dependent var	: 0.885169	Degree of Freedom	: 2593
Lag coeff. (Lambda)	: 0.616840		
R-squared	: 0.106487	R-squared (BUSE)	: -
Sq. Correlation	: -	Log likelihood	:-3245.147359
Sigma-square	: 0.700089	Akaike info criterion	: 6508.29
S.E of regression	: 0.836713	Schwarz criterion	: 6561.071039

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	0.3294342	0.04381415	7.518898	0.0000000
BSTOPS	0.02762343	0.006651422	4.153011	0.0000328
SIC58XX	0.04541828	0.0310444	1.463011	0.1434646
SIC55XX	0.05003853	0.05261713	0.9509932	0.3416078
SIC54XX	0.2333275	0.04509195	5.174482	0.0000002
SIC75XX	-0.09831036	0.03626307	-2.711032	0.0067075
SIC73XX	0.003822714	0.02974945	0.128497	0.8977556
SIC72XX	0.1310318	0.02883237	4.544607	0.0000055
FINANCE	0.06817934	0.08083943	0.8433922	0.3990090
LAMBDA	0.6168398	0.04789693	12.87848	0.0000000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	8	197.7441	0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : **TPSW1600ft.GWT**

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	139.504	0.0000000

========================= END OF REPORT ======================================

Table B- 8. SLM Model 4 output in GeoDa on aggravated assault

```
============================== BEGINNING OF REPORT=============================
Dependent Variable  :       ALT07  Number of Observations: 2602
Mean dependent var  :    0.408532  Number of Variables   :   16
S.D. dependent var  :    0.885169  Degrees of Freedom    : 2586
Lag coeff.   (Rho)  :     0.52244

R-squared           :    0.125377  Log likelihood        :    -3211.74
Sq. Correlation     : -            Akaike info criterion :     6455.48
Sigma-square        :    0.685288  Schwarz criterion     :      6549.3
S.E of regression   :    0.827821
-----------------------------------------------------------------------------
    Variable          Coefficient    Std.Error      z-value      Probability
-----------------------------------------------------------------------------
     W_ALT07           0.52244       0.05114153     10.21557       0.0000000
    CONSTANT           0.02384714    0.03611787      0.6602588     0.5090876
      BSTOPS           0.02716805    0.006531753     4.159381      0.0000319
     SIC58XX           0.03736157    0.03057273      1.222055      0.2216868
     SIC55XX           0.05415537    0.05184304      1.044603      0.2962067
     SIC54XX           0.2305167     0.04465302      5.1624        0.0000002
     SIC75XX          -0.1119645     0.03508561     -3.191181      0.0014171
     SIC73XX          -0.0009160637  0.02930902     -0.03125535    0.9750658
     SIC72XX           0.1294674     0.02838283      4.561468      0.0000051
     FINANCE           0.06850288    0.08051342      0.8508257     0.3948661
    MIXEDUSE           0.02804469    0.0520143       0.5391728     0.5897676
   PC1FVLAND           0.08627212    0.03414917      2.526332      0.0115261
    PHOUSING           0.4136716     0.05437869      7.607237      0.0000000
    UNIVCOLL          -0.1809864     0.1145288      -1.58027       0.1140451
     SCHOOLS           0.08704375    0.05444109      1.598861      0.1098514
   OPENSPACE          -0.0654458     0.0458319      -1.427953      0.1533055
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF      VALUE         PROB
Breusch-Pagan test                      14        635.3936     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF      VALUE         PROB
Likelihood Ratio Test                    1        90.15756     0.0000000
========================= END OF REPORT ======================================
```

Table B- 9. PRM Model 1 output in Stata on aggravated assault

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(2)      =     334.73
                                                  Prob > chi2     =     0.0000
Log likelihood = -2217.0403                       Pseudo R2       =     0.0702
------------------------------------------------------------------------------
       alt07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        walt |   1.794713    .096514    18.60   0.000     1.605549    1.983877
      bstops |   .0567224   .0068702     8.26   0.000      .043257    .0701877
       _cons |  -1.813792   .0637188   -28.47   0.000    -1.938678   -1.688905
------------------------------------------------------------------------------
```

Table B- 10. PRM Model 2 output in Stata on aggravated assault

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(8)      =     433.02
                                                  Prob > chi2     =     0.0000
Log likelihood = -2167.8913                       Pseudo R2       =     0.0908
------------------------------------------------------------------------------
       alt07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        walt |   1.740605   .0969258    17.96   0.000     1.550634    1.930576
     sic58xx |   .0624618   .0445929     1.40   0.161    -.0249387    .1498624
     sic55xx |   .1846903   .0907791     2.03   0.042     .0067665    .3626142
     sic54xx |   .3239465   .0552464     5.86   0.000     .2156655    .4322274
     sic75xx |  -.3154092   .0887928    -3.55   0.000    -.4894399   -.1413785
     sic73xx |   .0173685   .0496233     0.35   0.726    -.0798914    .1146285
     sic72xx |   .2245205   .0391944     5.73   0.000     .1477009    .3013402
     finance |   .1549985     .11754     1.32   0.187    -.0753756    .3853726
       _cons |  -1.860474   .0655288   -28.39   0.000    -1.988908    -1.73204
------------------------------------------------------------------------------
```

Table B- 11. PRM Model 3 output in Stata on aggravated assault

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(9)      =     457.34
                                                  Prob > chi2     =     0.0000
Log likelihood = -2155.7325                       Pseudo R2       =     0.0959
------------------------------------------------------------------------------
       alt07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        walt |   1.774684   .0977137    18.16   0.000     1.583169      1.9662
      bstops |   .0477342    .008211     5.81   0.000      .031641    .0638274
     sic58xx |   .0491398   .0454803     1.08   0.280         -.04    .1382796
     sic55xx |   .1601351   .0928361     1.72   0.085    -.0218204    .3420906
     sic54xx |    .311018   .0556995     5.58   0.000      .201849     .420187
     sic75xx |  -.3339892   .0893717    -3.74   0.000    -.5091545   -.1588239
     sic73xx |   .0115017   .0517393     0.22   0.824    -.0899054    .1129088
     sic72xx |   .2082349   .0392462     5.31   0.000     .1313138    .2851561
     finance |   .1224096   .1190031     1.03   0.304    -.1108321    .3556513
       _cons |  -1.908804   .0669311   -28.52   0.000    -2.039987   -1.777621
------------------------------------------------------------------------------
```

Table B- 12. PRM Model 4 output in Stata on aggravated assault

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(15)     =     535.21
                                                  Prob > chi2     =     0.0000
Log likelihood = -2116.7981                       Pseudo R2       =     0.1122
------------------------------------------------------------------------------
       alt07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        walt |   1.511387   .1050471    14.39   0.000     1.305498    1.717275
      bstops |   .0524148   .0082401     6.36   0.000     .0362646    .0685651
     sic58xx |   .0444253   .0460258     0.97   0.334    -.0457836    .1346342
     sic55xx |   .1622634   .0947899     1.71   0.087    -.0235213    .3480482
     sic54xx |   .3044762   .0567339     5.37   0.000     .1932798    .4156727
     sic75xx |    -.35536   .0897602    -3.96   0.000    -.5312869   -.1794332
     sic73xx |   .0385628   .0515531     0.75   0.454    -.0624794    .1396051
     sic72xx |   .2148814   .0398813     5.39   0.000     .1367156    .2930473
     finance |   .1023531   .1234255     0.83   0.407    -.1395565    .3442627
    mixeduse |   .0570354   .0937507     0.61   0.543    -.1267127    .2407835
   pc1fvland |   .1877574   .0698376     2.69   0.007     .0508781    .3246366
    phousing |   .6140455   .0787602     7.80   0.000     .4596784    .7684125
    univcoll |  -.8888076   .3817864    -2.33   0.020    -1.637095   -.1405199
     schools |   .1762128   .0911854     1.93   0.053    -.0025074    .3549329
   openspace |   -.174393   .0913738    -1.91   0.056    -.3534823    .0046964
       _cons |  -1.999962   .0802735   -24.91   0.000    -2.157296   -1.842629
------------------------------------------------------------------------------
```

Table B- 13. NBRM Model 1 output in Stata on aggravated assault

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(2)      =     210.60
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2080.0402                       Pseudo R2       =     0.0482

------------------------------------------------------------------------------
       alt07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        walt |   1.994747   .1466344    13.60   0.000     1.707349    2.282145
      bstops |   .0802457   .0147867     5.43   0.000     .0512644     .109227
       _cons |  -1.934505   .0847088   -22.84   0.000    -2.100532   -1.768479
-------------+----------------------------------------------------------------
    /lnalpha |   .3021508   .1035522                      .0991923    .5051093
-------------+----------------------------------------------------------------
       alpha |   1.352765   .1400818                      1.104279    1.657167
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  274.00 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2185.343    -2080.04      4      4168.08    4191.536
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table B- 14. NBRM Model 2 output in Stata on aggravated assault

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(8)      =     273.86
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2048.4137                       Pseudo R2       =     0.0627
------------------------------------------------------------------------------
       alt07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        walt |   1.978833   .1453916    13.61   0.000      1.69387    2.263795
     sic58xx |   .0922308   .0642258     1.44   0.151    -.0336495    .2181111
     sic55xx |   .1544447   .1223233     1.26   0.207    -.0853045     .394194
     sic54xx |   .4037592   .0898994     4.49   0.000     .2275597    .5799587
     sic75xx |  -.3329942   .1077067    -3.09   0.002    -.5440955   -.1218929
     sic73xx |   .0385213   .0699622     0.55   0.582    -.0986021    .1756447
     sic72xx |   .2485267   .0565668     4.39   0.000     .1376579    .3593956
     finance |   .1256024   .1724492     0.73   0.466    -.2123919    .4635967
       _cons |  -2.003462   .0872415   -22.96   0.000    -2.174453   -1.832472
-------------+----------------------------------------------------------------
    /lnalpha |    .181949   .1085921                     -.0308876    .3947857
-------------+----------------------------------------------------------------
       alpha |   1.199553    .130262                      .9695845    1.484066
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  238.96 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2185.343   -2048.414     10     4116.827    4175.468
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table B- 15. NBRM Model 3 output in Stata on aggravated assault

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(9)      =     285.77
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2042.4579                       Pseudo R2       =     0.0654
------------------------------------------------------------------------------
       alt07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        walt |   1.989232     .14426    13.79   0.000     1.706488    2.271977
      bstops |   .0501421   .0143842     3.49   0.000     .0219496    .0783346
     sic58xx |   .0659707     .06361     1.04   0.300    -.0587027    .1906441
     sic55xx |   .1243962   .1230286     1.01   0.312    -.1167354    .3655279
     sic54xx |   .3948732   .0891236     4.43   0.000     .2201942    .5695523
     sic75xx |   -.339902   .1071157    -3.17   0.002    -.5498449   -.1299592
     sic73xx |   .0212523     .07061     0.30   0.763    -.1171407    .1596453
     sic72xx |   .2168127   .0568739     3.81   0.000      .105342    .3282835
     finance |   .1029347   .1722967     0.60   0.550    -.2347607    .4406301
       _cons |  -2.035302   .0875021   -23.26   0.000    -2.206803   -1.863801
-------------+----------------------------------------------------------------
    /lnalpha |   .1457049    .110759                     -.0713788    .3627885
-------------+----------------------------------------------------------------
       alpha |   1.156855   .1281321                      .9311092    1.437332
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  226.55 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2185.343   -2042.458     11     4106.916     4171.42
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table B- 16. NBRM Model 4 output in Stata on aggravated assault

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(15)     =     340.54
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2015.0728                       Pseudo R2       =     0.0779
------------------------------------------------------------------------------
       alt07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        walt |   1.703214   .1476623    11.53   0.000     1.413801    1.992626
      bstops |   .0589894   .0144692     4.08   0.000     .0306304    .0873485
     sic58xx |   .0579277   .0625834     0.93   0.355    -.0647334    .1805888
     sic55xx |   .1229434   .1228229     1.00   0.317    -.1177852     .363672
     sic54xx |   .3620724   .0875054     4.14   0.000     .1905649    .5335799
     sic75xx |  -.3608237   .1064372    -3.39   0.001    -.5694367   -.1522107
     sic73xx |   .0548693    .070556     0.78   0.437    -.0834178    .1931565
     sic72xx |   .2337546    .056094     4.17   0.000     .1238125    .3436968
     finance |   .1283662    .168959     0.76   0.447    -.2027873    .4595197
    mixeduse |   .0641916   .1211161     0.53   0.596    -.1731916    .3015749
   pc1fvland |   .2198982   .0857967     2.56   0.010     .0517398    .3880566
    phousing |   .6839654   .1069764     6.39   0.000     .4742955    .8936354
    univcoll |  -.8487915   .4211363    -2.02   0.044    -1.674204   -.0233795
     schools |   .1815376   .1199269     1.51   0.130    -.0535148    .4165899
   openspace |  -.1813242   .1126236    -1.61   0.107    -.4020625     .039414
       _cons |  -2.154232   .1024063   -21.04   0.000    -2.354944   -1.953519
-------------+----------------------------------------------------------------
    /lnalpha |     .04542   .1150002                     -.1799761    .2708162
-------------+----------------------------------------------------------------
       alpha |   1.046467   .1203439                      .8352901    1.311034
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  203.45 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2185.343   -2015.073     17     4064.146    4163.834
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table B- 17. Model fit comparison of PRM & NBRM Model 1 on aggravated assault

```
----------------------------------------------------------
                        Variable |    PRM        NBRM
---------------------------------+------------------------
alt07                            |
                            WALT |    6.018       7.350
                                 |    18.60       13.60
                          BSTOPS |    1.058       1.084
                                 |     8.26        5.43
                        Constant |    0.163       0.144
                                 |   -28.47      -22.84
---------------------------------+------------------------
lnalpha                          |
                        Constant |                1.353
                                 |                 2.92
---------------------------------+------------------------
Statistics                       |
                           alpha |                1.353
                               N |  2602.000   2602.000
                              ll | -2217.040  -2080.040
                             bic |  4457.673   4191.536
                             aic |  4440.081   4168.080
```

```
----------------------------------------------------------
                                             legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM         -0.080         1      0.015
NBRM        -0.007         1      0.002

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.738       0.684      0.054    11.062
1        0.164       0.245      0.080    68.188
2        0.054       0.056      0.003     0.294
3        0.017       0.012      0.005     5.295
4        0.006       0.003      0.003     9.215
5        0.004       0.001      0.003    41.176
6        0.001       0.000      0.001     5.729
7        0.002       0.000      0.002   298.018
8        0.000       0.000      0.000    28.596
9        0.000       0.000      0.000     0.009
------------------------------------------------
Sum      0.986       1.000      0.151   467.582

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.738       0.742      0.004     0.057
1        0.164       0.171      0.007     0.677
2        0.054       0.052      0.002     0.131
3        0.017       0.019      0.002     0.527
4        0.006       0.008      0.002     1.433
5        0.004       0.004      0.000     0.024
6        0.001       0.002      0.001     1.700
7        0.002       0.001      0.001     4.068
8        0.000       0.001      0.000     0.209
9        0.000       0.000      0.000     0.963
------------------------------------------------
Sum      0.986       0.999      0.020     9.789

Tests and Fit Statistics

PRM            BIC=-16004.548  AIC=     1.706  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM      BIC=-16270.684  dif=   266.136  NBRM    PRM   Very strong
               AIC=     1.602  dif=     0.105  NBRM    PRM
               LRX2=  274.000  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM           BIC=-16270.684  AIC=     1.602  Prefer  Over  Evidence
```

Table B- 18. Model fit comparison of PRM & NBRM Model 2 on aggravated assault

Variable	PRM	NBRM
alt07		
WALT	5.692	7.204
	17.90	13.59
SIC58XX	1.071	1.104
	1.55	1.55
SIC55XX	1.200	1.165
	2.01	1.25
SIC54XX	1.390	1.512
	6.03	4.64
SIC75XX	0.733	0.718
	-3.50	-3.07
SIC73XX	1.017	1.043
	0.34	0.61
SIC72XX	1.259	1.286
	5.94	4.45
Constant	0.156	0.135
	-28.35	-22.95
lnalpha		
Constant		1.201
		1.69
Statistics		
alpha		1.201
N	2602.000	2602.000
ll	-2168.700	-2048.678
bic	4400.311	4168.132
aic	4353.399	4115.356

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
PRM	-0.074	1	0.014
NBRM	-0.008	1	0.002

PRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.738	0.689	0.049	8.912
1	0.164	0.238	0.074	59.560
2	0.054	0.055	0.002	0.129
3	0.017	0.013	0.004	4.033
4	0.006	0.003	0.003	5.574
5	0.004	0.001	0.003	23.139
6	0.001	0.000	0.000	1.373
7	0.002	0.000	0.002	80.980
8	0.000	0.000	0.000	3.487
9	0.000	0.000	0.000	0.094
Sum	0.986	1.000	0.137	187.282

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson

0	0.738	0.742	0.004	0.063
1	0.164	0.172	0.008	0.861
2	0.054	0.051	0.003	0.335
3	0.017	0.018	0.001	0.296
4	0.006	0.008	0.002	1.258
5	0.004	0.004	0.000	0.022
6	0.001	0.002	0.001	1.832
7	0.002	0.001	0.001	3.387
8	0.000	0.001	0.000	0.321
9	0.000	0.000	0.000	1.117
Sum	0.986	0.999	0.021	9.492

Tests and Fit Statistics

PRM	BIC=-16061.909	AIC= 1.673	Prefer	Over	Evidence
vs NBRM	BIC=-16294.089	dif= 232.179	NBRM	PRM	Very strong
	AIC= 1.582	dif= 0.091	NBRM	PRM	
	LRX2= 240.043	prob= 0.000	NBRM	PRM	p=0.000
NBRM	BIC=-16294.089	AIC= 1.582	Prefer	Over	Evidence

Table B- 19. Model fit comparison of PRM & NBRM Model 3 on aggravated assault

	Variable	PRM	NBRM
alt07			
	WALT	5.892	7.287
		18.12	13.78
	BSTOPS	1.049	1.052
		5.90	3.51
	SIC58XX	1.056	1.073
		1.22	1.12
	SIC55XX	1.170	1.131
		1.69	1.00
	SIC54XX	1.371	1.496
		5.72	4.56
	SIC75XX	0.719	0.713
		-3.70	-3.16
	SIC73XX	1.011	1.024
		0.21	0.34
	SIC72XX	1.236	1.245
		5.44	3.86
	Constant	0.149	0.131
		-28.49	-23.25
lnalpha			
	Constant		1.158
			1.32
Statistics			
	alpha		1.158
	N	2602.000	2602.000
	ll	-2156.233	-2042.636
	bic	4383.243	4163.912
	aic	4330.466	4105.272

legend: b/t

```
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM         -0.072         1      0.013
NBRM        -0.008         1      0.002

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.738       0.690      0.047     8.390
1        0.164       0.236      0.072    56.808
2        0.054       0.056      0.002     0.139
3        0.017       0.013      0.004     3.314
4        0.006       0.003      0.002     4.564
5        0.004       0.001      0.003    20.864
6        0.001       0.000      0.000     1.268
7        0.002       0.000      0.002    84.139
8        0.000       0.000      0.000     4.071
9        0.000       0.000      0.000     0.079
------------------------------------------------
Sum      0.986       1.000      0.133   183.637

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.738       0.741      0.004     0.051
1        0.164       0.173      0.008     1.051
2        0.054       0.051      0.003     0.346
3        0.017       0.018      0.001     0.266
4        0.006       0.008      0.002     1.206
5        0.004       0.004      0.000     0.030
6        0.001       0.002      0.001     1.783
7        0.002       0.001      0.001     3.564
8        0.000       0.001      0.000     0.290
9        0.000       0.000      0.000     1.075
------------------------------------------------
Sum      0.986       0.999      0.021     9.661

Tests and Fit Statistics

PRM             BIC=-16078.978  AIC=     1.664  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM       BIC=-16298.309  dif=   219.331  NBRM    PRM   Very strong
                AIC=     1.578  dif=     0.087  NBRM    PRM
                LRX2=  227.195  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM            BIC=-16298.309  AIC=     1.578  Prefer  Over  Evidence
```

Table B- 20. Model fit comparison of PRM & NBRM Model 4 on aggravated assault

Variable	PRM	NBRM
alt07		
WALT	4.539	5.470
	14.39	11.52
BSTOPS	1.054	1.061
	6.40	4.11
SIC58XX	1.050	1.066
	1.07	1.02
SIC55XX	1.173	1.128
	1.68	0.98
SIC54XX	1.361	1.451
	5.48	4.29
SIC75XX	0.703	0.699
	-3.93	-3.37
SIC73XX	1.038	1.060
	0.73	0.82
SIC72XX	1.245	1.267
	5.55	4.23
MIXEDUSE	1.067	1.070
	0.70	0.56
PC1FVLAND	1.204	1.242
	2.67	2.53
PHOUSING	1.851	1.981
	7.82	6.39
UNIVCOLL	0.413	0.428
	-2.31	-2.01
SCHOOLS	1.191	1.199
	1.92	1.51
OPENSPACE	0.839	0.837
	-1.92	-1.59
Constant	0.135	0.116
	-24.89	-21.03
lnalpha		
Constant		1.047
		0.40
Statistics		
alpha		1.047
N	2602.000	2602.000
ll	-2117.128	-2015.359
bic	4352.216	4156.543
aic	4264.255	4062.718

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
PRM	-0.066	1	0.012
NBRM	-0.008	1	0.002

PRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.738	0.695	0.043	6.916
1	0.164	0.231	0.066	49.783

2	0.054	0.055	0.001	0.092
3	0.017	0.013	0.003	2.262
4	0.006	0.004	0.002	2.727
5	0.004	0.001	0.003	14.580
6	0.001	0.000	0.000	0.563
7	0.002	0.000	0.002	61.085
8	0.000	0.000	0.000	2.656
9	0.000	0.000	0.000	0.106
Sum	0.986	1.000	0.122	140.770

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.738	0.742	0.004	0.062
1	0.164	0.173	0.008	1.055
2	0.054	0.051	0.003	0.537
3	0.017	0.018	0.001	0.197
4	0.006	0.008	0.002	1.215
5	0.004	0.004	0.000	0.014
6	0.001	0.002	0.001	1.934
7	0.002	0.001	0.001	3.062
8	0.000	0.001	0.000	0.373
9	0.000	0.000	0.000	1.170
Sum	0.986	0.999	0.022	9.618

Tests and Fit Statistics

PRM	BIC=-16110.005	AIC= 1.639	Prefer	Over	Evidence
vs NBRM	BIC=-16305.678	dif= 195.673	NBRM	PRM	Very strong
	AIC= 1.561	dif= 0.077	NBRM	PRM	
	LRX2= 203.537	prob= 0.000	NBRM	PRM	p=0.000
NBRM	BIC=-16305.678	AIC= 1.561	Prefer	Over	Evidence

Table B- 21. Model fit comparison of NBRM & ZINBRM model 1 on assault

Variable	NBRM	ZINBRM
alt07		
WALT	7.350	4.230
	13.60	7.62
BSTOPS	1.084	1.055
	5.43	3.37
Constant	0.144	0.230
	-22.84	-9.87
lnalpha		
Constant	1.353	1.091
	2.92	0.62
inflate		
WALT		0.002
		-2.61
BSTOPS		0.041
		-0.45
Constant		2.128

```
                                |                 1.62
--------------------------------+-----------------------
Statistics                      |
                          alpha |     1.353
                              N |  2602.000    2602.000
                             ll | -2080.040   -2069.838
                            bic |  4191.536    4194.723
                            aic |  4168.080    4153.675
--------------------------------------------------------
                                             legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.007         1      0.002
ZINBRM       -0.005         0      0.002

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.738       0.742      0.004     0.057
1        0.164       0.171      0.007     0.677
2        0.054       0.052      0.002     0.131
3        0.017       0.019      0.002     0.527
4        0.006       0.008      0.002     1.433
5        0.004       0.004      0.000     0.024
6        0.001       0.002      0.001     1.700
7        0.002       0.001      0.001     4.068
8        0.000       0.001      0.000     0.209
9        0.000       0.000      0.000     0.963
------------------------------------------------
Sum      0.986       0.999      0.020     9.789

ZINB: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.738       0.742      0.005     0.074
1        0.164       0.167      0.002     0.090
2        0.054       0.056      0.002     0.243
3        0.017       0.020      0.004     1.576
4        0.006       0.008      0.002     1.648
5        0.004       0.003      0.000     0.166
6        0.001       0.002      0.001     0.970
7        0.002       0.001      0.002     8.912
8        0.000       0.000      0.000     0.002
9        0.000       0.000      0.000     0.505
------------------------------------------------
Sum      0.986       1.000      0.018    14.185

Tests and Fit Statistics

-------------------------------------------------------------------------
NBRM           BIC=-16270.684  AIC=     1.602  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-16267.497  dif=    -3.187  NBRM    ZINBRM  Positive
               AIC=     1.596  dif=     0.006  ZINBRM    NBRM
               Vuong=   2.704  prob=    0.003  ZINBRM    NBRM  p=0.003
```

Table B- 22. Model fit comparison of NBRM & ZINBRM model 2 on assault

Variable	NBRM	ZINBRM
alt07		
WALT	7.310	4.091
	13.79	7.31
BSTOPS	1.051	1.117
	3.49	4.68
SIC58XX	1.068	1.035
	1.04	0.55
SIC55XX	1.132	0.974
	1.01	-0.21
SIC54XX	1.484	1.345
	4.43	3.35
SIC75XX	0.712	0.676
	-3.17	-3.52
SIC73XX	1.021	0.947
	0.30	-0.73
SIC72XX	1.242	1.144
	3.81	2.33
FINANCE	1.108	1.067
	0.60	0.35
Constant	0.131	0.211
	-23.26	-9.23
lnalpha		
Constant	1.157	0.870
	1.32	-0.88
inflate		
WALT		0.001
		-2.80
BSTOPS		1.362
		2.63
SIC58XX		0.585
		-1.22
SIC55XX		0.032
		-0.66
SIC54XX		0.133
		-1.78
SIC75XX		0.254
		-1.23
SIC73XX		0.108
		-2.15
SIC72XX		0.383
		-1.68
FINANCE		0.611
		-0.25
Constant		3.172
		2.47
Statistics		
alpha	1.157	
N	2602.000	2602.000
ll	-2042.458	-2017.978
bic	4171.420	4201.100
aic	4106.916	4077.955

legend: b/t

Comparison of Mean Observed and Predicted Count

Maximum At Mean

Model	Difference	Value	|Diff|
NBRM	-0.008	1	0.002
ZINBRM	-0.005	1	0.002

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.738	0.741	0.004	0.051
1	0.164	0.173	0.008	1.054
2	0.054	0.051	0.003	0.347
3	0.017	0.018	0.001	0.265
4	0.006	0.008	0.002	1.205
5	0.004	0.004	0.000	0.030
6	0.001	0.002	0.001	1.785
7	0.002	0.001	0.001	3.552
8	0.000	0.001	0.000	0.292
9	0.000	0.000	0.000	1.077
Sum	0.986	0.999	0.021	9.657

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.738	0.742	0.004	0.060
1	0.164	0.170	0.005	0.404
2	0.054	0.055	0.001	0.036
3	0.017	0.019	0.002	0.799
4	0.006	0.008	0.002	1.159
5	0.004	0.003	0.001	0.214
6	0.001	0.002	0.001	1.121
7	0.002	0.001	0.001	6.673
8	0.000	0.000	0.000	0.044
9	0.000	0.000	0.000	0.741
Sum	0.986	0.999	0.018	11.251

Tests and Fit Statistics

```
-------------------------------------------------------------------------
NBRM           BIC=-16290.801  AIC=     1.578  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-16261.121  dif=   -29.680  NBRM    ZINBRM  Very strong
               AIC=     1.567  dif=     0.011  ZINBRM    NBRM
               Vuong=   4.092  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table B- 23. Model fit comparison of NBRM & ZINBRM model 3 on assault

Variable	NBRM	ZINBRM
alt07		
WALT	7.287	4.085
	13.78	7.31
BSTOPS	1.052	1.119
	3.51	5.49
SIC58XX	1.073	1.039
	1.12	0.62
SIC55XX	1.131	0.971

```
                                 |     1.00       -0.23
                          SIC54XX |     1.496       1.350
                                 |     4.56        3.42
                          SIC75XX |     0.713       0.675
                                 |    -3.16       -3.60
                          SIC73XX |     1.024       0.947
                                 |     0.34       -0.72
                          SIC72XX |     1.245       1.146
                                 |     3.86        2.39
                         Constant |     0.131       0.211
                                 |   -23.25       -9.25
---------------------------------+-----------------------
lnalpha                          |
                         Constant |     1.158       0.866
                                 |     1.32       -0.92
---------------------------------+-----------------------
inflate                          |
                             WALT |                 0.001
                                 |                 -2.84
                           BSTOPS |                 1.363
                                 |                  2.76
                          SIC58XX |                 0.593
                                 |                 -1.33
                          SIC55XX |                 0.029
                                 |                 -0.59
                          SIC54XX |                 0.129
                                 |                 -2.01
                          SIC75XX |                 0.235
                                 |                 -1.38
                          SIC73XX |                 0.109
                                 |                 -2.19
                          SIC72XX |                 0.386
                                 |                 -1.76
                         Constant |                 3.064
                                 |                  2.57
---------------------------------+-----------------------
Statistics                       |
                            alpha |     1.158
                                N |  2602.000    2602.000
                               ll | -2042.636   -2018.172
                              bic |  4163.912    4185.760
                              aic |  4105.272    4074.343
---------------------------------------------------------
                                             legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum        At       Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.008         1      0.002
ZINBRM       -0.005         1      0.002

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.738       0.741      0.004     0.051
1        0.164       0.173      0.008     1.051
2        0.054       0.051      0.003     0.346
3        0.017       0.018      0.001     0.266
4        0.006       0.008      0.002     1.206
5        0.004       0.004      0.000     0.030
6        0.001       0.002      0.001     1.783
```

7	0.002	0.001	0.001	3.564
8	0.000	0.001	0.000	0.290
9	0.000	0.000	0.000	1.075
Sum	0.986	0.999	0.021	9.661

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.738	0.742	0.004	0.059
1	0.164	0.170	0.005	0.414
2	0.054	0.055	0.001	0.035
3	0.017	0.019	0.002	0.790
4	0.006	0.008	0.002	1.149
5	0.004	0.003	0.001	0.218
6	0.001	0.002	0.001	1.116
7	0.002	0.001	0.001	6.700
8	0.000	0.000	0.000	0.043
9	0.000	0.000	0.000	0.740
Sum	0.986	0.999	0.018	11.265

Tests and Fit Statistics

```
-------------------------------------------------------------------------
NBRM           BIC=-16298.309  AIC=     1.578  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-16276.461  dif=   -21.848  NBRM    ZINBRM  Very strong
               AIC=     1.566  dif=     0.012  ZINBRM    NBRM
               Vuong=   4.054  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table B- 24. Model fit comparison of NBRM & ZINBRM model 4 on assault

Variable	NBRM	ZINBRM
alt07		
WALT	5.470	3.480
	11.52	7.50
BSTOPS	1.061	1.117
	4.11	5.66
SIC58XX	1.066	1.040
	1.02	0.65
SIC55XX	1.128	0.987
	0.98	-0.10
SIC54XX	1.451	1.361
	4.29	3.61
SIC75XX	0.699	0.677
	-3.37	-3.63
SIC73XX	1.060	0.988
	0.82	-0.17
SIC72XX	1.267	1.167
	4.23	2.75
MIXEDUSE	1.070	1.049
	0.56	0.37
PC1FVLAND	1.242	1.048
	2.53	0.46
PHOUSING	1.981	1.741
	6.39	5.15
UNIVCOLL	0.428	0.407

```
                                 |     -2.01       -1.64
                         SCHOOLS |     1.199       1.091
                                 |      1.51        0.70
                       OPENSPACE |     0.837       0.817
                                 |     -1.59       -1.69
                        Constant |     0.116       0.200
                                 |    -21.03      -10.24
---------------------------------+-----------------------
lnalpha                          |
                        Constant |     1.047       0.823
                                 |      0.40       -1.42
---------------------------------+-----------------------
inflate                          |
                            WALT |                 0.000
                                 |                 -3.58
                          BSTOPS |                 1.336
                                 |                  2.64
                         SIC58XX |                 0.560
                                 |                 -1.32
                         SIC55XX |                 0.015
                                 |                 -1.09
                         SIC54XX |                 0.353
                                 |                 -0.89
                         SIC75XX |                 0.322
                                 |                 -1.18
                         SIC73XX |                 0.153
                                 |                 -2.00
                         SIC72XX |                 0.367
                                 |                 -1.89
                        MIXEDUSE |                 0.671
                                 |                 -0.40
                       PC1FVLAND |                 0.176
                                 |                 -2.93
                        PHOUSING |                 0.000
                                 |                 -0.02
                        UNIVCOLL |                 0.445
                                 |                 -0.45
                         SCHOOLS |                 0.131
                                 |                 -1.01
                       OPENSPACE |                 0.517
                                 |                 -0.85
                        Constant |                13.738
                                 |                  4.15
---------------------------------+-----------------------
Statistics                       |
                           alpha |     1.047
                               N |  2602.000    2602.000
                              ll | -2015.359   -1986.493
                             bic |  4156.543    4216.770
                             aic |  4062.718    4034.985
---------------------------------------------------------
                                             legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.008         1      0.002
ZINBRM       -0.006         1      0.002

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
```

```
------------------------------------------------
0        0.738       0.742       0.004     0.062
1        0.164       0.173       0.008     1.055
2        0.054       0.051       0.003     0.537
3        0.017       0.018       0.001     0.197
4        0.006       0.008       0.002     1.215
5        0.004       0.004       0.000     0.014
6        0.001       0.002       0.001     1.934
7        0.002       0.001       0.001     3.062
8        0.000       0.001       0.000     0.373
9        0.000       0.000       0.000     1.170
------------------------------------------------
Sum      0.986       0.999       0.022     9.618

ZINB: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.738       0.741       0.004     0.055
1        0.164       0.171       0.006     0.553
2        0.054       0.054       0.000     0.005
3        0.017       0.019       0.002     0.642
4        0.006       0.008       0.002     1.097
5        0.004       0.003       0.001     0.197
6        0.001       0.002       0.001     1.190
7        0.002       0.001       0.001     6.181
8        0.000       0.000       0.000     0.066
9        0.000       0.000       0.000     0.782
------------------------------------------------
Sum      0.986       0.999       0.017    10.768

Tests and Fit Statistics

-------------------------------------------------------------------------
NBRM           BIC=-16305.678  AIC=     1.561  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-16245.450  dif=   -60.228  NBRM    ZINBRM  Very strong
               AIC=     1.551  dif=     0.011  ZINBRM    NBRM
               Vuong=   4.464  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table B- 25. Fit comparisons of NBRM Model 1 and Model 2 on aggravated assault

Measures of Fit for nbreg of alt07, Current = M2, Saved = M1

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-2185.343	-2185.343	0.000
Log-Lik Full Model	-2048.414	-2080.040	31.626
D	4096.827(2592)	4160.080(2598)	63.253(6)
LR	273.858(8)	210.605(2)	63.253(6)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.063	0.048	0.014
McFadden's Adj R2	0.058	0.046	0.012
ML (Cox-Snell) R2	0.100	0.078	0.022
Cragg-Uhler(Nagelkerke) R2	0.123	0.096	0.027
AIC	1.582	1.602	-0.020
AIC*n	4116.827	4168.080	-51.253
BIC	-16286.753	-16270.684	-16.069
BIC'	-210.946	-194.877	-16.069
BIC used by Stata	4175.468	4191.536	-16.069
AIC used by Stata	4116.827	4168.080	-51.253

Difference of 16.069 in BIC' provides very strong support for current model.

Note: p-value for difference in LR is only valid if models are nested.

Table B- 26. Fit comparisons of NBRM Model 2 and Model 3 on aggravated assault

Measures of Fit for nbreg of alt07, Current = M3, Saved = M2

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-2185.343	-2185.343	0.000
Log-Lik Full Model	-2042.458	-2048.414	5.956
D	4084.916(2591)	4096.827(2592)	11.912(1)
LR	285.770(9)	273.858(8)	11.912(1)
Prob > LR	0.000	0.000	0.001
McFadden's R2	0.065	0.063	0.003
McFadden's Adj R2	0.060	0.058	0.002
ML (Cox-Snell) R2	0.104	0.100	0.004
Cragg-Uhler(Nagelkerke) R2	0.128	0.123	0.005
AIC	1.578	1.582	-0.004
AIC*n	4106.916	4116.827	-9.912
BIC	-16290.801	-16286.753	-4.048
BIC'	-214.993	-210.946	-4.048
BIC used by Stata	4171.420	4175.468	-4.048
AIC used by Stata	4106.916	4116.827	-9.912

Difference of 4.048 in BIC' provides positive support for current model.

Note: p-value for difference in LR is only valid if models are nested.

Table B- 27. Fit comparisons of NBRM Model 3 and Model 4 on aggravated assault

Measures of Fit for nbreg of alt07, Current = M4, Saved = M3

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-2185.343	-2185.343	0.000
Log-Lik Full Model	-2015.073	-2042.458	27.385
D	4030.146(2585)	4084.916(2591)	54.770(6)
LR	340.540(15)	285.770(9)	54.770(6)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.078	0.065	0.013
McFadden's Adj R2	0.070	0.060	0.010
ML (Cox-Snell) R2	0.123	0.104	0.019
Cragg-Uhler(Nagelkerke) R2	0.151	0.128	0.023
AIC	1.562	1.578	-0.016
AIC*n	4064.146	4106.916	-42.770
BIC	-16298.387	-16290.801	-7.586
BIC'	-222.579	-214.993	-7.586
BIC used by Stata	4163.834	4171.420	-7.586
AIC used by Stata	4064.146	4106.916	-42.770

Difference of 7.586 in BIC' provides strong support for current model.

Note: p-value for difference in LR is only valid if models are nested.

Table B- 28. Expected percent change on aggravated assault by predictors in NBRM

nbreg (N=2602): Percentage Change in Expected Count

Observed SD: .8853393

alt07	b	z	P>|z|	%	%StdX	SDofX
walt	1.75941	12.176	0.000	480.9	57.9	0.2595
bstops	0.06027	4.200	0.000	6.2	16.8	2.5823
sic54xx	0.41356	4.811	0.000	51.2	17.7	0.3935
sic75xx	-0.32204	-3.098	0.002	-27.5	-14.3	0.4805
sic72xx	0.25812	4.891	0.000	29.4	18.2	0.6492
phousing	0.65226	6.111	0.000	92.0	22.0	0.3050
univcoll	-0.79316	-1.891	0.059	-54.8	-10.8	0.1439
ln alpha	0.07100					
alpha	1.07358	SE(alpha) = 0.12226				

LR test of alpha=0: 209.27 Prob>=LRX2 = 0.000

b = raw coefficient
z = z-score for test of b=0
P>|z| = p-value for z-test
% = percent change in expected count for unit increase in X
%StdX = percent change in expected count for SD increase in X
SDofX = standard deviation of X

Appendix C. Data Analysis Results on Motor Vehicle Theft

Table C- 1. OLS regression Model 1 output in GeoDa on motor vehicle theft

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :      AUTO07      Number of Observations: 2602
Mean dependent var  :     1.62452      Number of Variables   :    2
S.D. dependent var  :     1.91279      Degrees of Freedom    : 2600

R-squared           :    0.011629      F-statistic           :      30.591
Adjusted R-squared  :    0.011249      Prob(F-statistic)     :3.50344e-008
Sum squared residual:     9409.45      Log Likelihood        :    -5364.43
Sigma-square        :     3.61902      Akaike info criterion :     10732.9
S.E. of regression  :     1.90237      Schwarz criterion     :     10744.6
Sigma-square ML     :     3.61624
S.E of regression ML:     1.90164
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                 1.563663       0.03888351      40.21403    0.0000000
Bus stops                0.07989396     0.01444499      5.530909    0.0000000
-------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   1.337648
TEST ON NORMALITY OF ERRORS
TEST                     DF           VALUE            PROB
Jarque-Bera               2            4896.468        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF           VALUE            PROB
Breusch-Pagan test        1            100.5818        0.0000000
Koenker-Bassett test      1            26.63233        0.0000002
SPECIFICATION ROBUST TEST
TEST                     DF           VALUE            PROB
White                     2            33.36351        0.0000001

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                              MI/DF       VALUE          PROB
Moran's I (error)               0.098664    22.5726661      0.0000000
Lagrange Multiplier (lag)          1       476.9845130      0.0000000
Robust LM (lag)                    1         2.3399639      0.1260925
Lagrange Multiplier (error)        1       496.7068036      0.0000000
Robust LM (error)                  1        22.0622545      0.0000026
Lagrange Multiplier (SARMA)        2       499.0467675      0.0000000
========================= END OF REPORT ==============================
```

Table C- 2. OLS regression Model 2 output in GeoDa on motor vehicle theft

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :      AUTO07      Number of Observations: 2602
Mean dependent var  :     1.62452      Number of Variables   :    8
S.D. dependent var  :     1.91279      Degrees of Freedom    : 2594

R-squared           :    0.033840      F-statistic           :     12.9796
Adjusted R-squared  :    0.031233      Prob(F-statistic)     :1.65262e-016
Sum squared residual:     9197.99      Log Likelihood        :    -5334.86
Sigma-square        :     3.54587      Akaike info criterion :     10685.7
S.E. of regression  :     1.88305      Schwarz criterion     :     10732.6
Sigma-square ML     :     3.53497
S.E of regression ML:     1.88015
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                 1.457145       0.04177434     34.88135      0.0000000
Eating/drinking places   0.1579614      0.06912636      2.285111     0.0223865
Automotive retail        0.0521372      0.1175913       0.4433763    0.6575735
Food store retail        0.5578811      0.1011079       5.517681     0.0000000
Automotive service       0.1196586      0.07950448      1.505055     0.1324289
Business service         0.09619279     0.06643106      1.448009     0.1477363
Personal service         0.1784833      0.06412089      2.783544     0.0054160
Banks                   -0.03251342     0.1826896      -0.1779708    0.8587096
-----------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   2.207984
TEST ON NORMALITY OF ERRORS
TEST                  DF           VALUE            PROB
Jarque-Bera            2            5177.612        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF           VALUE            PROB
Breusch-Pagan test     7            111.3321        0.0000000
Koenker-Bassett test   7            28.67748        0.0001656
SPECIFICATION ROBUST TEST
TEST                  DF           VALUE            PROB
White                 35            90.43012        0.0000009

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                         MI/DF      VALUE           PROB
Moran's I (error)          0.091247    20.9442306      0.0000000
Lagrange Multiplier (lag)       1      425.7449419      0.0000000
Robust LM (lag)                 1        5.8637880      0.0154555
Lagrange Multiplier (error)     1      424.8365818      0.0000000
Robust LM (error)               1        4.9554279      0.0260089
Lagrange Multiplier (SARMA)     2      430.7003698      0.0000000
========================= END OF REPORT ==============================
```

Table C- 3. OLS regression Model 3 output in GeoDa on motor vehicle theft

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :      AUTO07      Number of Observations: 2602
Mean dependent var  :     1.62452      Number of Variables   :    9
S.D. dependent var  :     1.91279      Degrees of Freedom    : 2593

R-squared           :    0.039245      F-statistic           :       13.24
Adjusted R-squared  :    0.036281      Prob(F-statistic)     :6.69113e-019
Sum squared residual:     9146.53      Log Likelihood        :    -5327.56
Sigma-square        :     3.52739      Akaike info criterion :     10673.1
S.E. of regression  :     1.87814      Schwarz criterion     :     10725.9
Sigma-square ML     :     3.51519
S.E of regression ML:     1.87489
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                 1.430793       0.04223282     33.87869      0.0000000
Bus stops                0.05585689     0.0146249       3.819302     0.0001370
Eating/drinking places   0.134595       0.06921694      1.944538     0.0519371
Automotive retail        0.03682307     0.1173531       0.3137802    0.7537166
Food store retail        0.5423851      0.1009257       5.374102     0.0000001
Automotive service       0.1029787      0.07941725      1.296679     0.1948525
Business service         0.08706679     0.06630083      1.313208     0.1892199
Personal service         0.1585193      0.06416688      2.470422     0.0135592
Banks                   -0.0587303      0.1823423      -0.3220882    0.7474476
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   2.283673
TEST ON NORMALITY OF ERRORS
TEST                     DF         VALUE            PROB
Jarque-Bera               2          4818.962        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF         VALUE            PROB
Breusch-Pagan test        8          207.9491        0.0000000
Koenker-Bassett test      8          55.23928        0.0000000
SPECIFICATION ROBUST TEST
TEST                     DF         VALUE            PROB
White                    44          133.2957        0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                            MI/DF        VALUE            PROB
Moran's I (error)             0.093578    21.4980328       0.0000000
Lagrange Multiplier (lag)         1      429.3066385       0.0000000
Robust LM (lag)                   1        1.0588004       0.3034890
Lagrange Multiplier (error)       1      446.8150857       0.0000000
Robust LM (error)                 1       18.5672475       0.0000164
Lagrange Multiplier (SARMA)       2      447.8738861       0.0000000
========================= END OF REPORT ==============================
```

Table C- 4. OLS regression Model 4 output in GeoDa on motor vehicle theft

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :      AUTO07      Number of Observations: 2602
Mean dependent var  :     1.62452      Number of Variables   :   15
S.D. dependent var  :     1.91279      Degrees of Freedom    : 2587

R-squared           :    0.055168      F-statistic           :     10.7896
Adjusted R-squared  :    0.050055      Prob(F-statistic)     :2.67185e-024
Sum squared residual:     8994.94      Log Likelihood        :    -5305.82
Sigma-square        :     3.47698      Akaike info criterion :     10641.6
S.E. of regression  :     1.86467      Schwarz criterion     :     10729.6
Sigma-square ML     :     3.45693
S.E of regression ML:     1.85928
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                 1.324525      0.06975919     18.9871       0.0000000
Bus stops                0.06061439    0.01470729      4.121384     0.0000388
Eating/drinking places   0.1227509     0.06885354      1.782782     0.0747387
Automotive retail        0.03816439    0.1167763       0.3268162    0.7437706
Food store retail        0.5088944     0.100533        5.061962     0.0000004
Automotive service       0.09533688    0.07901539      1.206561     0.2277061
Business service         0.09784195    0.06601844      1.48204      0.1384555
Personal service         0.1426614     0.06393012      2.231521     0.0257325
Banks                   -0.06928206    0.1813524      -0.38203      0.7025366
Mixed landuse            0.2767783     0.1171589       2.362419     0.0182300
Vacant land              0.1490171     0.07683537      1.939434     0.0525560
Public housing          -0.2753086     0.1203892      -2.286822     0.0222872
Colleges                -0.6720935     0.2578094      -2.606939     0.0091879
Grade K-12               0.5389582     0.1225482       4.397929     0.0000114
Parks and cemeteries    -0.1587856     0.103235       -1.538099     0.1241508
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   4.446725
TEST ON NORMALITY OF ERRORS
TEST                     DF          VALUE            PROB
Jarque-Bera               2           4855.262        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF          VALUE            PROB
Breusch-Pagan test       14           226.3399        0.0000000
Koenker-Bassett test     14           59.83387        0.0000001
SPECIFICATION ROBUST TEST
TEST                     DF          VALUE            PROB
White                    119           N/A             N/A

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                            MI/DF      VALUE           PROB
Moran's I (error)             0.087321    20.3761350      0.0000000
Lagrange Multiplier (lag)        1       391.0440990      0.0000000
Robust LM (lag)                  1        10.0992778      0.0014833
Lagrange Multiplier (error)      1       389.0665803      0.0000000
Robust LM (error)                1         8.1217590      0.0043737
Lagrange Multiplier (SARMA)      2       399.1658580      0.0000000
========================= END OF REPORT ==============================
```

Table C- 5. SEM Model 1 output in GeoDa on motor vehicle theft

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :      AUTO07      Number of Observations: 2602
Mean dependent var  :    1.624520      Number of Variables   :    2
S.D. dependent var  :    1.912795      Degree of Freedom     : 2600
Lag coeff. (Lambda) :    0.669099

R-squared           :    0.095702      R-squared (BUSE)      : -
Sq. Correlation     : -                Log Likelihood        :-5269.600145
Sigma-square        :    3.308630      Akaike info criterion : 0543.2
S.E of regression   :     1.81896      Schwarz criterion     : 10554.928362
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Constant                 1.522368       0.1082592      14.06224      0.0000000
Bus stops                0.08740776     0.01416989      6.168555     0.0000000
LAMBDA                   0.6690995      0.04348492     15.38693      0.0000000
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF     VALUE         PROB
Breusch-Pagan test                       1       76.74729     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF     VALUE         PROB
Likelihood Ratio Test                    1       189.6544     0.0000000
========================= END OF REPORT ======================================
```

Table C- 6. SLM Model 2 output in GeoDa on motor vehicle theft

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :      AUTO07      Number of Observations: 2602
Mean dependent var  :     1.62452      Number of Variables   :    9
S.D. dependent var  :     1.91279      Degrees of Freedom    : 2593
Lag coeff.   (Rho)  :    0.639459

R-squared           :    0.107857      Log Likelihood        :    -5249.71
Sq. Correlation     : -                Akaike info criterion :     10517.4
Sigma-square        :     3.26416      Schwarz criterion     :     10570.2
S.E of regression   :      1.8067
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Spatial lag              0.6394588     0.04454482     14.3554        0.0000000
Constant                 0.4200166     0.08050832      5.217058      0.0000002
Eating/drinking places   0.1504602     0.06637926      2.266675      0.0234100
Automotive retail        0.08778918    0.1128238       0.7781084     0.4365050
Food store retail        0.472274      0.09701563      4.86802       0.0000011
Automotive service       0.1184379     0.07632825      1.551691      0.1207362
Business service         0.119605      0.06375829      1.875913      0.0606671
Personal service         0.142786      0.06153388      2.320445      0.0203168
Banks                   -0.01097763    0.1752875      -0.06262645    0.9500638
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
```

```
TEST                                        DF     VALUE          PROB
Breusch-Pagan test                          7       134.9337      0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                        DF     VALUE          PROB
Likelihood Ratio Test                       1       170.2965      0.0000000
========================= END OF REPORT =======================================
```

Table C- 7. SEM Model 3 output in GeoDa on motor vehicle theft

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :      AUTO07       Number of Observations: 2602
Mean dependent var  :    1.624520       Number of Variables   :    9
S.D. dependent var  :    1.912795       Degree of Freedom     : 2593
Lag coeff. (Lambda) :    0.656169

R-squared           :    0.115982       R-squared (BUSE)      : -
Sq. Correlation     : -                 Log Likelihood        :-5239.068237
Sigma-square        :    3.234432       Akaike info criterion :10496.1
S.E of regression   :     1.79845       Schwarz criterion     :10548.912796
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                 1.401489       0.104465       13.41588      0.0000000
Bus stops                0.06779554     0.01429952      4.741105     0.0000021
Eating/drinking places   0.1450621      0.0667158       2.17433      0.0296803
Automotive retail        0.09043266     0.1130573       0.7998832    0.4237783
Food store retail        0.4244243      0.09688837      4.38055      0.0000118
Automotive service       0.135834       0.07798405      1.741818     0.0815402
Business service         0.1409503      0.0639378       2.20449      0.0274898
Personal service         0.1110615      0.06196951      1.792196     0.0731015
Banks                   -0.03260133     0.1736451      -0.1877469    0.8510751
LAMBDA                   0.6561692      0.04460062     14.71211      0.0000000
-------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                        DF     VALUE          PROB
Breusch-Pagan test                          8       194.2545      0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                        DF     VALUE          PROB
Likelihood Ratio Test                       1       176.9801      0.0000000
========================= END OF REPORT =======================================
```

Table C- 8. SEM Model 4 output in GeoDa on motor vehicle theft

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :      AUTO07  Number of Observations: 2602
Mean dependent var  :    1.624520  Number of Variables   :   15
S.D. dependent var  :    1.912795  Degree of Freedom     : 2587
Lag coeff. (Lambda) :    0.648936

R-squared           :    0.126121  R-squared (BUSE)      : -
Sq. Correlation     : -            Log likelihood        :-5223.506023
Sigma-square        :    3.197336  Akaike info criterion :      10477
S.E of regression   :     1.78811  Schwarz criterion     :10564.972582

-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    z-value      Probability
-----------------------------------------------------------------------
    CONSTANT     1.271128      0.1168418      10.87905    0.0000000
      BSTOPS    0.0688834     0.01433696      4.804605    0.0000016
     SIC58XX    0.1334867     0.06647743         2.008    0.0446431
     SIC55XX    0.1008492      0.1125944     0.8956859    0.3704205
     SIC54XX    0.4044951     0.09652412      4.190612    0.0000278
     SIC75XX    0.1190973     0.07769303      1.532921    0.1252953
     SIC73XX    0.1404077     0.06361282      2.207223    0.0272984
     SIC72XX    0.1039268     0.06170405      1.684279    0.0921277
     FINANCE   -0.03668571      0.1728541    -0.2122352    0.8319236
    MIXEDUSE    0.2815263       0.117218      2.401733    0.0163176
   PC1FVLAND    0.1740516     0.07859895      2.214427    0.0267993
    PHOUSING   -0.05920298      0.1345203    -0.4401044    0.6598615
    UNIVCOLL   -0.2355281      0.3011031    -0.7822174    0.4340867
     SCHOOLS    0.4599226      0.1195758      3.846286    0.0001200
   OPENSPACE   -0.2102705       0.108365     -1.940391    0.0523320
      LAMBDA    0.6489358     0.04521763      14.35139    0.0000000
-----------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF     VALUE         PROB
Breusch-Pagan test                      14       219.3957     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF     VALUE         PROB
Likelihood Ratio Test                    1       164.6188     0.0000000
========================= END OF REPORT =======================================
```

Table C- 9. SLM Model 4 output in GeoDa on motor vehicle theft

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :      AUTO07       Number of Observations: 2602
Mean dependent var  :     1.62452       Number of Variables   :   16
S.D. dependent var  :     1.91279       Degrees of Freedom    : 2586
Lag coeff.   (Rho)  :    0.631722

R-squared           :    0.125163       Log Likelihood        :    -5223.66
Sq. Correlation     : -                 Akaike info criterion :     10479.3
Sigma-square        :     3.20084       Schwarz criterion     :     10573.1
S.E of regression   :     1.78909
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
```

```
------------------------------------------------------------------------------
Spatial lag              0.6317223      0.04413697     14.31277       0.0000000
Constant                 0.2774102      0.09598893      2.890023      0.0038523
Bus stops                0.06259321     0.01411983      4.433001      0.0000093
Eating/drinking places   0.1150947      0.0661151       1.740823      0.0817146
Automotive retail        0.0762706      0.1120435       0.6807233     0.4960464
Food store retail        0.4259091      0.09647726      4.414607      0.0000101
Automotive service       0.09117948     0.0758288       1.202439      0.2291937
Business service         0.1193346      0.06335573      1.883565      0.0596238
Personal service         0.1134885      0.06134284      1.850069      0.0643035
Banks                   -0.04626835     0.1740042      -0.2659036     0.7903136
Mixed landuse            0.2724526      0.1124567       2.422733      0.0154042
Vacant land              0.1546408      0.07374882      2.096858      0.0360060
Public housing          -0.1067994      0.1155899      -0.9239509     0.3555118
Colleges                -0.3709669      0.2473611      -1.499698      0.1336928
Grade K-12               0.4804256      0.1176206       4.084538      0.0000442
Parks and cemeteries    -0.1665901      0.09914169     -1.680324      0.0928943
------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF      VALUE          PROB
Breusch-Pagan test                      14        225.944      0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF      VALUE          PROB
Likelihood Ratio Test                    1       164.3086      0.0000000
========================= END OF REPORT ======================================
```

Table C- 10. PRM Model 1 output in Stata on motor vehicle theft

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(2)      =     470.96
                                                  Prob > chi2     =     0.0000
Log likelihood = -4749.6726                       Pseudo R2       =     0.0472
------------------------------------------------------------------------------
      auto07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       wauto |   .4038536   .0168102    24.02   0.000     .3709061    .4368011
      bstops |   .0322448   .0039087     8.25   0.000     .0245839    .0399058
       _cons |  -.2455546   .0350706    -7.00   0.000    -.3142916   -.1768175
------------------------------------------------------------------------------
```

Table C- 11. PRM Model 2 output in Stata on motor vehicle theft

Poisson regression — Number of obs = 2602; LR chi2(8) = 520.15; Prob > chi2 = 0.0000; Pseudo R2 = 0.0522

Log likelihood = -4725.0783

auto07	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wauto	.3722304	.0173364	21.47	0.000	.3382516	.4062091
sic58xx	.0763885	.0243838	3.13	0.002	.0285971	.1241799
sic55xx	.0619864	.043836	1.41	0.157	-.0239306	.1479034
sic54xx	.1516524	.0339224	4.47	0.000	.0851657	.2181392
sic75xx	.0739705	.0303679	2.44	0.015	.0144505	.1334905
sic73xx	.0611571	.0214883	2.85	0.004	.0190409	.1032734
sic72xx	.0458213	.0232479	1.97	0.049	.0002561	.0913864
finance	.0346976	.064571	0.54	0.591	-.0918593	.1612545
_cons	-.2387387	.0353388	-6.76	0.000	-.3080015	-.1694759

Table C- 12. PRM Model 3 output in Stata on motor vehicle theft

Poisson regression — Number of obs = 2602; LR chi2(9) = 550.03; Prob > chi2 = 0.0000; Pseudo R2 = 0.0552

Log likelihood = -4710.1421

auto07	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wauto	.3748432	.0173688	21.58	0.000	.3408009	.4088854
bstops	.0261577	.0042949	6.09	0.000	.0177398	.0345756
sic58xx	.0669355	.0246232	2.72	0.007	.0186749	.1151961
sic55xx	.0543391	.0443303	1.23	0.220	-.0325467	.141225
sic54xx	.1453686	.0339988	4.28	0.000	.0787322	.212005
sic75xx	.0627254	.0305834	2.05	0.040	.0027831	.1226678
sic73xx	.0601198	.0218261	2.75	0.006	.0173414	.1028982
sic72xx	.0362126	.0232839	1.56	0.120	-.009423	.0818482
finance	.0216906	.0651917	0.33	0.739	-.1060827	.149464
_cons	-.2585265	.0355914	-7.26	0.000	-.3282843	-.1887688

Table C- 13. PRM Model 4 output in Stata on motor vehicle theft

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(15)     =     617.46
                                                  Prob > chi2     =     0.0000
Log likelihood = -4676.4251                       Pseudo R2       =     0.0619
------------------------------------------------------------------------------
      auto07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       wauto |   .3704311    .017813    20.80   0.000     .3355181     .405344
      bstops |   .0270725    .004285     6.32   0.000     .0186741    .0354709
     sic58xx |    .058329   .0247938     2.35   0.019     .0097342    .1069239
     sic55xx |   .0521327   .0450534     1.16   0.247    -.0361703    .1404357
     sic54xx |   .1290828   .0344158     3.75   0.000     .0616291    .1965366
     sic75xx |   .0543681   .0307611     1.77   0.077    -.0059226    .1146588
     sic73xx |   .0695455   .0219131     3.17   0.002     .0265966    .1124944
     sic72xx |   .0331124   .0233556     1.42   0.156    -.0126637    .0788885
     finance |    .010233   .0665712     0.15   0.878    -.1202442    .1407102
    mixeduse |   .1757712   .0457052     3.85   0.000     .0861907    .2653517
   pc1fvland |   .0883597   .0329708     2.68   0.007      .023738    .1529814
    phousing |  -.0750925   .0555865    -1.35   0.177    -.1840399     .033855
    univcoll |  -.2899677   .1311683    -2.21   0.027    -.5470528   -.0328826
     schools |   .2584405   .0459652     5.62   0.000     .1683504    .3485307
   openspace |  -.1154764   .0456243    -2.53   0.011    -.2048984   -.0260544
       _cons |  -.3258915   .0436102    -7.47   0.000    -.4113659   -.2404172
------------------------------------------------------------------------------
```

Table C- 14. NBRM Model 1 output in Stata on motor vehicle theft

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(2)      =     238.62
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -4404.7415                       Pseudo R2       =     0.0264
------------------------------------------------------------------------------
      auto07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       wauto |   .5033949   .0347809    14.47   0.000     .4352257    .5715642
      bstops |     .03901   .0084359     4.62   0.000     .0224759     .055544
       _cons |   -.427422   .0645574    -6.62   0.000    -.5539523   -.3008918
-------------+----------------------------------------------------------------
    /lnalpha |  -.5329449   .0653141                     -.6609581   -.4049317
-------------+----------------------------------------------------------------
       alpha |   .5868741   .0383311                      .5163564    .6670224
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  689.86 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -4524.053   -4404.741      4     8817.483    8840.939
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table C- 15. NBRM Model 2 output in Stata on motor vehicle theft

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(8)      =     284.89
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -4381.606                        Pseudo R2       =     0.0315
------------------------------------------------------------------------------
      auto07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       wauto |    .494173   .0345491    14.30   0.000     .4264581    .5618879
     sic58xx |   .0902676   .0376994     2.39   0.017      .016378    .1641571
     sic55xx |   .0745058   .0668797     1.11   0.265     -.056576    .2055875
     sic54xx |   .1969991   .0545336     3.61   0.000     .0901153    .3038829
     sic75xx |   .0744106   .0456111     1.63   0.103    -.0149855    .1638068
     sic73xx |   .0793253   .0381076     2.08   0.037     .0046357    .1540148
     sic72xx |   .0962801   .0341759     2.82   0.005     .0292966    .1632636
     finance |  -.0188342   .0990016    -0.19   0.849    -.2128738    .1752055
       _cons |  -.4792408   .0649446    -7.38   0.000      -.60653   -.3519517
-------------+----------------------------------------------------------------
    /lnalpha |  -.5702033   .0661152                     -.6997866   -.4406199
-------------+----------------------------------------------------------------
       alpha |   .5654105   .0373822                      .4966913    .6436373
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  686.94 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -4524.053   -4381.606     10     8783.212    8841.852
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table C- 16. NBRM Model 3 output in Stata on motor vehicle theft

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(9)      =     295.11
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -4376.4962                       Pseudo R2       =     0.0326
------------------------------------------------------------------------------
      auto07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       wauto |    .492989   .0344703    14.30   0.000     .4254284    .5605496
      bstops |   .0248779   .0080594     3.09   0.002     .0090817    .0406741
     sic58xx |   .0799849   .0376489     2.12   0.034     .0061943    .1537754
     sic55xx |   .0631023   .0666242     0.95   0.344    -.0674787    .1936833
     sic54xx |   .1891056   .0543239     3.48   0.000     .0826327    .2955786
     sic75xx |   .0662728   .0455655     1.45   0.146    -.0230339    .1555795
     sic73xx |   .0756471   .0377299     2.00   0.045     .0016979    .1495963
     sic72xx |   .0872883   .0341895     2.55   0.011     .0202781    .1542985
     finance |  -.0395061    .098811    -0.40   0.689    -.2331721    .1541599
       _cons |  -.4902554   .0649471    -7.55   0.000    -.6175494   -.3629614
-------------+----------------------------------------------------------------
    /lnalpha |  -.5834063   .0666692                     -.7140757    -.452737
-------------+----------------------------------------------------------------
       alpha |   .5579944   .0372011                      .4896445    .6358853
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  667.29 Prob>=chibar2 = 0.000
----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -4524.053   -4376.496     11     8774.992    8839.497
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table C- 17. NBRM Model 4 output in Stata on motor vehicle theft

Negative binomial regression		Number of obs =	2602
		LR chi2(15) =	325.96
Dispersion = mean		Prob > chi2 =	0.0000
Log likelihood = -4361.074		Pseudo R2 =	0.0360

auto07	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wauto	.4812051	.0346986	13.87	0.000	.413197	.5492132
bstops	.0274302	.0082916	3.31	0.001	.0111789	.0436814
sic58xx	.066259	.0376129	1.76	0.078	-.007461	.1399789
sic55xx	.0595522	.0663023	0.90	0.369	-.0703979	.1895022
sic54xx	.1765434	.0539821	3.27	0.001	.0707404	.2823463
sic75xx	.0560485	.0454079	1.23	0.217	-.0329493	.1450462
sic73xx	.0829645	.0374667	2.21	0.027	.0095312	.1563978
sic72xx	.0872561	.0341286	2.56	0.011	.0203652	.154147
finance	-.0280908	.0982697	-0.29	0.775	-.2206957	.1645142
mixeduse	.1762824	.0659679	2.67	0.008	.0469876	.3055771
pc1fvland	.1056715	.0456671	2.31	0.021	.0161657	.1951773
phousing	-.0420058	.0745056	-0.56	0.573	-.1880341	.1040224
univcoll	-.1785098	.1687008	-1.06	0.290	-.5091573	.1521377
schools	.2345366	.0681091	3.44	0.001	.1010452	.3680281
openspace	-.1105502	.062413	-1.77	0.077	-.2328775	.0117771
_cons	-.5605797	.0747891	-7.50	0.000	-.7071636	-.4139959
/lnalpha	-.6167156	.0678447			-.7496887	-.4837425
alpha	.5397142	.0366167			.4725136	.6164719

Likelihood-ratio test of alpha=0: chibar2(01) = 630.70 Prob>=chibar2 = 0.000

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	2602	-4524.053	-4361.074	17	8756.148	8855.837

Note: N=Obs used in calculating BIC; see [R] BIC note

Table C- 18. Model fit comparison of PRM & NBRM Model 1 on motor vehicle theft

	Variable	PRM	NBRM
auto07			
	WAUTO	1.498	1.654
		24.02	14.47
	BSTOPS	1.033	1.040
		8.25	4.62
	Constant	0.782	0.652
		-7.00	-6.62
lnalpha			
	Constant		0.587
			-8.16
Statistics			
	alpha		0.587
	N	2602.000	2602.000
	ll	-4749.673	-4404.741
	bic	9522.937	8840.939
	aic	9505.345	8817.483

```
----------------------------------------------------------
                                                legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM          0.123         0      0.036
NBRM        -0.029         1      0.005

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.341       0.218      0.123   180.329
1        0.234       0.318      0.083    56.948
2        0.161       0.243      0.082    71.968
3        0.089       0.131      0.042    35.143
4        0.055       0.057      0.002     0.176
5        0.029       0.021      0.007     6.773
6        0.019       0.008      0.012    46.551
7        0.010       0.003      0.008    57.420
8        0.005       0.001      0.004    33.619
9        0.002       0.000      0.001     7.562
------------------------------------------------
Sum      0.944       0.999      0.364   496.490

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.341       0.339      0.001     0.016
1        0.234       0.264      0.029     8.416
2        0.161       0.166      0.005     0.396
3        0.089       0.098      0.009     2.071
4        0.055       0.056      0.001     0.079
5        0.029       0.032      0.003     0.715
6        0.019       0.018      0.001     0.160
7        0.010       0.011      0.000     0.004
8        0.005       0.006      0.002     1.035
9        0.002       0.004      0.002     3.323
------------------------------------------------
Sum      0.944       0.993      0.054    16.215

Tests and Fit Statistics

PRM             BIC=-10939.283  AIC=     3.653  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM       BIC=-11621.282  dif=   681.998  NBRM    PRM   Very strong
                AIC=     3.389  dif=     0.264  NBRM    PRM
                LRX2=  689.862  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM            BIC=-11621.282  AIC=     3.389  Prefer  Over  Evidence
```

Table C- 19. Model fit comparison of PRM & NBRM Model 2 on motor vehicle theft

Variable	PRM	NBRM
auto07		
WAUTO	1.450	1.639
	21.49	14.31
SIC58XX	1.081	1.093
	3.24	2.39
SIC55XX	1.063	1.078
	1.39	1.13
SIC54XX	1.165	1.216
	4.54	3.61
SIC75XX	1.078	1.077
	2.47	1.62
SIC73XX	1.063	1.082
	2.86	2.07
SIC72XX	1.048	1.100
	2.02	2.81
Constant	0.789	0.619
	-6.74	-7.38
lnalpha		
Constant		0.565
		-8.62
Statistics		
alpha		0.565
N	2602.000	2602.000
ll	-4725.220	-4381.624
bic	9513.352	8834.024
aic	9466.440	8781.248

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
PRM	0.121	0	0.036
NBRM	-0.030	1	0.005

PRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.341	0.219	0.121	174.360
1	0.234	0.318	0.083	56.679
2	0.161	0.242	0.081	69.985
3	0.089	0.130	0.041	34.103
4	0.055	0.057	0.002	0.182
5	0.029	0.022	0.007	6.137
6	0.019	0.008	0.011	42.742
7	0.010	0.003	0.008	51.024
8	0.005	0.001	0.003	28.446
9	0.002	0.000	0.001	6.150
Sum	0.944	0.999	0.359	469.808

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.341	0.339	0.001	0.009

1	0.234	0.265	0.030	8.967
2	0.161	0.166	0.005	0.343
3	0.089	0.097	0.008	1.707
4	0.055	0.055	0.001	0.016
5	0.029	0.031	0.003	0.539
6	0.019	0.018	0.001	0.208
7	0.010	0.011	0.000	0.005
8	0.005	0.006	0.002	1.163
9	0.002	0.004	0.002	3.625
Sum	0.944	0.992	0.052	16.583

Tests and Fit Statistics

PRM	BIC=-10948.868	AIC= 3.638	Prefer	Over	Evidence
vs NBRM	BIC=-11628.196	dif= 679.328	NBRM	PRM	Very strong
	AIC= 3.375	dif= 0.263	NBRM	PRM	
	LRX2= 687.192	prob= 0.000	NBRM	PRM	p=0.000
NBRM	BIC=-11628.196	AIC= 3.375	Prefer	Over	Evidence

Table C- 20. Model fit comparison of PRM & NBRM Model 3 on motor vehicle theft

	Variable	PRM	NBRM
auto07			
	WAUTO	1.454	1.638
		21.61	14.31
	BSTOPS	1.027	1.025
		6.11	3.07
	SIC58XX	1.070	1.081
		2.80	2.09
	SIC55XX	1.055	1.067
		1.21	0.97
	SIC54XX	1.158	1.205
		4.32	3.46
	SIC75XX	1.065	1.067
		2.07	1.43
	SIC73XX	1.062	1.077
		2.76	1.98
	SIC72XX	1.037	1.090
		1.58	2.53
	Constant	0.773	0.612
		-7.26	-7.56
lnalpha			
	Constant		0.558
			-8.75
Statistics			
	alpha		0.558
	N	2602.000	2602.000
	ll	-4710.197	-4376.576
	bic	9491.170	8831.792
	aic	9438.394	8773.152

legend: b/t

Comparison of Mean Observed and Predicted Count

Maximum	At	Mean

Model	Difference	Value	|Diff|
PRM	0.120	0	0.035
NBRM	-0.031	1	0.005

PRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.341	0.221	0.120	169.090
1	0.234	0.318	0.083	56.607
2	0.161	0.241	0.080	68.371
3	0.089	0.129	0.040	33.004
4	0.055	0.056	0.002	0.147
5	0.029	0.022	0.007	5.956
6	0.019	0.008	0.011	40.781
7	0.010	0.003	0.007	47.024
8	0.005	0.001	0.003	24.552
9	0.002	0.001	0.001	4.545
Sum	0.944	0.999	0.355	450.078

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.341	0.339	0.002	0.021
1	0.234	0.265	0.031	9.302
2	0.161	0.166	0.005	0.396
3	0.089	0.097	0.008	1.728
4	0.055	0.055	0.000	0.012
5	0.029	0.031	0.002	0.493
6	0.019	0.018	0.001	0.245
7	0.010	0.010	0.000	0.002
8	0.005	0.006	0.002	1.109
9	0.002	0.004	0.002	3.568
Sum	0.944	0.992	0.054	16.875

Tests and Fit Statistics

PRM	BIC=-10971.051	AIC= 3.627	Prefer	Over	Evidence
vs NBRM	BIC=-11630.428	dif= 659.378	NBRM	PRM	Very strong
	AIC= 3.372	dif= 0.256	NBRM	PRM	
	LRX2= 667.242	prob= 0.000	NBRM	PRM	p=0.000
NBRM	BIC=-11630.428	AIC= 3.372	Prefer	Over	Evidence

Table C- 21. Model fit comparison of PRM & NBRM Model 4 on motor vehicle theft

	Variable	PRM	NBRM
auto07			
	WAUTO	1.448	1.618
		20.83	13.87
	BSTOPS	1.027	1.028
		6.33	3.30
	SIC58XX	1.061	1.067

	2.39	1.74
SIC55XX	1.053	1.063
	1.15	0.92
SIC54XX	1.138	1.191
	3.78	3.26
SIC75XX	1.056	1.057
	1.78	1.22
SIC73XX	1.072	1.086
	3.18	2.20
SIC72XX	1.034	1.090
	1.44	2.54
MIXEDUSE	1.193	1.192
	3.87	2.67
PC1FVLAND	1.092	1.112
	2.68	2.33
PHOUSING	0.928	0.959
	-1.35	-0.56
UNIVCOLL	0.748	0.836
	-2.21	-1.06
SCHOOLS	1.295	1.264
	5.62	3.44
OPENSPACE	0.891	0.895
	-2.53	-1.78
Constant	0.722	0.570
	-7.48	-7.51
lnalpha		
Constant		0.540
		-9.09
Statistics		
alpha		0.540
N	2602.000	2602.000
ll	-4676.437	-4361.115
bic	9470.834	8848.054
aic	9382.874	8754.230

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
PRM	0.116	0	0.034
NBRM	-0.032	1	0.005

PRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.341	0.225	0.116	154.491
1	0.234	0.316	0.081	54.632
2	0.161	0.237	0.076	63.575
3	0.089	0.128	0.039	31.361
4	0.055	0.057	0.002	0.212
5	0.029	0.022	0.006	4.771
6	0.019	0.008	0.011	35.678
7	0.010	0.003	0.007	41.015
8	0.005	0.001	0.003	20.981
9	0.002	0.001	0.001	3.553
Sum	0.944	0.999	0.343	410.271

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.341	0.338	0.002	0.038
1	0.234	0.266	0.032	9.779
2	0.161	0.166	0.005	0.435
3	0.089	0.097	0.008	1.682
4	0.055	0.055	0.000	0.005
5	0.029	0.031	0.002	0.444
6	0.019	0.018	0.001	0.275
7	0.010	0.010	0.000	0.000
8	0.005	0.006	0.002	1.091
9	0.002	0.004	0.002	3.557
Sum	0.944	0.992	0.055	17.306

Tests and Fit Statistics

```
PRM            BIC=-10991.387  AIC=     3.606  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM      BIC=-11614.167  dif=   622.780  NBRM    PRM   Very strong
               AIC=     3.364  dif=     0.242  NBRM    PRM
               LRX2=  630.644  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM           BIC=-11614.167  AIC=     3.364  Prefer  Over  Evidence
```

Table C- 22. Model fit comparison of NBRM & ZINBRM model 1 on motor vehicle theft

```
----------------------------------------------------------
                        Variable |   NBRM        ZINBRM
---------------------------------+------------------------
auto07                           |
                           WAUTO |    1.657       1.608
                                 |    14.46       10.02
                        Constant |    0.673       0.717
                                 |    -6.15       -3.37
---------------------------------+------------------------
lnalpha                          |
                        Constant |    0.601       0.586
                                 |    -7.88       -6.55
---------------------------------+------------------------
inflate                          |
                           WAUTO |                0.004
                                 |                -1.03
                        Constant |                6.654
                                 |                 0.75
---------------------------------+------------------------
Statistics                       |
                           alpha |    0.601
                               N |  2602.000    2602.000
                              ll | -4416.718   -4411.372
                             bic |  8857.028    8862.064
                             aic |  8839.436    8832.744
----------------------------------------------------------
                                                legend: b/t
```

Comparison of Mean Observed and Predicted Count

```
            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
```

```
NBRM       -0.028         1      0.005
ZINBRM       -0.025         1      0.005

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.341       0.340      0.001     0.006
1        0.234       0.263      0.028     7.873
2        0.161       0.166      0.005     0.339
3        0.089       0.098      0.009     2.132
4        0.055       0.056      0.002     0.118
5        0.029       0.032      0.003     0.848
6        0.019       0.018      0.001     0.104
7        0.010       0.011      0.000     0.015
8        0.005       0.006      0.002     1.111
9        0.002       0.004      0.002     3.383
------------------------------------------------
Sum      0.944       0.993      0.052    15.930

ZINB: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.341       0.341      0.000     0.001
1        0.234       0.259      0.025     6.179
2        0.161       0.166      0.005     0.391
3        0.089       0.099      0.010     2.609
4        0.055       0.057      0.002     0.253
5        0.029       0.032      0.004     1.061
6        0.019       0.019      0.001     0.070
7        0.010       0.011      0.000     0.017
8        0.005       0.006      0.002     1.048
9        0.002       0.004      0.002     3.216
------------------------------------------------
Sum      0.944       0.993      0.051    14.844

Tests and Fit Statistics

-------------------------------------------------------------------------
NBRM           BIC=-11605.193  AIC=     3.397  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-11600.157  dif=    -5.036  NBRM    ZINBRM  Positive
               AIC=     3.395  dif=     0.003  ZINBRM    NBRM
               Vuong=   2.347  prob=    0.009  ZINBRM    NBRM  p=0.009
```

Table C- 23. Model fit comparison of NBRM & ZINBRM model 2 on motor vehicle theft

```
----------------------------------------------------------
                      Variable |   NBRM       ZINBRM
---------------------------------+------------------------
auto07                           |
                           WAUTO |     1.639       1.484
                                 |     14.30       11.20
                         SIC58XX |     1.094       1.073
                                 |      2.39        1.95
                         SIC55XX |     1.077       1.037
                                 |      1.11        0.58
                         SIC54XX |     1.218       1.181
                                 |      3.61        3.25
                         SIC75XX |     1.077       1.035
```

```
                                |      1.63        0.79
                        SIC73XX |     1.083       1.061
                                |      2.08        1.67
                        SIC72XX |     1.101       1.071
                                |      2.82        2.11
                        FINANCE |     0.981       0.990
                                |     -0.19       -0.11
                       Constant |     0.619       0.818
                                |     -7.38       -2.68
--------------------------------+-----------------------
lnalpha                         |
                       Constant |     0.565       0.452
                                |     -8.62       -8.91
--------------------------------+-----------------------
inflate                         |
                          WAUTO |                 0.132
                                |                 -4.10
                        SIC58XX |                 0.332
                                |                 -0.69
                        SIC55XX |                 0.002
                                |                 -0.31
                        SIC54XX |                 0.000
                                |                 -0.00
                        SIC75XX |                 0.001
                                |                 -0.21
                        SIC73XX |                 0.003
                                |                 -0.32
                        SIC72XX |                 0.000
                                |                 -0.00
                        FINANCE |                 0.873
                                |                 -0.08
                       Constant |                 2.712
                                |                  1.87
--------------------------------+-----------------------
Statistics                      |
                          alpha |     0.565
                              N |  2602.000    2602.000
                             ll | -4381.606   -4360.555
                            bic |  8841.852    8870.526
                            aic |  8783.212    8759.110
--------------------------------------------------------
                                              legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.030         1      0.005
ZINBRM       -0.014         3      0.005

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.341       0.339      0.001     0.009
1        0.234       0.265      0.030     8.965
2        0.161       0.166      0.005     0.343
3        0.089       0.097      0.008     1.706
4        0.055       0.055      0.001     0.016
5        0.029       0.031      0.003     0.539
6        0.019       0.018      0.001     0.208
7        0.010       0.011      0.000     0.005
8        0.005       0.006      0.002     1.162
```

9	0.002	0.004	0.002	3.625
Sum	0.944	0.992	0.052	16.577

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.341	0.349	0.008	0.531
1	0.234	0.244	0.009	0.913
2	0.161	0.167	0.006	0.620
3	0.089	0.102	0.014	4.709
4	0.055	0.059	0.005	1.001
5	0.029	0.034	0.005	1.731
6	0.019	0.019	0.000	0.024
7	0.010	0.011	0.000	0.009
8	0.005	0.006	0.001	0.832
9	0.002	0.003	0.002	2.779
Sum	0.944	0.994	0.051	13.147

Tests and Fit Statistics

```
-------------------------------------------------------------------------
NBRM           BIC=-11620.368  AIC=     3.376  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-11591.695  dif=   -28.674  NBRM    ZINBRM  Very strong
               AIC=     3.366  dif=     0.009  ZINBRM    NBRM
               Vuong=   3.677  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table C- 24. Model fit comparison of NBRM & ZINBRM model 3 on motor vehicle theft

	Variable	NBRM	ZINBRM
auto07			
	WAUTO	1.638	1.486
		14.31	11.35
	BSTOPS	1.025	1.019
		3.07	2.65
	SIC58XX	1.081	1.060
		2.09	1.66
	SIC55XX	1.067	1.028
		0.97	0.44
	SIC54XX	1.205	1.169
		3.46	3.11
	SIC75XX	1.067	1.024
		1.43	0.56
	SIC73XX	1.077	1.057
		1.98	1.58
	SIC72XX	1.090	1.061
		2.53	1.84
	Constant	0.612	0.815
		-7.56	-2.80
lnalpha			
	Constant	0.558	0.436
		-8.75	-9.37
inflate			

```
                         WAUTO |                 0.171
                               |                 -4.09
                        BSTOPS |                 0.496
                               |                 -1.21
                       SIC58XX |                 0.292
                               |                 -0.68
                       SIC55XX |                 0.004
                               |                 -0.38
                       SIC54XX |                 0.000
                               |                 -0.00
                       SIC75XX |                 0.002
                               |                 -0.34
                       SIC73XX |                 0.006
                               |                 -0.35
                       SIC72XX |                 0.000
                               |                 -0.00
                      Constant |                 2.358
                               |                  1.70
-------------------------------+-----------------------
Statistics                     |
                         alpha |     0.558
                             N |  2602.000    2602.000
                            ll | -4376.576   -4353.046
                           bic |  8831.792    8855.509
                           aic |  8773.152    8744.092
-------------------------------------------------------
                                            legend: b/t
```

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
NBRM	-0.031	1	0.005
ZINBRM	-0.014	3	0.005

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.341	0.339	0.002	0.021
1	0.234	0.265	0.031	9.302
2	0.161	0.166	0.005	0.396
3	0.089	0.097	0.008	1.728
4	0.055	0.055	0.000	0.012
5	0.029	0.031	0.002	0.493
6	0.019	0.018	0.001	0.245
7	0.010	0.010	0.000	0.002
8	0.005	0.006	0.002	1.109
9	0.002	0.004	0.002	3.568
Sum	0.944	0.992	0.054	16.875

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.341	0.349	0.009	0.568
1	0.234	0.243	0.008	0.729
2	0.161	0.168	0.007	0.700
3	0.089	0.103	0.014	4.977
4	0.055	0.060	0.005	1.071
5	0.029	0.034	0.005	1.739
6	0.019	0.019	0.000	0.030

```
7        0.010       0.010      0.000     0.004
8        0.005       0.006      0.001     0.773
9        0.002       0.003      0.002     2.694
------------------------------------------------
Sum      0.944       0.994      0.051    13.286
```

Tests and Fit Statistics

```
-------------------------------------------------------------------------
NBRM           BIC=-11630.428  AIC=     3.372  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-11606.712  dif=   -23.716  NBRM    ZINBRM  Very strong
               AIC=     3.361  dif=     0.011  ZINBRM    NBRM
               Vuong=   3.860  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table C- 25. Model fit comparison of NBRM & ZINBRM model 4 on motor vehicle theft

	Variable	NBRM	ZINBRM
auto07			
	WAUTO	1.618	1.514
		13.87	10.93
	BSTOPS	1.028	1.018
		3.30	2.34
	SIC58XX	1.067	1.052
		1.74	1.44
	SIC55XX	1.063	1.032
		0.92	0.50
	SIC54XX	1.191	1.159
		3.26	2.99
	SIC75XX	1.057	1.019
		1.22	0.44
	SIC73XX	1.086	1.058
		2.20	1.62
	SIC72XX	1.090	1.058
		2.54	1.78
	MIXEDUSE	1.192	1.123
		2.67	1.76
	PC1FVLAND	1.112	1.028
		2.33	0.57
	PHOUSING	0.959	0.958
		-0.56	-0.53
	UNIVCOLL	0.836	0.861
		-1.06	-0.78
	SCHOOLS	1.264	1.184
		3.44	2.45
	OPENSPACE	0.895	1.076
		-1.78	0.98
	Constant	0.570	0.758
		-7.51	-3.11
lnalpha			
	Constant	0.540	0.405
		-9.09	-9.98
inflate			
	WAUTO		0.359
			-2.83
	BSTOPS		0.501

```
                                 |               -1.69
                         SIC58XX |               0.352
                                 |               -0.55
                         SIC55XX |               0.207
                                 |               -0.72
                         SIC54XX |               0.000
                                 |               -0.00
                         SIC75XX |               0.054
                                 |               -0.55
                         SIC73XX |               0.122
                                 |               -0.77
                         SIC72XX |               0.000
                                 |               -0.00
                        MIXEDUSE |               0.238
                                 |               -1.29
                       PC1FVLAND |               0.443
                                 |               -2.40
                        PHOUSING |               0.843
                                 |               -0.35
                        UNIVCOLL |               1.404
                                 |                0.30
                         SCHOOLS |               0.268
                                 |               -1.37
                       OPENSPACE |               4.813
                                 |                4.54
                        Constant |               1.223
                                 |                0.31
---------------------------------+----------------------
Statistics                       |
                           alpha |     0.540
                               N |  2602.000    2602.000
                              ll | -4361.115   -4325.999
                             bic |  8848.054    8895.783
                             aic |  8754.230    8713.998
--------------------------------------------------------
                                            legend: b/t
Comparison of Mean Observed and Predicted Count

           Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.032         1      0.005
ZINBRM       -0.014         3      0.005

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.341       0.338      0.002     0.038
1        0.234       0.266      0.032     9.779
2        0.161       0.166      0.005     0.435
3        0.089       0.097      0.008     1.682
4        0.055       0.055      0.000     0.005
5        0.029       0.031      0.002     0.444
6        0.019       0.018      0.001     0.275
7        0.010       0.010      0.000     0.000
8        0.005       0.006      0.002     1.091
9        0.002       0.004      0.002     3.557
------------------------------------------------
Sum      0.944       0.992      0.055    17.306

ZINB: Predicted and actual probabilities
```

Count	Actual	Predicted	|Diff|	Pearson
0	0.341	0.350	0.009	0.652
1	0.234	0.242	0.007	0.552
2	0.161	0.168	0.007	0.777
3	0.089	0.103	0.014	5.126
4	0.055	0.060	0.005	1.064
5	0.029	0.033	0.005	1.669
6	0.019	0.019	0.001	0.044
7	0.010	0.010	0.000	0.001
8	0.005	0.006	0.001	0.752
9	0.002	0.003	0.002	2.698
Sum	0.944	0.994	0.051	13.335

Tests and Fit Statistics

```
-------------------------------------------------------------------------
NBRM           BIC=-11614.167  AIC=     3.364  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-11566.438  dif=   -47.729  NBRM    ZINBRM  Very strong
               AIC=     3.349  dif=     0.015  ZINBRM    NBRM
               Vuong=   4.702  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table C- 26. Fit comparisons of Model 1 and Model 2 on motor vehicle theft

Measures of Fit for nbreg of auto07, Current = M2, Saved = M1

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-4524.053	-4524.053	0.000
Log-Lik Full Model	-4381.606	-4404.741	23.136
D	8763.212(2592)	8809.483(2598)	46.271(6)
LR	284.895(8)	238.624(2)	46.271(6)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.031	0.026	0.005
McFadden's Adj R2	0.029	0.025	0.004
ML (Cox-Snell) R2	0.104	0.088	0.016
Cragg-Uhler(Nagelkerke) R2	0.107	0.090	0.017
AIC	3.376	3.389	-0.013
AIC*n	8783.212	8817.483	-34.271
BIC	-11620.368	-11621.282	0.913
BIC'	-221.983	-222.896	0.913
BIC used by Stata	8841.852	8840.939	0.913
AIC used by Stata	8783.212	8817.483	-34.271

Difference of   0.913 in BIC' provides weak support for saved model.

Note: p-value for difference in LR is only valid if models are nested.

Table C- 27. Fit comparisons of Model 1 and Model 3 on motor vehicle theft

Measures of Fit for nbreg of auto07, Current = M3, Saved = M1

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-4524.053	-4524.053	0.000
Log-Lik Full Model	-4376.496	-4404.741	28.245
D	8752.992(2591)	8809.483(2598)	56.491(7)
LR	295.115(9)	238.624(2)	56.491(7)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.033	0.026	0.006
McFadden's Adj R2	0.030	0.025	0.005
ML (Cox-Snell) R2	0.107	0.088	0.020
Cragg-Uhler(Nagelkerke) R2	0.111	0.090	0.020
AIC	3.372	3.389	-0.016
AIC*n	8774.992	8817.483	-42.491
BIC	-11622.724	-11621.282	-1.442
BIC'	-224.338	-222.896	-1.442
BIC used by Stata	8839.497	8840.939	-1.442
AIC used by Stata	8774.992	8817.483	-42.491

Difference of 1.442 in BIC' provides weak support for current model.

Note: p-value for difference in LR is only valid if models are nested.

Table C- 28. Fit comparisons of Model 3 and Model 4 on motor vehicle theft

Measures of Fit for nbreg of auto07, Current = M4, Saved = M3

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-4524.053	-4524.053	0.000
Log-Lik Full Model	-4361.074	-4376.496	15.422
D	8722.148(2585)	8752.992(2591)	30.844(6)
LR	325.959(15)	295.115(9)	30.844(6)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.036	0.033	0.003
McFadden's Adj R2	0.032	0.030	0.002
ML (Cox-Snell) R2	0.118	0.107	0.011
Cragg-Uhler(Nagelkerke) R2	0.121	0.111	0.011
AIC	3.365	3.372	-0.007
AIC*n	8756.148	8774.992	-18.844
BIC	-11606.384	-11622.724	16.340
BIC'	-207.998	-224.338	16.340
BIC used by Stata	8855.837	8839.497	16.340
AIC used by Stata	8756.148	8774.992	-18.844

Difference of 16.340 in BIC' provides very strong support for saved model.

Note: p-value for difference in LR is only valid if models are nested.

Table C- 29. Expected percent change on motor vehicle theft by predictors in NBRM

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(6)      =     291.26
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -4378.4252                       Pseudo R2       =     0.0322

------------------------------------------------------------------------------
      auto07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       wauto |    .491023   .0344575    14.25   0.000     .4234876    .5585585
      bstops |   .0259204   .0080241     3.23   0.001     .0101934    .0416473
     sic58xx |   .0840793   .0372081     2.26   0.024     .0111527    .1570059
     sic54xx |   .1850158   .0540881     3.42   0.001      .079005    .2910266
     sic73xx |   .0813595   .0375505     2.17   0.030     .0077618    .1549571
     sic72xx |   .0857121   .0340956     2.51   0.012     .0188861    .1525382
       _cons |  -.4755698   .0644536    -7.38   0.000    -.6018966   -.3492431
-------------+----------------------------------------------------------------
    /lnalpha |  -.5799487   .0665649                     -.7104136   -.4494838
-------------+----------------------------------------------------------------
       alpha |   .5599271   .0372715                      .4914409    .6379574
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  670.55 Prob>=chibar2 = 0.000
```

```
nbreg (N=2602): Percentage Change in Expected Count

 Observed SD: 1.9131625

----------------------------------------------------------------------
      auto07 |      b         z     P>|z|      %      %StdX      SDofX
-------------+--------------------------------------------------------
       wauto |   0.49102   14.250   0.000     63.4     35.3     0.6161
      bstops |   0.02592    3.230   0.001      2.6      6.9     2.5823
     sic58xx |   0.08408    2.260   0.024      8.8      5.3     0.6119
     sic54xx |   0.18502    3.421   0.001     20.3      7.6     0.3935
     sic73xx |   0.08136    2.167   0.030      8.5      4.8     0.5776
     sic72xx |   0.08571    2.514   0.012      8.9      5.7     0.6492
-------------+--------------------------------------------------------
    ln alpha |  -0.57995
       alpha |   0.55993   SE(alpha) = 0.03727
----------------------------------------------------------------------
 LR test of alpha=0: 670.55   Prob>=LRX2 = 0.000
----------------------------------------------------------------------
       b = raw coefficient
       z = z-score for test of b=0
   P>|z| = p-value for z-test
       % = percent change in expected count for unit increase in X
   %StdX = percent change in expected count for SD increase in X
   SDofX = standard deviation of X
```

Appendix D. Data Analysis Results on Theft From Motor Vehicle

Table D- 1. OLS regression Model 1 output on theft from vehicle theft

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :       TFA07       Number of Observations: 2602
Mean dependent var  :    0.936587       Number of Variables   :    2
S.D. dependent var  :     1.56373       Degrees of Freedom    : 2600

R-squared           :    0.006697       F-statistic           :     17.5301
Adjusted R-squared  :    0.006315       Prob(F-statistic)     :2.92201e-005
Sum squared residual:     6319.93       Log Likelihood        :    -4846.62
Sigma-square        :     2.43074       Akaike info criterion :     9697.24
S.E. of regression  :     1.55908       Schwarz criterion     :     9708.97
Sigma-square ML     :     2.42887
S.E of regression ML:     1.55848
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                 0.8988318      0.03186687      28.20584    0.0000000
Bus stops                0.04956589     0.01183835       4.18689    0.0000292
-------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   1.337648
TEST ON NORMALITY OF ERRORS
TEST                     DF          VALUE            PROB
Jarque-Bera               2           333273.5        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF          VALUE            PROB
Breusch-Pagan test        1           8.745729        0.0031033
Koenker-Bassett test      1          0.3088871        0.5783639
SPECIFICATION ROBUST TEST
TEST                     DF          VALUE            PROB
White                     2           5.410065        0.0668681

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                              MI/DF       VALUE           PROB
Moran's I (error)               0.081735    18.7155860       0.0000000
Lagrange Multiplier (lag)          1       354.9213663       0.0000000
Robust LM (lag)                    1        20.0092516       0.0000077
Lagrange Multiplier (error)        1       340.8748091       0.0000000
Robust LM (error)                  1         5.9626943       0.0146117
Lagrange Multiplier (SARMA)        2       360.8840607       0.0000000
========================= END OF REPORT ==============================
```

Table D- 2. OLS regression Model 2 output on theft from vehicle theft

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :       TFA07       Number of Observations: 2602
Mean dependent var  :    0.936587       Number of Variables   :    8
S.D. dependent var  :     1.56373       Degrees of Freedom    : 2594

R-squared           :    0.019652       F-statistic           :     7.42834
Adjusted R-squared  :    0.017006       Prob(F-statistic)     :7.21413e-009
Sum squared residual:      6237.5       Log Likelihood        :    -4829.54
Sigma-square        :     2.40459       Akaike info criterion :     9675.08
S.E. of regression  :     1.55067       Schwarz criterion     :        9722
Sigma-square ML     :      2.3972
S.E of regression ML:     1.54829
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                 0.8248344      0.0344008      23.97719      0.0000000
Eating/drinking places   0.1334509      0.05692495      2.34433      0.0191353
Automotive retail        0.1261232      0.0968354       1.302449     0.1928880
Food store retail        0.2661877      0.08326145      3.197011     0.0014053
Automotive service       0.1065138      0.06547123      1.62688      0.1038827
Business service         0.03304397     0.05470539      0.604035     0.5459126
Personal service         0.1099954      0.05280298      2.083129     0.0373371
Banks                    0.06203442     0.1504433       0.4123442    0.6800685
-------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   2.207984
TEST ON NORMALITY OF ERRORS
TEST                     DF           VALUE            PROB
Jarque-Bera               2            322009.9        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF           VALUE            PROB
Breusch-Pagan test        7            277.2323        0.0000000
Koenker-Bassett test      7            9.955547        0.1911066
SPECIFICATION ROBUST TEST
TEST                     DF           VALUE            PROB
White                    35            112.5923        0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                             MI/DF        VALUE           PROB
Moran's I (error)              0.082578    18.9648806       0.0000000
Lagrange Multiplier (lag)          1      348.3752366       0.0000000
Robust LM (lag)                    1        3.1032029       0.0781384
Lagrange Multiplier (error)        1      347.9472149       0.0000000
Robust LM (error)                  1        2.6751812       0.1019237
Lagrange Multiplier (SARMA)        2      351.0504178       0.0000000
========================= END OF REPORT ==============================
```

Table D- 3. OLS regression Model 3 output on theft from vehicle theft

Data set : **IncTPAllDataDec31SW1600**

Dependent Variable	: **TFA07**	Number of Observations	: 2602
Mean dependent var	: 0.936587	Number of Variables	: 9
S.D. dependent var	: 1.56373	Degrees of Freedom	: 2593

R-squared	: 0.022515	F-statistic	: 7.46571
Adjusted R-squared	: 0.019499	Prob(F-statistic)	:6.94547e-010
Sum squared residual	: 6219.29	Log likelihood	: -4825.74
Sigma-square	: 2.39849	Akaike info criterion	: 9669.47
S.E. of regression	: 1.54871	Schwarz criterion	: 9722.25
Sigma-square ML	: 2.39019		
S.E of regression ML	: 1.54603		

Variable	**Coefficient**	**Std.Error**	**t-Statistic**	**Probability**
CONSTANT	0.8091543	0.03482507	23.23482	0.0000000
BSTOPS	0.03323548	0.01205965	2.755924	0.0058938
SIC58XX	0.1195476	0.05707611	2.09453	0.0363100
SIC55XX	0.1170111	0.09676904	1.209179	0.2267002
SIC54XX	0.2569675	0.08322308	3.087695	0.0020385
SIC75XX	0.09658909	0.06548725	1.47493	0.1403468
SIC73XX	0.0276139	0.05467149	0.5050877	0.6135909
SIC72XX	0.09811656	0.05291184	1.85434	0.0638020
FINANCE	0.04643508	0.150359	0.3088281	0.7574906

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 2.283673

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	323032.2	0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	8	279.2838	0.0000000
Koenker-Bassett test	8	10.01336	0.2640895

SPECIFICATION ROBUST TEST

TEST	DF	VALUE	PROB
White	44	119.812	0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

FOR WEIGHT MATRIX : **TPSW1600ft.GWT** (row-standardized weights)

TEST	**MI/DF**	**VALUE**	**PROB**
Moran's I (error)	0.080370	18.4800588	0.0000000
Lagrange Multiplier (lag)	1	336.8999878	0.0000000
Robust LM (lag)	1	8.0087742	0.0046551
Lagrange Multiplier (error)	1	329.5872597	0.0000000
Robust LM (error)	1	0.6960460	0.4041155
Lagrange Multiplier (SARMA)	2	337.5960339	0.0000000

========================= END OF REPORT ==============================

Table D- 4. OLS regression Model 4 output on theft from vehicle theft

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :       TFA07      Number of Observations: 2602
Mean dependent var  :    0.936587      Number of Variables   :   15
S.D. dependent var  :     1.56373      Degrees of Freedom    : 2587

R-squared           :    0.031444      F-statistic           :     5.99906
Adjusted R-squared  :    0.026203      Prob(F-statistic)     :8.37186e-012
Sum squared residual:     6162.47      Log Likelihood        :     -4813.8
Sigma-square        :     2.38209      Akaike info criterion :     9657.59
S.E. of regression  :      1.5434      Schwarz criterion     :     9745.56
Sigma-square ML     :     2.36836
S.E of regression ML:     1.53895
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                 0.6908888      0.05774039     11.96543      0.0000000
Bus stops                0.03141053     0.01217337      2.580265     0.0099272
Eating/drinking places   0.1089209      0.05699077      1.911202     0.0560894
Automotive retail        0.1111778      0.09665692      1.150232     0.2501479
Food store retail        0.2462308      0.08321221      2.95907      0.0031136
Automotive service       0.08602604     0.06540183      1.315346     0.1885058
Business service         0.02763145     0.05464413      0.5056619    0.6131439
Personal service         0.09969401     0.05291561      1.884019     0.0596762
Banks                    0.03271197     0.1501072       0.2179241    0.8275006
Mixed landuse            0.3176857      0.09697359      3.276002     0.0010668
Vacant land              0.1311368      0.06359741      2.061984     0.0393097
Public housing          -0.1646193      0.09964733     -1.65202      0.0986475
Colleges                 0.1439265      0.2133914       0.6744718    0.5000484
Grade K-12               0.196567       0.1014344       1.937873     0.0527480
Parks and cemeteries     0.003852445    0.08544862      0.04508493   0.9642922
-------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   4.446725
TEST ON NORMALITY OF ERRORS
TEST                  DF          VALUE            PROB
Jarque-Bera            2           333530.9        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF          VALUE            PROB
Breusch-Pagan test    14           333.0184        0.0000000
Koenker-Bassett test  14           11.75339        0.6260997
SPECIFICATION ROBUST TEST
TEST                  DF          VALUE            PROB
White                 119            N/A            N/A

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                          MI/DF      VALUE           PROB
Moran's I (error)           0.078045    18.2341010      0.0000000
Lagrange Multiplier (lag)       1      320.0806606      0.0000000
Robust LM (lag)                 1       10.5836533      0.0011409
Lagrange Multiplier (error)     1      310.7943790      0.0000000
Robust LM (error)               1        1.2973716      0.2546939
Lagrange Multiplier (SARMA)     2      321.3780323      0.0000000
========================= END OF REPORT ==============================
```

Table D- 5. SEM Model 1 output on theft from vehicle theft

```
============================== BEGINNING OF REPORT============================
Data set            : TPAll07DataZCNov10
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :       TFA07  Number of Observations: 2602
Mean dependent var  :    0.936587  Number of Variables   :    3
S.D. dependent var  :     1.56373  Degrees of Freedom    : 2599
Lag coeff.   (Rho)  :    0.646786

R-squared           :    0.077856  Log likelihood        :       -4769
Sq. Correlation     : -            Akaike info criterion :        9544
Sigma-square        :     2.25487  Schwarz criterion     :      9561.6
S.E of regression   :     1.50162
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    z-value      Probability
-----------------------------------------------------------------------
     W_TFA07     0.6467861     0.04515349      14.32417    0.0000000
    CONSTANT     0.2947527     0.05136508      5.738387    0.0000000
      BSTOPS    0.03816011     0.01140703      3.345317    0.0008220
-----------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF     VALUE          PROB
Breusch-Pagan test                       1       8.417929     0.0037154

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF     VALUE          PROB
Likelihood Ratio Test                    1       155.2372     0.0000000
========================= END OF REPORT ======================================
```

Table D- 6. SLM Model 2 output on theft from vehicle theft

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :       TFA07      Number of Observations: 2602
Mean dependent var  :    0.936587      Number of Variables   :    9
S.D. dependent var  :     1.56373      Degrees of Freedom    : 2593
Lag coeff.   (Rho)  :    0.643612

R-squared           :    0.089447      Log Likelihood        :    -4752.31
Sq. Correlation     : -                Akaike info criterion :     9522.61
Sigma-square        :     2.22653      Schwarz criterion     :     9575.39
S.E of regression   :     1.49216
------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
------------------------------------------------------------------------------
Spatial lag              0.6436116      0.04473303     14.38784     0.0000000
Constant                 0.2256322      0.05211863      4.329203    0.0000150
Eating/drinking places   0.1015427      0.05480271      1.852877    0.0638999
Automotive retail        0.1559777      0.09318559      1.673839    0.0941622
Food store retail        0.2345806      0.08012232      2.92778     0.0034140
Automotive service       0.1109223      0.06303977      1.759561    0.0784822
Business service         0.02137063     0.05264814      0.4059142   0.6848057
Personal service         0.09856038     0.05081504      1.939591    0.0524293
Banks                    0.09136636     0.1447698       0.6311148   0.5279653
------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
```

```
RANDOM COEFFICIENTS
TEST                                    DF     VALUE         PROB
Breusch-Pagan test                      7       253.5523     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                    DF     VALUE         PROB
Likelihood Ratio Test                   1       154.4694     0.0000000
========================= END OF REPORT ======================================
```

Table D- 7. SLM Model 3 output on theft from vehicle theft

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :       TFA07      Number of Observations: 2602
Mean dependent var  :    0.936587      Number of Variables   :   10
S.D. dependent var  :     1.56373      Degrees of Freedom    : 2592
Lag coeff.   (Rho)  :    0.638301

R-squared           :    0.090598      Log Likelihood        :    -4750.27
Sq. Correlation     : -                Akaike info criterion :     9520.54
Sigma-square        :     2.22371      Schwarz criterion     :     9579.18
S.E of regression   :     1.49121
------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
------------------------------------------------------------------------------
Spatial lag              0.6383007      0.04504148     14.1714       0.0000000
Constant                 0.2194944      0.05250034      4.180818     0.0000291
Bus stops                0.02348984     0.01161506      2.022361     0.0431389
Eating/drinking places   0.09197957     0.05498658      1.672764     0.0943737
Automotive retail        0.1492912      0.09317819      1.602212     0.1091088
Food store retail        0.2283248      0.08013557      2.849231     0.0043826
Automotive service       0.1038714      0.06308881      1.646432     0.0996748
Business service         0.01762914     0.0526554       0.3348022    0.7377745
Personal service         0.09025912     0.05095096      1.77149      0.0764791
Banks                    0.08009917     0.1447791       0.5532509    0.5800916
------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                    DF     VALUE         PROB
Breusch-Pagan test                      8       256.0991     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                    DF     VALUE         PROB
Likelihood Ratio Test                   1       150.9345     0.0000000
========================= END OF REPORT ======================================
```

Table D- 8. SLM Model 4 output on theft from vehicle theft

```
============================= BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :       TFA07      Number of Observations: 2602
Mean dependent var  :    0.936587      Number of Variables   :   16
S.D. dependent var  :     1.56373      Degrees of Freedom    : 2586
Lag coeff.   (Rho)  :    0.638142

R-squared           :    0.098123      Log Likelihood        :    -4739.45
Sq. Correlation     : -                Akaike info criterion :     9510.89
Sigma-square        :     2.20531      Schwarz criterion     :     9604.72
S.E of regression   :     1.48503
-----------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Spatial lag              0.6381418      0.04452337     14.33274      0.0000000
Constant                 0.09262946     0.06781953      1.365823     0.1719947
Bus stops                0.02420222     0.0117144       2.066024     0.0388261
Eating/drinking places   0.08066994     0.05486027      1.470462     0.1414368
Automotive retail        0.1400792      0.09300508      1.506145     0.1320300
Food store retail        0.2114365      0.08007164      2.640592     0.0082762
Automotive service       0.09127151     0.06295085      1.449885     0.1470906
Business service         0.02379384     0.0525808       0.4525195    0.6508948
Personal service         0.09104713     0.05091437      1.78824      0.0737371
Banks                    0.0714728      0.1444326       0.4948521    0.6207045
Mixed landuse            0.293629       0.09339247      3.144033     0.0016665
Vacant land              0.1545088      0.06122889      2.523462     0.0116206
Public housing          -0.083367       0.0959049      -0.8692673    0.3847009
Colleges                -0.1332621      0.2053208      -0.6490436    0.5163101
Grade K-12               0.1519745      0.09760485      1.557038     0.1194615
Parks and cemeteries    -0.01483834     0.08222786     -0.180454     0.8567962
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF      VALUE         PROB
Breusch-Pagan test                       14       309.9603     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF      VALUE         PROB
Likelihood Ratio Test                    1        148.6997     0.0000000
========================= END OF REPORT ======================================
```

Table D- 9. PRM Model 1 output on theft from vehicle theft

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(2)      =     479.07
                                                  Prob > chi2     =     0.0000
Log likelihood = -3697.1399                       Pseudo R2       =     0.0608
------------------------------------------------------------------------------
       tfa07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wtfa |   .7624133   .0326103    23.38   0.000     .6984984    .8263283
      bstops |   .0230555   .0054522     4.23   0.000     .0123695    .0337416
       _cons |  -.8858541   .0431204   -20.54   0.000    -.9703685   -.8013397
------------------------------------------------------------------------------
```

Table D- 10. PRM Model 2 output on theft from vehicle theft

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(8)      =     534.02
                                                  Prob > chi2     =     0.0000
Log likelihood = -3669.6635                       Pseudo R2       =     0.0678
------------------------------------------------------------------------------
       tfa07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wtfa |   .7419644   .0327715    22.64   0.000     .6777333    .8061954
     sic58xx |    .071116   .0311058     2.29   0.022     .0101499    .1320822
     sic55xx |   .1449152   .0530332     2.73   0.006      .040972    .2488584
     sic54xx |   .1239278   .0438438     2.83   0.005     .0379955      .20986
     sic75xx |   .0987675   .0377937     2.61   0.009     .0246932    .1728417
     sic73xx |   .0212963   .0284841     0.75   0.455    -.0345314    .0771241
     sic72xx |   .0635329   .0292246     2.17   0.030     .0062537     .120812
     finance |   .1199683   .0795361     1.51   0.131    -.0359195    .2758561
       _cons |  -.9292118   .0434917   -21.37   0.000    -1.014454   -.8439695
------------------------------------------------------------------------------
```

Table D- 11. PRM Model 3 output on theft from vehicle theft

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(9)      =     538.92
                                                  Prob > chi2     =     0.0000
Log likelihood = -3667.2148                       Pseudo R2       =     0.0684
------------------------------------------------------------------------------
       tfa07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wtfa |   .7355121   .0329385    22.33   0.000     .6709538    .8000703
      bstops |    .014193   .0060913     2.33   0.020     .0022542    .0261318
     sic58xx |    .064558   .0314323     2.05   0.040     .0029519    .1261642
     sic55xx |   .1387072   .0534934     2.59   0.010      .033862    .2435523
     sic54xx |   .1214951   .0439166     2.77   0.006     .0354201      .20757
     sic75xx |    .094107    .037924     2.48   0.013     .0197774    .1684366
     sic73xx |   .0207929   .0287379     0.72   0.469    -.0355323    .0771182
     sic72xx |   .0584083    .029335     1.99   0.046     .0009128    .1159038
     finance |   .1148292   .0800021     1.44   0.151    -.0419722    .2716305
       _cons |  -.9306929    .043513   -21.39   0.000    -1.015977   -.8454091
------------------------------------------------------------------------------
```

Table D- 12. PRM Model 4 output on theft from vehicle theft

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(15)     =     592.98
                                                  Prob > chi2     =     0.0000
Log likelihood =  -3640.182                       Pseudo R2       =     0.0753
------------------------------------------------------------------------------
       tfa07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wtfa |   .7455865   .0336015    22.19   0.000     .6797288    .8114442
      bstops |   .0146238   .0061162     2.39   0.017     .0026362    .0266113
     sic58xx |   .0489796   .0316417     1.55   0.122    -.0130369    .1109961
     sic55xx |   .1182347   .0549318     2.15   0.031     .0105703    .2258991
     sic54xx |   .1008131   .0449395     2.24   0.025     .0127333    .1888929
     sic75xx |   .0817146   .0382393     2.14   0.033     .0067669    .1566623
     sic73xx |   .0316869   .0288835     1.10   0.273    -.0249237    .0882974
     sic72xx |   .0645499   .0297001     2.17   0.030     .0063388    .1227609
     finance |   .1127089   .0821305     1.37   0.170    -.0482639    .2736817
    mixeduse |   .3138419   .0571078     5.50   0.000     .2019126    .4257712
   pc1fvland |   .1768948   .0439588     4.02   0.000     .0907372    .2630524
    phousing |  -.0847053   .0738231    -1.15   0.251    -.2293958    .0599853
    univcoll |  -.1098804   .1310139    -0.84   0.402    -.3666628    .1469021
     schools |   .0890149   .0631674     1.41   0.159     -.034791    .2128208
   openspace |  -.0273631   .0581768    -0.47   0.638    -.1413876    .0866615
       _cons |  -1.087234   .0557452   -19.50   0.000    -1.196493   -.9779755
------------------------------------------------------------------------------
```

Table D- 13. NBRM Model 1 output on theft from vehicle theft

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(2)      =     225.79
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -3371.2904                       Pseudo R2       =     0.0324
------------------------------------------------------------------------------
       tfa07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wtfa |   .8223858    .058541    14.05   0.000     .7076475     .937124
      bstops |   .0357182   .0111687     3.20   0.001     .0138279    .0576085
       _cons |  -.9607516   .0672033   -14.30   0.000    -1.092468   -.8290354
-------------+----------------------------------------------------------------
    /lnalpha |  -.0382454   .0713243                     -.1780385    .1015477
-------------+----------------------------------------------------------------
       alpha |   .9624767    .068648                      .8369102    1.106883
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  651.70 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -3484.186    -3371.29      4     6750.581    6774.037
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table D- 14. NBRM Model 2 output on theft from vehicle theft

Negative binomial regression		Number of obs =	2602
		LR chi2(8) =	252.62
Dispersion = mean		Prob > chi2 =	0.0000
Log likelihood = -3357.8743		Pseudo R2 =	0.0363

tfa07	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wtfa	.8191865	.0586901	13.96	0.000	.704156	.934217
sic58xx	.0762345	.0479446	1.59	0.112	-.0177351	.1702042
sic55xx	.1930399	.0856006	2.26	0.024	.0252659	.360814
sic54xx	.148443	.0721504	2.06	0.040	.0070308	.2898553
sic75xx	.1256812	.0576615	2.18	0.029	.0126667	.2386957
sic73xx	.0146134	.0473794	0.31	0.758	-.0782485	.1074754
sic72xx	.0701719	.0450363	1.56	0.119	-.0180976	.1584413
finance	.1323618	.1302301	1.02	0.309	-.1228844	.387608
_cons	-1.025219	.0686596	-14.93	0.000	-1.159789	-.8906487
/lnalpha	-.0732176	.0725384			-.2153903	.0689551
alpha	.9293986	.0674171			.8062267	1.071388

Likelihood-ratio test of alpha=0: chibar2(01) = 623.58 Prob>=chibar2 = 0.000

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	2602	-3484.186	-3357.874	10	6735.749	6794.389

Note: N=Obs used in calculating BIC; see [R] BIC note

Table D- 15. NBRM Model 3 output on theft from vehicle theft

Negative binomial regression		Number of obs =	2602
		LR chi2(9) =	257.23
Dispersion = mean		Prob > chi2 =	0.0000
Log likelihood = -3355.5717		Pseudo R2 =	0.0369

tfa07	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wtfa	.8108566	.0587312	13.81	0.000	.6957455	.9259677
bstops	.0230707	.0109738	2.10	0.036	.0015625	.0445788
sic58xx	.0667981	.0480407	1.39	0.164	-.0273598	.1609561
sic55xx	.1788186	.0855742	2.09	0.037	.0110963	.3465409
sic54xx	.1448005	.0720604	2.01	0.044	.0035648	.2860363
sic75xx	.1234167	.0576801	2.14	0.032	.0103657	.2364676
sic73xx	.0135935	.0470911	0.29	0.773	-.0787033	.1058903
sic72xx	.0633605	.045128	1.40	0.160	-.0250888	.1518098
finance	.1113249	.1301677	0.86	0.392	-.1437992	.3664489
_cons	-1.03064	.0686707	-15.01	0.000	-1.165232	-.8960479
/lnalpha	-.0761591	.072549			-.2183525	.0660344
alpha	.9266688	.0672289			.803842	1.068263

Likelihood-ratio test of alpha=0: chibar2(01) = 623.29 Prob>=chibar2 = 0.000

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	2602	-3484.186	-3355.572	11	6733.143	6797.648

Note: N=Obs used in calculating BIC; see [R] BIC note

Table D- 16. NBRM Model 4 output on theft from vehicle theft

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(15)     =     287.04
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -3340.6666                       Pseudo R2       =     0.0412
------------------------------------------------------------------------------
       tfa07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wtfa |   .8104284   .0586767    13.81   0.000     .6954242    .9254326
      bstops |   .0234795   .0110542     2.12   0.034     .0018137    .0451453
     sic58xx |   .0490981   .0481266     1.02   0.308    -.0452282    .1434244
     sic55xx |   .1716465   .0849244     2.02   0.043     .0051978    .3380953
     sic54xx |   .1362503   .0716401     1.90   0.057    -.0041617    .2766623
     sic75xx |   .1143395   .0574678     1.99   0.047     .0017047    .2269743
     sic73xx |   .0219549   .0465139     0.47   0.637    -.0692106    .1131204
     sic72xx |   .0721484   .0450685     1.60   0.109    -.0161842     .160481
     finance |   .1029109   .1284985     0.80   0.423    -.1489415    .3547634
    mixeduse |   .3092541   .0838195     3.69   0.000      .144971    .4735372
   pc1fvland |    .181904   .0605982     3.00   0.003     .0631337    .3006742
    phousing |  -.1312385   .0981573    -1.34   0.181    -.3236234    .0611464
    univcoll |  -.1280428   .1869605    -0.68   0.493    -.4944786    .2383931
     schools |   .1415191   .0910861     1.55   0.120    -.0370064    .3200447
   openspace |  -.0105669   .0806596    -0.13   0.896    -.1686567     .147523
       _cons |  -1.185533   .0829913   -14.29   0.000    -1.348193   -1.022873
-------------+----------------------------------------------------------------
    /lnalpha |   -.111875    .073689                     -.2563028    .0325528
-------------+----------------------------------------------------------------
       alpha |    .894156   .0658895                      .7739076    1.033088
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  599.03 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -3484.186   -3340.667     17     6715.333    6815.022
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table D- 17. Model fit comparison of PRM & NBRM Model 1 on theft from vehicle theft

```
----------------------------------------------------------
                        Variable |    PRM        NBRM
---------------------------------+------------------------
tfa07                            |
                            WTFA |    2.143       2.276
                                 |    23.38       14.05
                          BSTOPS |    1.023       1.036
                                 |     4.23        3.20
                        Constant |    0.412       0.383
                                 |   -20.54      -14.30
---------------------------------+------------------------
lnalpha                          |
                        Constant |                0.962
                                 |                -0.54
---------------------------------+------------------------
Statistics                       |
                           alpha |                0.962
                               N |  2602.000    2602.000
                              ll | -3697.140   -3371.290
                             bic |  7417.872    6774.037
                             aic |  7400.280    6750.581
----------------------------------------------------------
```

```
                                                 legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM        -0.106         1      0.028
NBRM       -0.009         0      0.003

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.422      0.104    67.261
1        0.238       0.343      0.106    84.413
2        0.105       0.155      0.050    42.070
3        0.053       0.053      0.000     0.000
4        0.021       0.017      0.005     3.207
5        0.011       0.005      0.005    14.085
6        0.006       0.002      0.004    21.782
7        0.005       0.001      0.004    46.139
8        0.002       0.000      0.002    17.280
9        0.003       0.000      0.002    53.524
------------------------------------------------
Sum      0.971       1.000      0.282   349.759

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.535      0.009     0.381
1        0.238       0.242      0.004     0.194
2        0.105       0.111      0.006     0.983
3        0.053       0.053      0.000     0.012
4        0.021       0.026      0.005     2.474
5        0.011       0.013      0.003     1.382
6        0.006       0.007      0.001     0.409
7        0.005       0.004      0.001     0.573
8        0.002       0.002      0.000     0.006
9        0.003       0.001      0.001     2.685
------------------------------------------------
Sum      0.971       0.997      0.031     9.099

Tests and Fit Statistics

PRM            BIC=-13044.349  AIC=     2.844  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM      BIC=-13688.184  dif=   643.835  NBRM    PRM   Very strong
               AIC=     2.594  dif=     0.250  NBRM    PRM
               LRX2=  651.699  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM           BIC=-13688.184  AIC=     2.594  Prefer  Over  Evidence
```

Table D- 18. Model fit comparison of PRM & NBRM Model 2 on theft from vehicle theft

```
----------------------------------------------------------
                        Variable |    PRM        NBRM
---------------------------------+------------------------
tfa07                            |
                            WTFA |     2.094       2.266
```

```
                                 |     22.61       13.93
                         SIC58XX |     1.081       1.087
                                 |      2.55        1.74
                         SIC55XX |     1.151       1.208
                                 |      2.65        2.21
                         SIC54XX |     1.138       1.167
                                 |      2.99        2.15
                         SIC75XX |     1.107       1.138
                                 |      2.70        2.24
                         SIC73XX |     1.023       1.019
                                 |      0.79        0.40
                         SIC72XX |     1.068       1.077
                                 |      2.25        1.64
                        Constant |     0.397       0.360
                                 |    -21.34      -14.89
---------------------------------+-----------------------
lnalpha                          |
                        Constant |                 0.931
                                 |                 -0.99
---------------------------------+-----------------------
Statistics                       |
                           alpha |                 0.931
                               N |  2602.000    2602.000
                              ll | -3670.729   -3358.397
                             bic |  7404.370    6787.571
                             aic |  7357.457    6734.795
---------------------------------------------------------
                                              legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM         -0.104        1      0.028
NBRM        -0.008        0      0.003

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.424      0.102    64.004
1        0.238       0.342      0.104    82.687
2        0.105       0.154      0.049    40.796
3        0.053       0.053      0.000     0.004
4        0.021       0.017      0.005     3.367
5        0.011       0.005      0.005    14.702
6        0.006       0.002      0.004    22.758
7        0.005       0.001      0.004    46.942
8        0.002       0.001      0.002    16.596
9        0.003       0.000      0.002    47.449
------------------------------------------------
Sum      0.971       0.999      0.278   339.305

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.535      0.008     0.337
1        0.238       0.243      0.005     0.321
2        0.105       0.111      0.007     0.998
3        0.053       0.053      0.001     0.027
4        0.021       0.026      0.005     2.276
5        0.011       0.013      0.003     1.245
```

6	0.006	0.007	0.001	0.345
7	0.005	0.004	0.001	0.656
8	0.002	0.002	0.000	0.001
9	0.003	0.001	0.001	2.845
Sum	0.971	0.996	0.032	9.051

Tests and Fit Statistics

PRM	BIC=-13057.851	AIC= 2.828	Prefer	Over	Evidence
vs NBRM	BIC=-13674.650	dif= 616.799	NBRM	PRM	Very strong
	AIC= 2.588	dif= 0.239	NBRM	PRM	
	LRX2= 624.663	prob= 0.000	NBRM	PRM	p=0.000
NBRM	BIC=-13674.650	AIC= 2.588	Prefer	Over	Evidence

Table D- 19. Model fit comparison of PRM & NBRM Model 3 on theft from vehicle theft

Variable	PRM	NBRM
tfa07		
WTFA	2.081	2.247
	22.29	13.78
BSTOPS	1.015	1.024
	2.38	2.17
SIC58XX	1.074	1.075
	2.30	1.51
SIC55XX	1.144	1.191
	2.51	2.04
SIC54XX	1.136	1.162
	2.92	2.08
SIC75XX	1.101	1.134
	2.56	2.19
SIC73XX	1.022	1.017
	0.76	0.36
SIC72XX	1.062	1.068
	2.06	1.47
Constant	0.396	0.358
	-21.37	-14.98
lnalpha		
Constant		0.928
		-1.03
Statistics		
alpha		0.928
N	2602.000	2602.000
ll	-3668.183	-3355.941
bic	7407.142	6790.522
aic	7354.366	6731.882

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
PRM	-0.104	1	0.028
NBRM	-0.009	0	0.003

PRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.527	0.425	0.102	63.560
1	0.238	0.342	0.104	82.293
2	0.105	0.154	0.049	40.527
3	0.053	0.053	0.000	0.003
4	0.021	0.017	0.005	3.215
5	0.011	0.005	0.005	14.267
6	0.006	0.002	0.004	22.387
7	0.005	0.001	0.004	47.080
8	0.002	0.000	0.002	17.052
9	0.003	0.000	0.002	49.643
Sum	0.971	0.999	0.277	340.027

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.527	0.535	0.009	0.356
1	0.238	0.243	0.005	0.309
2	0.105	0.111	0.006	0.940
3	0.053	0.053	0.001	0.037
4	0.021	0.026	0.005	2.237
5	0.011	0.013	0.003	1.251
6	0.006	0.007	0.001	0.365
7	0.005	0.004	0.001	0.604
8	0.002	0.002	0.000	0.006
9	0.003	0.001	0.001	2.655
Sum	0.971	0.996	0.032	8.758

Tests and Fit Statistics

PRM	BIC=-13055.079	AIC= 2.826	Prefer	Over	Evidence
vs NBRM	BIC=-13671.698	dif= 616.620	NBRM	PRM	Very strong
	AIC= 2.587	dif= 0.239	NBRM	PRM	
	LRX2= 624.484	prob= 0.000	NBRM	PRM	p=0.000
NBRM	BIC=-13671.698	AIC= 2.587	Prefer	Over	Evidence

Table D- 20. Model fit comparison of PRM & NBRM Model 4 on theft from vehicle theft

Variable	PRM	NBRM
tfa07		
WTFA	2.101	2.245
	22.15	13.79
BSTOPS	1.015	1.024
	2.43	2.18
SIC58XX	1.057	1.055
	1.77	1.13
SIC55XX	1.120	1.183
	2.07	1.98
SIC54XX	1.112	1.152
	2.39	1.97

```
                         SIC75XX |     1.088       1.124
                                 |      2.20        2.04
                         SIC73XX |     1.033       1.025
                                 |      1.12        0.54
                         SIC72XX |     1.069       1.077
                                 |      2.27        1.66
                        MIXEDUSE |     1.374       1.366
                                 |      5.57        3.72
                       PC1FVLAND |     1.190       1.197
                                 |      3.97        2.97
                        PHOUSING |     0.917       0.876
                                 |     -1.18       -1.35
                        UNIVCOLL |     0.897       0.882
                                 |     -0.83       -0.67
                         SCHOOLS |     1.092       1.154
                                 |      1.39        1.57
                       OPENSPACE |     0.973       0.989
                                 |     -0.47       -0.13
                        Constant |     0.339       0.307
                                 |    -19.48      -14.26
---------------------------------+-----------------------
lnalpha                          |
                        Constant |                 0.895
                                 |                 -1.50
---------------------------------+-----------------------
Statistics                       |
                           alpha |                 0.895
                               N |  2602.000    2602.000
                              ll | -3641.071   -3340.989
                             bic |  7400.103    6807.803
                             aic |  7312.143    6713.979
---------------------------------------------------------
                                              legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
------------------------------------------------
PRM         -0.100        1      0.027
NBRM        -0.008        0      0.003

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.428      0.098    58.770
1        0.238       0.338      0.100    77.569
2        0.105       0.152      0.048    38.546
3        0.053       0.054      0.000     0.002
4        0.021       0.017      0.004     2.389
5        0.011       0.006      0.005    12.003
6        0.006       0.002      0.004    19.568
7        0.005       0.001      0.004    42.902
8        0.002       0.001      0.002    15.922
9        0.003       0.000      0.002    49.365
------------------------------------------------
Sum      0.971       0.999      0.268   317.037

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.535      0.008     0.343
```

```
1        0.238       0.244      0.006     0.388
2        0.105       0.111      0.006     0.882
3        0.053       0.052      0.001     0.061
4        0.021       0.026      0.005     2.132
5        0.011       0.013      0.002     1.219
6        0.006       0.007      0.001     0.368
7        0.005       0.004      0.001     0.580
8        0.002       0.002      0.000     0.009
9        0.003       0.001      0.001     2.568
------------------------------------------------
Sum      0.971       0.996      0.032     8.550

Tests and Fit Statistics

PRM            BIC=-13062.117  AIC=     2.810  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM      BIC=-13654.418  dif=   592.300  NBRM    PRM   Very strong
               AIC=     2.580  dif=     0.230  NBRM    PRM
               LRX2=  600.164  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM           BIC=-13654.418  AIC=     2.580  Prefer  Over  Evidence
```

Table D- 21. Model fit comparison of NBRM & ZINBRM Model 1 on theft from vehicle theft

```
----------------------------------------------------------
                        Variable |   NBRM       ZINBRM
---------------------------------+------------------------
tfa07                            |
                            WTFA |    2.276       2.144
                                 |    14.05       11.87
                          BSTOPS |    1.036       1.034
                                 |     3.20        2.99
                        Constant |    0.383       0.418
                                 |   -14.30      -10.82
---------------------------------+------------------------
lnalpha                          |
                        Constant |    0.962       0.928
                                 |    -0.54       -0.98
---------------------------------+------------------------
inflate                          |
                            WTFA |                0.002
                                 |                -2.34
                          BSTOPS |                0.000
                                 |                -0.02
                        Constant |                2.006
                                 |                 0.75
---------------------------------+------------------------
Statistics                       |
                           alpha |    0.962
                               N |  2602.000    2602.000
                              ll | -3371.290   -3368.140
                             bic |  6774.037    6791.328
                             aic |  6750.581    6750.280
----------------------------------------------------------
                                                legend: b/t
Comparison of Mean Observed and Predicted Count

           Maximum        At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.009         0      0.003
```

ZINBRM	-0.010	0	0.003

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.527	0.535	0.009	0.381
1	0.238	0.242	0.004	0.194
2	0.105	0.111	0.006	0.983
3	0.053	0.053	0.000	0.012
4	0.021	0.026	0.005	2.474
5	0.011	0.013	0.003	1.382
6	0.006	0.007	0.001	0.409
7	0.005	0.004	0.001	0.573
8	0.002	0.002	0.000	0.006
9	0.003	0.001	0.001	2.685
Sum	0.971	0.997	0.031	9.099

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.527	0.536	0.010	0.459
1	0.238	0.239	0.001	0.018
2	0.105	0.112	0.008	1.302
3	0.053	0.054	0.001	0.013
4	0.021	0.027	0.006	2.964
5	0.011	0.014	0.003	1.567
6	0.006	0.007	0.001	0.424
7	0.005	0.004	0.001	0.645
8	0.002	0.002	0.000	0.000
9	0.003	0.001	0.001	3.188
Sum	0.971	0.997	0.031	10.580

Tests and Fit Statistics

```
-------------------------------------------------------------------------
NBRM           BIC=-13688.184  AIC=     2.594  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM    BIC=-13670.893  dif=   -17.291  NBRM    ZINBRM  Very strong
               AIC=     2.594  dif=     0.000  ZINBRM    NBRM
               Vuong=   1.114  prob=    0.133  ZINBRM    NBRM  p=0.133
```

Table D- 22. Model fit comparison of NBRM & ZINBRM Model 2 on theft from vehicle theft

Variable	NBRM	ZINBRM
tfa07		
WTFA	2.269	2.051
	13.96	10.93
SIC58XX	1.079	1.062
	1.59	1.28
SIC55XX	1.213	1.175
	2.26	1.90
SIC54XX	1.160	1.139
	2.06	1.85
SIC75XX	1.134	1.116
	2.18	1.90

```
                         SIC73XX |     1.015       1.040
                                 |      0.31        0.77
                         SIC72XX |     1.073       1.059
                                 |      1.56        1.28
                         FINANCE |     1.142       1.153
                                 |      1.02        1.08
                        Constant |     0.359       0.422
                                 |    -14.93       -9.41
---------------------------------+------------------------
lnalpha                          |
                        Constant |     0.929       0.864
                                 |     -1.01       -1.74
---------------------------------+------------------------
inflate                          |
                            WTFA |                 0.007
                                 |                 -2.21
                         SIC58XX |                 0.002
                                 |                 -0.29
                         SIC55XX |                 0.017
                                 |                 -0.15
                         SIC54XX |                 0.000
                                 |                 -0.02
                         SIC75XX |                 0.744
                                 |                 -0.44
                         SIC73XX |                 1.729
                                 |                  1.66
                         SIC72XX |                 0.000
                                 |                 -0.02
                         FINANCE |                 3.769
                                 |                  0.80
                        Constant |                 2.412
                                 |                  1.03
---------------------------------+------------------------
Statistics                       |
                           alpha |     0.929
                               N |  2602.000    2602.000
                              ll | -3357.874   -3349.603
                             bic |  6794.389    6848.623
                             aic |  6735.749    6737.206
----------------------------------------------------------
                                              legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.008         0      0.003
ZINBRM       -0.010         0      0.003

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.535      0.008     0.336
1        0.238       0.243      0.006     0.324
2        0.105       0.111      0.007     0.994
3        0.053       0.053      0.001     0.028
4        0.021       0.026      0.005     2.273
5        0.011       0.013      0.003     1.248
6        0.006       0.007      0.001     0.348
7        0.005       0.004      0.001     0.648
8        0.002       0.002      0.000     0.002
9        0.003       0.001      0.001     2.822
```

```
-----------------------------------------------
Sum      0.971       0.996      0.032     9.022

ZINB: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
-----------------------------------------------
0        0.527       0.537      0.010     0.503
1        0.238       0.238      0.000     0.001
2        0.105       0.113      0.008     1.602
3        0.053       0.055      0.001     0.056
4        0.021       0.027      0.006     3.163
5        0.011       0.014      0.003     1.578
6        0.006       0.007      0.001     0.379
7        0.005       0.004      0.001     0.773
8        0.002       0.002      0.000     0.007
9        0.003       0.001      0.001     3.704
-----------------------------------------------
Sum      0.971       0.997      0.032    11.767

Tests and Fit Statistics

-------------------------------------------------------------------------
NBRM           BIC=-13667.832  AIC=     2.589  Prefer  Over  Evidence
-------------------------------------------------------------------------
vs ZINBRM      BIC=-13613.598  dif=   -54.234  NBRM    ZINBRM  Very strong
               AIC=     2.589  dif=    -0.001  NBRM    ZINB
               Vuong=   1.989  prob=    0.023  ZINBRM    NBRM  p=0.023
```

Table D- 23. Model fit comparison of NBRM & ZINBRM Model 3 on theft from vehicle theft

```
----------------------------------------------------------
                        Variable |   NBRM       ZINBRM
---------------------------------+------------------------
tfa07                            |
                            WTFA |    2.247       2.036
                                 |    13.78       10.77
                          BSTOPS |    1.024       1.023
                                 |     2.17        2.04
                         SIC58XX |    1.075       1.060
                                 |     1.51        1.24
                         SIC55XX |    1.191       1.156
                                 |     2.04        1.72
                         SIC54XX |    1.162       1.143
                                 |     2.08        1.90
                         SIC75XX |    1.134       1.117
                                 |     2.19        1.93
                         SIC73XX |    1.017       1.041
                                 |     0.36        0.81
                         SIC72XX |    1.068       1.056
                                 |     1.47        1.23
                        Constant |    0.358       0.419
                                 |   -14.98       -9.48
---------------------------------+------------------------
lnalpha                          |
                        Constant |    0.928       0.866
                                 |    -1.03       -1.73
---------------------------------+------------------------
inflate                          |
                            WTFA |                0.005
                                 |                -2.29
```

```
                  BSTOPS |                  0.938
                         |                  -0.20
                 SIC58XX |                  0.025
                         |                  -0.41
                 SIC55XX |                  0.001
                         |                  -0.07
                 SIC54XX |                  0.000
                         |                  -0.02
                 SIC75XX |                  0.778
                         |                  -0.46
                 SIC73XX |                  1.742
                         |                   1.64
                 SIC72XX |                  0.000
                         |                  -0.02
                Constant |                  2.829
                         |                   1.23
-------------------------+----------------------
Statistics               |
                   alpha |     0.928
                       N |  2602.000    2602.000
                      ll | -3355.941   -3347.944
                     bic |  6790.522    6845.305
                     aic |  6731.882    6733.888
------------------------------------------------
                                      legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.009         0      0.003
ZINBRM       -0.010         0      0.003

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.535      0.009     0.356
1        0.238       0.243      0.005     0.309
2        0.105       0.111      0.006     0.940
3        0.053       0.053      0.001     0.037
4        0.021       0.026      0.005     2.237
5        0.011       0.013      0.003     1.251
6        0.006       0.007      0.001     0.365
7        0.005       0.004      0.001     0.604
8        0.002       0.002      0.000     0.006
9        0.003       0.001      0.001     2.655
------------------------------------------------
Sum      0.971       0.996      0.032     8.758

ZINB: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.537      0.010     0.514
1        0.238       0.238      0.000     0.000
2        0.105       0.113      0.008     1.522
3        0.053       0.054      0.001     0.040
4        0.021       0.027      0.006     3.091
5        0.011       0.014      0.003     1.567
6        0.006       0.007      0.001     0.392
7        0.005       0.004      0.001     0.728
8        0.002       0.002      0.000     0.003
```

```
9        0.003       0.001      0.001     3.497
------------------------------------------------
Sum      0.971       0.997      0.031    11.354

Tests and Fit Statistics
-------------------------------------------------------------------------
NBRM           BIC=-13671.698  AIC=     2.587  Prefer  Over  Evidence
-------------------------------------------------------------------------
vs ZINBRM      BIC=-13616.916  dif=   -54.783  NBRM    ZINBRM  Very strong
               AIC=     2.588  dif=    -0.001  NBRM    ZINB
               Vuong=   1.910  prob=    0.028  ZINBRM    NBRM  p=0.028
```

Table D- 24. Model fit comparison of NBRM & ZINBRM Model 4 on theft from vehicle theft

	Variable	NBRM	ZINBRM
tfa07			
	WTFA	2.245	2.005
		13.79	10.98
	BSTOPS	1.024	1.032
		2.18	2.27
	SIC58XX	1.055	1.034
		1.13	0.72
	SIC55XX	1.183	1.140
		1.98	1.52
	SIC54XX	1.152	1.097
		1.97	1.36
	SIC75XX	1.124	1.075
		2.04	1.13
	SIC73XX	1.025	1.064
		0.54	1.25
	SIC72XX	1.077	1.045
		1.66	1.01
	MIXEDUSE	1.366	1.251
		3.72	2.63
	PC1FVLAND	1.197	1.189
		2.97	2.33
	PHOUSING	0.876	1.258
		-1.35	1.78
	UNIVCOLL	0.882	0.968
		-0.67	-0.15
	SCHOOLS	1.154	1.159
		1.57	1.43
	OPENSPACE	0.989	0.940
		-0.13	-0.71
	Constant	0.307	0.380
		-14.26	-9.60
lnalpha			
	Constant	0.895	0.757
		-1.50	-3.08
inflate			
	WTFA		0.089
			-2.42
	BSTOPS		1.120
			2.04
	SIC58XX		0.236
			-0.85
	SIC55XX		0.275

```
                                 |                 -0.67
                         SIC54XX |                 0.000
                                 |                 -0.02
                         SIC75XX |                 0.491
                                 |                 -0.58
                         SIC73XX |                 1.519
                                 |                  0.71
                         SIC72XX |                 0.000
                                 |                 -0.00
                        MIXEDUSE |                 0.096
                                 |                 -1.92
                       PC1FVLAND |                 1.019
                                 |                  0.03
                        PHOUSING |                12.977
                                 |                  4.81
                        UNIVCOLL |                 4.806
                                 |                  1.11
                         SCHOOLS |                 0.893
                                 |                 -0.13
                       OPENSPACE |                 0.453
                                 |                 -1.00
                        Constant |                 0.715
                                 |                 -0.50
---------------------------------+------------------------
Statistics                       |
                           alpha |     0.895
                               N |  2602.000    2602.000
                              ll | -3340.989   -3319.323
                             bic |  6807.803    6882.431
                             aic |  6713.979    6700.646
----------------------------------------------------------
                                                legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------------
NBRM       -0.008         0      0.003
ZINBRM       -0.012         0      0.004

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.535      0.008     0.343
1        0.238       0.244      0.006     0.388
2        0.105       0.111      0.006     0.882
3        0.053       0.052      0.001     0.061
4        0.021       0.026      0.005     2.132
5        0.011       0.013      0.002     1.219
6        0.006       0.007      0.001     0.368
7        0.005       0.004      0.001     0.580
8        0.002       0.002      0.000     0.009
9        0.003       0.001      0.001     2.568
------------------------------------------------
Sum      0.971       0.996      0.032     8.550

ZINB: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.539      0.012     0.702
1        0.238       0.234      0.004     0.150
```

2	0.105	0.114	0.009	1.829
3	0.053	0.055	0.002	0.125
4	0.021	0.027	0.006	3.422
5	0.011	0.014	0.003	1.683
6	0.006	0.007	0.001	0.407
7	0.005	0.004	0.001	0.751
8	0.002	0.002	0.000	0.007
9	0.003	0.001	0.001	3.711
Sum	0.971	0.997	0.039	12.787

Tests and Fit Statistics

NBRM	BIC=-13654.418	AIC= 2.580	Prefer	Over	Evidence
vs ZINBRM	BIC=-13579.790	dif= -74.628	NBRM	ZINBRM	Very strong
	AIC= 2.575	dif= 0.005	ZINBRM	NBRM	
	Vuong= 3.258	prob= 0.001	ZINBRM	NBRM	p=0.001

Table D- 25. Fit comparisons of Model 1 and Model 2 on theft from vehicle theft

	Model 2	Model 1	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-3484.186	-3484.186	0.000
Log-Lik Full Model	-3357.874	-3371.290	13.416
D	6715.749(2592)	6742.581(2598)	26.832(6)
LR	252.623(8)	225.791(2)	26.832(6)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.036	0.032	0.004
McFadden's Adj R2	0.033	0.031	0.002
ML (Cox-Snell) R2	0.093	0.083	0.009
Cragg-Uhler(Nagelkerke) R2	0.099	0.089	0.010
AIC	2.589	2.594	-0.006
AIC*n	6735.749	6750.581	-14.832
BIC	-13667.832	-13688.184	20.352
BIC'	-189.711	-210.063	20.352
BIC used by Stata	6794.389	6774.037	20.352
AIC used by Stata	6735.749	6750.581	-14.832

Difference of 20.352 in BIC' provides very strong support for Model 1.

Note: p-value for difference in LR is only valid if models are nested.

Table D- 26. Fit comparisons of Model 1 and Model 3 on theft from vehicle theft

	Model 3	Model 1	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-3484.186	-3484.186	0.000
Log-Lik Full Model	-3355.572	-3371.290	15.719
D	6711.143(2591)	6742.581(2598)	31.437(7)
LR	257.228(9)	225.791(2)	31.437(7)
Prob > LR	0.000	0.000	0.000

McFadden's R2	0.037	0.032	0.005
McFadden's Adj R2	0.034	0.031	0.003
ML (Cox-Snell) R2	0.094	0.083	0.011
Cragg-Uhler(Nagelkerke) R2	0.101	0.089	0.012
AIC	2.588	2.594	-0.007
AIC*n	6733.143	6750.581	-17.437
BIC	-13664.573	-13688.184	23.611
BIC'	-186.452	-210.063	23.611
BIC used by Stata	6797.648	6774.037	23.611
AIC used by Stata	6733.143	6750.581	-17.437

Difference of 23.611 in BIC' provides very strong support for Model 1.

Note: p-value for difference in LR is only valid if models are nested.

Table D- 27. Fit comparisons of Model 1 and Model 4 on theft from vehicle theft

	Model 4	Model 1	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-3484.186	-3484.186	0.000
Log-Lik Full Model	-3340.667	-3371.290	30.624
D	6681.333(2585)	6742.581(2598)	61.248(13)
LR	287.038(15)	225.791(2)	61.248(13)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.041	0.032	0.009
McFadden's Adj R2	0.036	0.031	0.005
ML (Cox-Snell) R2	0.104	0.083	0.021
Cragg-Uhler(Nagelkerke) R2	0.112	0.089	0.023
AIC	2.581	2.594	-0.014
AIC*n	6715.333	6750.581	-35.248
BIC	-13647.199	-13688.184	40.985
BIC'	-169.078	-210.063	40.985
BIC used by Stata	6815.022	6774.037	40.985
AIC used by Stata	6715.333	6750.581	-35.248

Difference of 40.985 in BIC' provides very strong support for Model 1.

Note: p-value for difference in LR is only valid if models are nested.

Table D- 28. Ad hoc model NBRM estimation on theft from vehicle theft

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(7)      =     274.61
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -3346.8804                       Pseudo R2       =     0.0394
------------------------------------------------------------------------------
       tfa07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wtfa |    .823117   .0581292    14.16   0.000     .7091858    .9370481
      bstops |   .0282512    .010875     2.60   0.009     .0069365    .0495658
     sic55xx |   .1782087   .0848169     2.10   0.036     .0119707    .3444467
     sic54xx |   .2071124   .0665436     3.11   0.002     .0766894    .3375354
     sic75xx |   .1160749   .0571151     2.03   0.042     .0041314    .2280184
    mixeduse |    .318401   .0835643     3.81   0.000      .154618    .4821841
    pc1fvland |   .1709057   .0602833     2.84   0.005     .0527526    .2890588
```

```
       _cons |  -1.171884   .0802024   -14.61   0.000    -1.329078    -1.01469
-------------+----------------------------------------------------------------
    /lnalpha |  -.0983598   .0732683                      -.241963    .0452434
-------------+----------------------------------------------------------------
       alpha |   .9063228   .0664047                      .7850853    1.046282
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  608.17 Prob>=chibar2 = 0.000

Measures of Fit for nbreg of tfa07

Log-Lik Intercept Only:      -3484.186   Log-Lik Full Model:           -3346.880
D(2593):                      6693.761   LR(7):                          274.611
                                         Prob > LR:                        0.000
McFadden's R2:                   0.039   McFadden's Adj R2:                0.037
ML (Cox-Snell) R2:               0.100   Cragg-Uhler(Nagelkerke) R2:       0.108
AIC:                             2.579   AIC*n:                         6711.761
BIC:                        -13697.684   BIC':                          -219.563
BIC used by Stata:            6764.537   AIC used by Stata:             6711.761
```

Table D- 29. Fit statistic comparisons between PRM and NBRM on ad hoc Model

```
----------------------------------------------------------
                        Variable |    PRM        NBRM
---------------------------------+------------------------
tfa07                            |
                            WTFA |    2.129       2.278
                                 |    22.89       14.16
                          BSTOPS |    1.018       1.029
                                 |     3.08        2.60
                         SIC55XX |    1.128       1.195
                                 |     2.19        2.10
                         SIC54XX |    1.202       1.230
                                 |     4.61        3.11
                         SIC75XX |    1.100       1.123
                                 |     2.51        2.03
                        MIXEDUSE |    1.388       1.375
                                 |     5.78        3.81
                       PC1FVLAND |    1.182       1.186
                                 |     3.84        2.84
                        Constant |    0.340       0.310
                                 |   -20.20      -14.61
---------------------------------+------------------------
lnalpha                          |
                        Constant |                0.906
                                 |                -1.34
---------------------------------+------------------------
Statistics                       |
                           alpha |                0.906
                               N |  2602.000    2602.000
                              ll | -3650.966   -3346.880
                             bic |  7364.844    6764.537
                             aic |  7317.932    6711.761
----------------------------------------------------------
                                                legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM         -0.101        1      0.027
```

```
NBRM       -0.008         0      0.003

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.427      0.099    60.066
1        0.238       0.339      0.101    78.423
2        0.105       0.153      0.048    39.126
3        0.053       0.054      0.000     0.003
4        0.021       0.017      0.004     2.422
5        0.011       0.006      0.005    12.086
6        0.006       0.002      0.004    19.619
7        0.005       0.001      0.004    43.335
8        0.002       0.001      0.002    16.614
9        0.003       0.000      0.002    53.318
------------------------------------------------
Sum      0.971       0.999      0.270   325.011

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.535      0.008     0.339
1        0.238       0.244      0.006     0.363
2        0.105       0.111      0.006     0.915
3        0.053       0.052      0.001     0.047
4        0.021       0.026      0.005     2.197
5        0.011       0.013      0.003     1.249
6        0.006       0.007      0.001     0.374
7        0.005       0.004      0.001     0.583
8        0.002       0.002      0.000     0.008
9        0.003       0.001      0.001     2.610
------------------------------------------------
Sum      0.971       0.996      0.032     8.684

Tests and Fit Statistics

PRM            BIC=-13097.377  AIC=     2.812  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM      BIC=-13697.684  dif=   600.307  NBRM    PRM   Very strong
               AIC=     2.579  dif=     0.233  NBRM    PRM
               LRX2=  608.171  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM           BIC=-13697.684  AIC=     2.579  Prefer  Over  Evidence
```

Table D- 30. Fit statistic comparisons between NBRM and ZINBRM on ad hoc Model

```
----------------------------------------------------------
                        Variable |   NBRM        ZINBRM
---------------------------------+------------------------
tfa07                            |
                            WTFA |     2.278       2.192
                                 |     14.16       12.96
                          BSTOPS |     1.029       1.029
                                 |      2.60        2.62
                         SIC55XX |     1.195       1.185
                                 |      2.10        2.01
                         SIC54XX |     1.230       1.221
                                 |      3.11        3.02
                         SIC75XX |     1.123       1.114
```

```
                                 |      2.03        1.90
                        MIXEDUSE |     1.375       1.355
                                 |      3.81        3.61
                       PC1FVLAND |     1.186       1.208
                                 |      2.84        3.09
                        Constant |     0.310       0.324
                                 |    -14.61      -13.61
---------------------------------+------------------------
lnalpha                          |
                        Constant |     0.906       0.891
                                 |     -1.34       -1.55
---------------------------------+------------------------
inflate                          |
                            WTFA |                 0.000
                                 |                 -1.29
                          BSTOPS |                 0.969
                                 |                 -0.07
                         SIC55XX |                 0.216
                                 |                 -0.51
                         SIC54XX |                 0.000
                                 |                 -0.01
                         SIC75XX |                 0.000
                                 |                 -0.05
                        MIXEDUSE |                 0.565
                                 |                 -0.05
                       PC1FVLAND |              9.20e+06
                                 |                  0.05
                        Constant |                 0.000
                                 |                 -0.03
---------------------------------+------------------------
Statistics                       |
                           alpha |     0.906
                               N |  2602.000    2602.000
                              ll | -3346.880   -3340.028
                             bic |  6764.537    6813.745
                             aic |  6711.761    6714.057
----------------------------------------------------------
                                                legend: b/t
Comparison of Mean Observed and Predicted Count
           Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.008         0      0.003
ZINBRM       -0.009         0      0.003

NBRM: Predicted and actual probabilities
Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.527       0.535      0.008     0.339
1        0.238       0.244      0.006     0.363
2        0.105       0.111      0.006     0.915
3        0.053       0.052      0.001     0.047
4        0.021       0.026      0.005     2.197
5        0.011       0.013      0.003     1.249
6        0.006       0.007      0.001     0.374
7        0.005       0.004      0.001     0.583
8        0.002       0.002      0.000     0.008
9        0.003       0.001      0.001     2.610
------------------------------------------------
Sum      0.971       0.996      0.032     8.684

ZINB: Predicted and actual probabilities
Count   Actual    Predicted    |Diff|   Pearson
```

```
-----------------------------------------------
0        0.527       0.535      0.009     0.376
1        0.238       0.242      0.004     0.186
2        0.105       0.112      0.007     1.107
3        0.053       0.053      0.000     0.007
4        0.021       0.026      0.005     2.461
5        0.011       0.013      0.003     1.342
6        0.006       0.007      0.001     0.376
7        0.005       0.004      0.001     0.634
8        0.002       0.002      0.000     0.001
9        0.003       0.001      0.001     2.902
-----------------------------------------------
Sum      0.971       0.997      0.031     9.392

Tests and Fit Statistics
------------------------------------------------------------------------
NBRM           BIC=-13697.684  AIC=     2.579  Prefer  Over  Evidence
------------------------------------------------------------------------
  vs ZINBRM    BIC=-13648.476  dif=   -49.208  NBRM    ZINBRM  Very strong
               AIC=     2.580  dif=    -0.001  NBRM    ZINB
               Vuong=   2.289  prob=    0.011  ZINBRM    NBRM  p=0.011
```

Table D- 31. Ad hoc NBRM Model showing discrete changes by predictors

```
nbreg (N=2602): Percentage Change in Expected Count
Observed SD: 1.5640296
----------------------------------------------------------------------
       tfa07 |      b         z     P>|z|      %      %StdX      SDofX
-------------+--------------------------------------------------------
        wtfa |   0.82312   14.160   0.000    127.8     46.6     0.4651
      bstops |   0.02825    2.598   0.009      2.9      7.6     2.5823
     sic55xx |   0.17821    2.101   0.036     19.5      5.9     0.3233
     sic54xx |   0.20711    3.112   0.002     23.0      8.5     0.3935
     sic75xx |   0.11607    2.032   0.042     12.3      5.7     0.4805
    mixeduse |   0.31840    3.810   0.000     37.5     10.6     0.3152
   pc1fvland |   0.17091    2.835   0.005     18.6      8.6     0.4811
-------------+--------------------------------------------------------
    ln alpha |  -0.09836
       alpha |   0.90632   SE(alpha) = 0.06640
----------------------------------------------------------------------
 LR test of alpha=0: 608.17   Prob>=LRX2 = 0.000
----------------------------------------------------------------------
       b = raw coefficient
       z = z-score for test of b=0
   P>|z| = p-value for z-test
       % = percent change in expected count for unit increase in X
   %StdX = percent change in expected count for SD increase in X
   SDofX = standard deviation of X
```

Appendix E. Data Analysis Results on Burglary

Table E- 1. OLS regression Model 1 output in GeoDa on burglary

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :      BUR07      Number of Observations: 2602
Mean dependent var  :   0.463874      Number of Variables   :    2
S.D. dependent var  :   0.900637      Degrees of Freedom    : 2600

R-squared           :   0.001434      F-statistic           :     3.73328
Adjusted R-squared  :   0.001050      Prob(F-statistic)     :     0.05345
Sum squared residual:    2107.58      Log Likelihood        :     -3417.9
Sigma-square        :   0.810607      Akaike info criterion :     6839.81
S.E. of regression  :   0.900337      Schwarz criterion     :     6851.54
Sigma-square ML     :   0.809984
S.E of regression ML:   0.899991
-------------------------------------------------------------------------------
Variable                Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                0.4538123     0.01840243       24.66045     0.0000000
Bus stops               0.01320908    0.006836395       1.932171    0.0534500
-------------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   1.337648
TEST ON NORMALITY OF ERRORS
TEST                     DF          VALUE             PROB
Jarque-Bera               2           20229.85        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF          VALUE             PROB
Breusch-Pagan test        1           1.350029        0.2452730
Koenker-Bassett test      1          0.1879399        0.6646368
SPECIFICATION ROBUST TEST
TEST                     DF          VALUE             PROB
White                     2             3.8475        0.1460582

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                            MI/DF      VALUE          PROB
Moran's I (error)            0.121533    27.7829837      0.0000000
Lagrange Multiplier (lag)        1      744.5878256      0.0000000
Robust LM (lag)                  1        6.7465465      0.0093929
Lagrange Multiplier (error)      1      753.6518430      0.0000000
Robust LM (error)                1       15.8105639      0.0000700
Lagrange Multiplier (SARMA)      2      760.3983895      0.0000000
========================= END OF REPORT ==============================
```

Table E- 2. OLS regression Model 2 output in GeoDa on burglary

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :       BUR07       Number of Observations: 2602
Mean dependent var  :    0.463874       Number of Variables   :    8
S.D. dependent var  :    0.900637       Degrees of Freedom    : 2594

R-squared           :    0.028863       F-statistic           :     11.0135
Adjusted R-squared  :    0.026242       Prob(F-statistic)     :8.79896e-014
Sum squared residual:     2049.69       Log Likelihood        :    -3381.67
Sigma-square        :    0.790164       Akaike info criterion :     6779.34
S.E. of regression  :    0.888912       Schwarz criterion     :     6826.25
Sigma-square ML     :    0.787735
S.E of regression ML:    0.887544
------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
------------------------------------------------------------------------------
Constant                 0.4167829      0.01972        21.13503      0.0000000
Eating/drinking places   0.02782588     0.0326318       0.852723     0.3938708
Automotive retail        0.004804472    0.05551017      0.0865512    0.9313552
Food store retail        0.3550165      0.047729        7.438171     0.0000000
Automotive service      -0.04278998     0.03753089     -1.140127     0.2543411
Business service         0.01071161     0.03135945      0.3415752    0.7326261
Personal service         0.03258503     0.03026891      1.076518     0.2817961
Banks                   -0.1060565      0.08624049     -1.229777     0.2188859
------------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   2.207984
TEST ON NORMALITY OF ERRORS
TEST                  DF          VALUE            PROB
Jarque-Bera            2           16945.04        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF          VALUE            PROB
Breusch-Pagan test     7           287.6912        0.0000000
Koenker-Bassett test   7           43.51743        0.0000003
SPECIFICATION ROBUST TEST
TEST                  DF          VALUE            PROB
White                 35           78.72214        0.0000328

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                          MI/DF       VALUE            PROB
Moran's I (error)           0.115276    26.4306830       0.0000000
Lagrange Multiplier (lag)       1      702.2009191       0.0000000
Robust LM (lag)                 1       24.3745530       0.0000008
Lagrange Multiplier (error)     1      678.0529771       0.0000000
Robust LM (error)               1        0.2266110       0.6340482
Lagrange Multiplier (SARMA)     2      702.4275300       0.0000000
========================= END OF REPORT ==============================
```

Table E- 3. OLS regression Model 3 output in GeoDa on burglary

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :       BUR07       Number of Observations: 2602
Mean dependent var  :    0.463874       Number of Variables   :    9
S.D. dependent var  :    0.900637       Degrees of Freedom    : 2593

R-squared           :    0.029172       F-statistic           :     9.73935
Adjusted R-squared  :    0.026176       Prob(F-statistic)     : 2.0946e-013
Sum squared residual:     2049.03       Log Likelihood        :    -3381.25
Sigma-square        :    0.790218       Akaike info criterion :     6780.51
S.E. of regression  :    0.888942       Schwarz criterion     :     6833.28
Sigma-square ML     :    0.787484
S.E of regression ML:    0.887403
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                 0.4138161      0.01998925     20.70194      0.0000000
Bus stops                0.00628836     0.00692212      0.9084443    0.3637264
Eating/drinking places   0.0251953      0.03276112      0.7690609    0.4419245
Automotive retail        0.00308041     0.05554447      0.05545844   0.9556666
Food store retail        0.3532719      0.04776922      7.395388     0.0000000
Automotive service      -0.0446678      0.03758903     -1.18832      0.2348063
Business service         0.009684209    0.0313809       0.3086021    0.7576811
Personal service         0.03033748     0.03037087      0.9989006    0.3179545
Banks                   -0.109008       0.08630457     -1.263062     0.2066743
-------------------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   2.283673
TEST ON NORMALITY OF ERRORS
TEST                     DF          VALUE            PROB
Jarque-Bera               2           16975.97        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF          VALUE            PROB
Breusch-Pagan test        8           288.5993        0.0000000
Koenker-Bassett test      8           43.61112        0.0000007
SPECIFICATION ROBUST TEST
TEST                     DF          VALUE            PROB
White                    44           81.54683        0.0004976

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                           MI/DF      VALUE          PROB
Moran's I (error)            0.115733    26.5604696      0.0000000
Lagrange Multiplier (lag)        1      703.7679960      0.0000000
Robust LM (lag)                  1       20.3365897      0.0000065
Lagrange Multiplier (error)      1      683.4327916      0.0000000
Robust LM (error)                1        0.0013853      0.9703099
Lagrange Multiplier (SARMA)      2      703.7693813      0.0000000
========================= END OF REPORT ==============================
```

Table E- 4. OLS regression Model 4 output in GeoDa on burglary

```
============================== BEGINNING OF REPORT============================
Dependent Variable  :       BUR07       Number of Observations: 2602
Mean dependent var  :    0.463874       Number of Variables   :   15
S.D. dependent var  :    0.900637       Degrees of Freedom    : 2587

R-squared           :    0.047477       F-statistic           :     9.21033
Adjusted R-squared  :    0.042322       Prob(F-statistic)     :3.94764e-020
Sum squared residual:      2010.4       Log Likelihood        :    -3356.49
Sigma-square        :    0.777116       Akaike info criterion :     6742.98
S.E. of regression  :    0.881542       Schwarz criterion     :     6830.94
Sigma-square ML     :    0.772636
S.E of regression ML:    0.878997
-------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
-------------------------------------------------------------------------------
Constant                 0.2813983      0.03297945      8.532535     0.0000000
Bus stops                0.009220147    0.006953038     1.32606      0.1849379
Eating/drinking places   0.02604145     0.03255129      0.8000127    0.4237565
Automotive retail        0.005690031    0.05520731      0.1030666    0.9178530
Food store retail        0.3311233      0.04752813      6.966892     0.0000000
Automotive service      -0.05114146     0.03735542     -1.369051     0.1711096
Business service         0.01816975     0.03121096      0.5821592    0.5604752
Personal service         0.03443068     0.03022369      1.139195     0.2547288
Banks                   -0.08405732     0.08573639     -0.9804159    0.3269864
Mixed landuse           -0.08629077     0.05538818     -1.557927     0.1193694
Vacant land              0.1979292      0.03632479      5.448875     0.0000001
Public housing           0.162233       0.05691534      2.850426     0.0044006
Colleges                -0.2057348      0.1218823      -1.687979     0.0915338
Grade K-12               0.1196506      0.05793604      2.065219     0.0390011
Parks and cemeteries    -0.06611477     0.04880549     -1.354658     0.1756366
-------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   4.446725
TEST ON NORMALITY OF ERRORS
TEST                     DF          VALUE            PROB
Jarque-Bera               2           16157.74        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF          VALUE            PROB
Breusch-Pagan test       14           432.1632        0.0000000
Koenker-Bassett test     14           66.71478        0.0000000
SPECIFICATION ROBUST TEST
TEST                     DF          VALUE            PROB
White                    119           N/A             N/A

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : TPSW1600ft.GWT  (row-standardized weights)
TEST                            MI/DF       VALUE          PROB
Moran's I (error)             0.102994    23.9951774      0.0000000
Lagrange Multiplier (lag)         1      589.3745488      0.0000000
Robust LM (lag)                   1       48.1378356      0.0000000
Lagrange Multiplier (error)       1      541.2620534      0.0000000
Robust LM (error)                 1        0.0253403      0.8735222
Lagrange Multiplier (SARMA)       2      589.3998890      0.0000000
========================= END OF REPORT ==============================
```

Table E- 5. SEM Model 1 output in GeoDa on burglary

```
============================== BEGINNING OF REPORT============================
Data set             : TPAll07DataZCNov10
Spatial Weight       : TPSW1600ft.GWT
Dependent Variable   :       BUR07      Number of Observations: 2602
Mean dependent var   :    0.463874      Number of Variables   :    2
S.D. dependent var   :    0.900637      Degree of Freedom     : 2600
Lag coeff. (Lambda)  :    0.703521

R-squared            :    0.108711      R-squared (BUSE)      : -
Sq. Correlation      : -                Log likelihood        :-3293.818754
Sigma-square         :    0.722967      Akaike info criterion : 6591.64
S.E of regression    :    0.850274      Schwarz criterion     : 6603.365580
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    z-value      Probability
-----------------------------------------------------------------------
    CONSTANT    0.4481377     0.05642725       7.941867    0.0000000
      BSTOPS   0.01777563     0.006624646       2.683257    0.0072910
      LAMBDA    0.7035212     0.04043008       17.40093    0.0000000
-----------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF     VALUE         PROB
Breusch-Pagan test                       1       2.717992     0.0992228

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                     DF     VALUE         PROB
Likelihood Ratio Test                    1       248.1706     0.0000000
========================= END OF REPORT ======================================
```

Table E- 6. SLM Model 2 output in GeoDa on burglary

```
============================== BEGINNING OF REPORT============================
Spatial Weight       : TPSW1600ft.GWT
Dependent Variable   :       BUR07      Number of Observations: 2602
Mean dependent var   :    0.463874      Number of Variables   :    9
S.D. dependent var   :    0.900637      Degrees of Freedom    : 2593
Lag coeff.   (Rho)   :    0.691362

R-squared            :    0.128882      Log Likelihood        :    -3262.96
Sq. Correlation      : -                Akaike info criterion :     6543.92
Sigma-square         :    0.706605      Schwarz criterion     :     6596.69
S.E of regression    :    0.840598
------------------------------------------------------------------------------
Variable                 Coefficient    Std.Error    t-Statistic   Probability
------------------------------------------------------------------------------
Spatial lag              0.6913619      0.04079273     16.94816     0.0000000
Constant                 0.0919695      0.02636481      3.488342    0.0004861
Eating/drinking places   0.03475052     0.03086639      1.125837    0.2602346
Automotive retail        0.03097911     0.05249312      0.5901556   0.5550862
Food store retail        0.3020495      0.04514306      6.690939    0.0000000
Automotive service      -0.02514171     0.0354925      -0.7083667   0.4787174
Business service         0.03665965     0.02965716      1.236115    0.2164161
Personal service         0.02867695     0.02862697      1.001746    0.3164662
Banks                   -0.0518077      0.08155389     -0.6352572   0.5252605
------------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
```

```
TEST                                      DF     VALUE         PROB
Breusch-Pagan test                        7       286.1114     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                      DF     VALUE         PROB
Likelihood Ratio Test                     1       237.4188     0.0000000
========================= END OF REPORT ======================================
```

Table E- 7. SLM Model 3 output in GeoDa on burglary

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :       BUR07      Number of Observations: 2602
Mean dependent var  :    0.463874      Number of Variables   :   10
S.D. dependent var  :    0.900637      Degrees of Freedom    : 2592
Lag coeff.   (Rho)  :    0.692348

R-squared           :    0.129487      Log Likelihood        :    -3262.14
Sq. Correlation     : -                Akaike info criterion :     6544.28
Sigma-square        :    0.706114      Schwarz criterion     :     6602.92
S.E of regression   :    0.840306
-----------------------------------------------------------------------------
Variable                Coefficient    Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------------
Spatial lag             0.6923482      0.04063315     17.039         0.0000000
Constant                0.08755438     0.02643301      3.312312      0.0009254
Bus stops               0.008376046    0.006544821     1.279798      0.2006165
Eating/drinking places  0.03125648     0.03097749      1.009006      0.3129717
Automotive retail       0.02872001     0.05250589      0.5469865     0.5843879
Food store retail       0.2996502      0.04516284      6.634885      0.0000000
Automotive service     -0.02761777     0.0355333      -0.7772364     0.4370192
Business service        0.03532817     0.02966772      1.190795      0.2337341
Personal service        0.02567766     0.02871167      0.8943281     0.3711463
Banks                  -0.05566167     0.08158295     -0.6822709     0.4950675
-----------------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                      DF     VALUE         PROB
Breusch-Pagan test                        8       287.2695     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                      DF     VALUE         PROB
Likelihood Ratio Test                     1       238.2282     0.0000000
========================= END OF REPORT ======================================
```

Table E- 8. SLM Model 4 output in GeoDa on burglary

```
============================== BEGINNING OF REPORT============================
Spatial Weight      : TPSW1600ft.GWT
Dependent Variable  :       BUR07      Number of Observations: 2602
Mean dependent var  :    0.463874      Number of Variables   :   16
S.D. dependent var  :    0.900637      Degrees of Freedom    : 2586
Lag coeff.   (Rho)  :    0.668053

R-squared           :    0.136565      Log likelihood        : -3249.48
Sq. Correlation     : -                Akaike info criterion :  6530.96
```

```
Sigma-square        :    0.700372       Schwarz criterion     :  6624.79
S.E of regression   :    0.836882
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    z-value      Probability
-----------------------------------------------------------------------
     W_BUR07     0.6680534     0.04207598       15.87731    0.0000000
    CONSTANT    0.02020398     0.03606055      0.5602792    0.5752889
      BSTOPS    0.01069813    0.006601912       1.620459    0.1051338
     SIC58XX    0.03096495     0.03090514       1.001935    0.3163749
     SIC55XX    0.03027268     0.05241048      0.5776073    0.5635292
     SIC54XX     0.2855058     0.04513801       6.325174    0.0000000
     SIC75XX   -0.03353401     0.03546333     -0.9455967    0.3443543
     SIC73XX    0.04015053      0.0296298       1.355072    0.1753946
     SIC72XX     0.0281423     0.02869255      0.9808225    0.3266803
     FINANCE   -0.04256432     0.08139296     -0.5229484    0.6010101
    MIXEDUSE   -0.04391984     0.05258221     -0.8352605    0.4035710
   PC1FVLAND     0.1235528     0.03455293       3.575756    0.0003493
    PHOUSING     0.1255667     0.05412404        2.31998    0.0203419
    UNIVCOLL    -0.1079247      0.1157525     -0.9323748    0.3511428
     SCHOOLS    0.06813892     0.05502301       1.238371    0.2155785
   OPENSPACE   -0.08848684     0.04635606      -1.908852    0.0562811
-----------------------------------------------------------------------
REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                    DF     VALUE         PROB
Breusch-Pagan test                     14       421.5651     0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : TPSW1600ft.GWT
TEST                                    DF     VALUE         PROB
Likelihood Ratio Test                   1       214.0136     0.0000000
========================= END OF REPORT ======================================
```

Table E- 9. PRM Model 1 output in Stata on burglary

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(2)      =     404.80
                                                  Prob > chi2     =     0.0000
Log likelihood = -2333.6126                       Pseudo R2       =     0.0798
------------------------------------------------------------------------------
       bur07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wbur |   1.397351   .0625182    22.35   0.000     1.274818    1.519884
      bstops |   .0326466   .0088967     3.67   0.000     .0152094    .0500838
       _cons |  -1.575516   .0523479   -30.10   0.000    -1.678116   -1.472916
------------------------------------------------------------------------------
```

Table E- 10. PRM Model 2 output in Stata on burglary

Poisson regression	Number of obs	=	2602
	LR chi2(8)	=	466.95
	Prob > chi2	=	0.0000
Log likelihood = -2302.5402	Pseudo R2	=	0.0921

bur07	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wbur	1.364159	.0635741	21.46	0.000	1.239556	1.488762
sic58xx	.0512501	.0466861	1.10	0.272	-.040253	.1427533
sic55xx	.1078632	.0866584	1.24	0.213	-.0619841	.2777105
sic54xx	.35261	.0530163	6.65	0.000	.2486999	.45652
sic75xx	-.0561893	.0684669	-0.82	0.412	-.190382	.0780034
sic73xx	.0814385	.0410541	1.98	0.047	.0009739	.1619032
sic72xx	.0609917	.0424931	1.44	0.151	-.0222932	.1442766
finance	-.0056993	.1355528	-0.04	0.966	-.2713778	.2599793
_cons	-1.628106	.0545891	-29.82	0.000	-1.735099	-1.521113

Table E- 11. PRM Model 3 output in Stata on burglary

Poisson regression	Number of obs	=	2602
	LR chi2(9)	=	470.70
	Prob > chi2	=	0.0000
Log likelihood = -2300.6624	Pseudo R2	=	0.0928

bur07	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wbur	1.370076	.0636755	21.52	0.000	1.245274	1.494878
bstops	.0213606	.0102965	2.07	0.038	.0011799	.0415414
sic58xx	.0434842	.0470923	0.92	0.356	-.048815	.1357834
sic55xx	.097858	.0876512	1.12	0.264	-.0739352	.2696512
sic54xx	.3474326	.0531592	6.54	0.000	.2432425	.4516227
sic75xx	-.062006	.0685993	-0.90	0.366	-.1964582	.0724462
sic73xx	.081177	.0415608	1.95	0.051	-.0002807	.1626346
sic72xx	.0539664	.042696	1.26	0.206	-.0297163	.1376491
finance	-.022673	.1364606	-0.17	0.868	-.2901309	.2447849
_cons	-1.643145	.0552718	-29.73	0.000	-1.751475	-1.534814

Table E- 12. PRM Model 4 output in Stata on burglary

```
Poisson regression                                Number of obs   =       2602
                                                  LR chi2(15)     =     505.07
                                                  Prob > chi2     =     0.0000
Log likelihood = -2283.4804                       Pseudo R2       =     0.0996
------------------------------------------------------------------------------
       bur07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wbur |   1.302041   .0666627    19.53   0.000     1.171385    1.432698
      bstops |   .0265764   .0104411     2.55   0.011     .0061122    .0470406
     sic58xx |   .0382203   .0475477     0.80   0.421    -.0549714    .1314121
     sic55xx |   .0986804   .0886543     1.11   0.266    -.0750788    .2724396
     sic54xx |   .3279138   .0534188     6.14   0.000      .223215    .4326126
     sic75xx |  -.0712178   .0685493    -1.04   0.299    -.2055719    .0631364
     sic73xx |   .0978133   .0423159     2.31   0.021     .0148756     .180751
     sic72xx |   .0556078   .0432242     1.29   0.198      -.02911    .1403256
     finance |  -.0030331    .138702    -0.02   0.983    -.2748841    .2688179
    mixeduse |  -.0577571    .096278    -0.60   0.549    -.2464584    .1309443
   pc1fvland |   .2506654   .0681318     3.68   0.000     .1171295    .3842013
    phousing |   .2698056   .0847422     3.18   0.001      .103714    .4358972
    univcoll |  -.5195032   .2821442    -1.84   0.066    -1.072496    .0334893
     schools |   .0569095   .0885856     0.64   0.521    -.1167151    .2305341
   openspace |  -.1913719   .0883685    -2.17   0.030     -.364571   -.0181727
       _cons |  -1.777966   .0714739   -24.88   0.000    -1.918052    -1.63788
------------------------------------------------------------------------------
```

Table E- 13. NBRM Model 1 output in Stata on burglary

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(2)      =     257.60
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2241.0895                       Pseudo R2       =     0.0543

------------------------------------------------------------------------------
       bur07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wbur |   1.568898   .0987613    15.89   0.000     1.375329    1.762466
      bstops |   .0381101   .0132641     2.87   0.004      .012113    .0641072
       _cons |   -1.67672   .0688379   -24.36   0.000     -1.81164     -1.5418
-------------+----------------------------------------------------------------
    /lnalpha |  -.0665841   .1147474                     -.2914849    .1583168
-------------+----------------------------------------------------------------
       alpha |   .9355843   .1073559                      .7471533    1.171537
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  185.05 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2369.888    -2241.09      4     4490.179    4513.635
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table E- 14. NBRM Model 2 output in Stata on burglary

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(8)      =     294.13
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2222.8219                       Pseudo R2       =     0.0621
------------------------------------------------------------------------------
       bur07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wbur |   1.525724    .097647    15.62   0.000      1.33434    1.717109
     sic58xx |   .0629867   .0614915     1.02   0.306    -.0575344    .1835078
     sic55xx |   .1397723   .1074337     1.30   0.193    -.0707939    .3503385
     sic54xx |   .3339438    .078383     4.26   0.000     .1803159    .4875717
     sic75xx |  -.0643158   .0838636    -0.77   0.443    -.2286854    .1000538
     sic73xx |   .0873252   .0570473     1.53   0.126    -.0244854    .1991357
     sic72xx |   .0914647   .0553405     1.65   0.098    -.0170006      .19993
     finance |  -.0237969   .1743573    -0.14   0.891     -.365531    .3179372
       _cons |  -1.727933   .0706473   -24.46   0.000      -1.8664   -1.589467
-------------+----------------------------------------------------------------
    /lnalpha |  -.1617067   .1210705                     -.3990006    .0755871
-------------+----------------------------------------------------------------
       alpha |   .8506907   .1029936                      .6709903    1.078517
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  159.44 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2369.888   -2222.822     10     4465.644    4524.284
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table E- 15. NBRM Model 3 output in Stata on burglary

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(9)      =     296.59
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2221.5939                       Pseudo R2       =     0.0626
------------------------------------------------------------------------------
       bur07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wbur |    1.53122   .0976745    15.68   0.000     1.339781    1.722658
      bstops |   .0219768   .0136805     1.61   0.108    -.0048365    .0487901
     sic58xx |   .0501908   .0620165     0.81   0.418    -.0713592    .1717408
     sic55xx |   .1333689    .107645     1.24   0.215    -.0776114    .3443491
     sic54xx |     .33138   .0782726     4.23   0.000     .1779686    .4847914
     sic75xx |  -.0706088   .0839056    -0.84   0.400    -.2350608    .0938432
     sic73xx |   .0851765   .0570204     1.49   0.135    -.0265813    .1969344
     sic72xx |   .0810579   .0556141     1.46   0.145    -.0279436    .1900595
     finance |   -.035853   .1747929    -0.21   0.837    -.3784408    .3067348
       _cons |  -1.741719   .0712487   -24.45   0.000    -1.881364   -1.602074
-------------+----------------------------------------------------------------
    /lnalpha |  -.1672231   .1214284                     -.4052184    .0707722
-------------+----------------------------------------------------------------
       alpha |   .8460108   .1027297                      .6668311    1.073337
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  158.14 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2369.888   -2221.594     11     4465.188    4529.692
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table E- 16. NBRM Model 4 output in Stata on burglary

```
Negative binomial regression                      Number of obs   =       2602
                                                  LR chi2(15)     =     317.85
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2210.9648                       Pseudo R2       =     0.0671
------------------------------------------------------------------------------
       bur07 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wbur |   1.456667   .0993069    14.67   0.000     1.262029    1.651305
      bstops |   .0278821   .0137916     2.02   0.043     .0008511    .0549131
     sic58xx |   .0453187   .0617444     0.73   0.463    -.0756981    .1663355
     sic55xx |   .1336965   .1074262     1.24   0.213    -.0768549    .3442479
     sic54xx |   .3070132   .0775854     3.96   0.000     .1549485    .4590778
     sic75xx |  -.0833031   .0831937    -1.00   0.317    -.2463597    .0797535
     sic73xx |   .1015404   .0570227     1.78   0.075    -.0102222    .2133029
     sic72xx |   .0835601   .0556334     1.50   0.133    -.0254794    .1925996
     finance |  -.0186096     .17441    -0.11   0.915     -.360447    .3232277
    mixeduse |  -.0148542   .1151863    -0.13   0.897    -.2406153    .2109068
   pc1fvland |   .2384625   .0795341     3.00   0.003     .0825784    .3943466
    phousing |   .2764758   .1064775     2.60   0.009     .0677838    .4851678
    univcoll |  -.3854004   .3221001    -1.20   0.231    -1.016705    .2459041
     schools |   .0104189   .1138291     0.09   0.927    -.2126821      .23352
   openspace |  -.1759782   .1045683    -1.68   0.092    -.3809282    .0289719
       _cons |   -1.86822   .0877993   -21.28   0.000    -2.040303   -1.696136
-------------+----------------------------------------------------------------
    /lnalpha |  -.2200116   .1251974                      -.465394    .0253708
-------------+----------------------------------------------------------------
       alpha |   .8025095   .1004721                      .6278876    1.025695
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  145.03 Prob>=chibar2 = 0.000
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |   2602   -2369.888   -2210.965     17      4455.93    4555.618
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Table E- 17. ZINBRM Model 1 output in Stata on burglary

```
Zero-inflated negative binomial regression        Number of obs   =       2602
                                                  Nonzero obs     =        777
                                                  Zero obs        =       1825
Inflation model = logit                           LR chi2(2)      =      92.54
Log likelihood  = -2225.742                       Prob > chi2     =     0.0000
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
bur07        |
        wbur |   1.099278   .1312409     8.38   0.000     .8420511    1.356506
      bstops |   .0685478   .0187745     3.65   0.000     .0317505    .1053451
       _cons |  -1.215534   .1290071    -9.42   0.000    -1.468383    -.962685
-------------+----------------------------------------------------------------
inflate      |
        wbur |  -5.475533   1.306114    -4.19   0.000    -8.035469   -2.915597
      bstops |   .0616541   .0348864     1.77   0.077     -.006722    .1300301
       _cons |   .6964182   .3571146     1.95   0.051    -.0035136     1.39635
-------------+----------------------------------------------------------------
    /lnalpha |  -.4695622   .1808084    -2.60   0.009    -.8239402   -.1151842
-------------+----------------------------------------------------------------
       alpha |   .6252759   .1130552                      .4386997     .891202
```

Table E- 18. ZINBRM Model 2 output in Stata on burglary

```
Zero-inflated negative binomial regression        Number of obs    =       2602
                                                  Nonzero obs      =        777
                                                  Zero obs         =       1825

Inflation model = logit                           LR chi2(8)       =     123.24
Log likelihood  = -2198.927                       Prob > chi2      =     0.0000
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
bur07        |
        wbur |   1.134315   .1161499     9.77   0.000     .9066658    1.361965
     sic58xx |   .1119157   .0693515     1.61   0.107    -.0240107    .2478421
     sic55xx |   .0150031    .106354     0.14   0.888    -.1934468    .2234531
     sic54xx |   .2765683    .079009     3.50   0.000     .1217136     .431423
     sic75xx |  -.1704119   .0841316    -2.03   0.043    -.3353069    -.005517
     sic73xx |    .251014   .0754876     3.33   0.001      .103061     .398967
     sic72xx |   .0391328   .0558687     0.70   0.484    -.0703678    .1486335
     finance |  -.2439771   .1760213    -1.39   0.166    -.5889726    .1010184
       _cons |  -1.312416   .1113304   -11.79   0.000     -1.53062   -1.094212
-------------+----------------------------------------------------------------
inflate      |
        wbur |  -7.519324    1.63318    -4.60   0.000     -10.7203    -4.31835
     sic58xx |   .3084632   .3000621     1.03   0.304    -.2796479    .8965742
     sic55xx |   -26.6467   1498.887    -0.02   0.986     -2964.41    2911.117
     sic54xx |  -.7820281   .9594694    -0.82   0.415    -2.662554    1.098497
     sic75xx |  -2.765904   1.436345    -1.93   0.054    -5.581088    .0492801
     sic73xx |   .8547872   .3499557     2.44   0.015     .1688866    1.540688
     sic72xx |  -.6827077   .4204115    -1.62   0.104    -1.506699    .1412838
     finance |   -6.00347   7.120124    -0.84   0.399    -19.95866    7.951717
       _cons |   1.326981   .4075994     3.26   0.001      .528101    2.125861
-------------+----------------------------------------------------------------
    /lnalpha |  -.4768597   .1562045    -3.05   0.002     -.783015   -.1707045
-------------+----------------------------------------------------------------
       alpha |   .6207296   .0969608                       .457026    .8430707
------------------------------------------------------------------------------
```

Table E- 19. ZINBRM Model 3 output in Stata on burglary

```
Zero-inflated negative binomial regression        Number of obs    =       2602
                                                  Nonzero obs      =        777
                                                  Zero obs         =       1825
Inflation model = logit                           LR chi2(9)       =     121.99
Log likelihood  = -2199.534                       Prob > chi2      =     0.0000
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
bur07        |
        wbur |   1.108708   .1253191     8.85   0.000     .8630871    1.354329
      bstops |   .0370305   .0193286     1.92   0.055    -.0008529    .0749138
     sic58xx |   .0469415   .0687154     0.68   0.495    -.0877382    .1816212
     sic55xx |   .1117429   .2127367     0.53   0.599    -.3052133    .5286992
     sic54xx |   .2650024   .0794867     3.33   0.001     .1092113    .4207935
     sic75xx |  -.1505367   .0983315    -1.53   0.126     -.343263    .0421895
     sic73xx |   .2496972   .0779453     3.20   0.001     .0969272    .4024672
     sic72xx |   .0158264   .0572549     0.28   0.782    -.0963912    .1280439
     finance |  -.2661209   .1770517    -1.50   0.133    -.6131359    .0808941
       _cons |  -1.284816   .1264057   -10.16   0.000    -1.532567   -1.037066
-------------+----------------------------------------------------------------
inflate      |
        wbur |  -6.180579   1.546472    -4.00   0.000    -9.211608   -3.149549
```

```
      bstops |   .0595114   .0504836     1.18   0.238    -.0394346    .1584574
     sic58xx |  -.0883901   .3557835    -0.25   0.804     -.785713    .6089327
     sic55xx |  -.0588579   .8790516    -0.07   0.947    -1.781767    1.664052
     sic54xx |  -.9701652   1.207157    -0.80   0.422     -3.33615    1.395819
     sic75xx |  -.7764346   .8671506    -0.90   0.371    -2.476018    .9231493
     sic73xx |   .7057271    .276682     2.55   0.011     .1634404    1.248014
     sic72xx |  -.7547203   .4066589    -1.86   0.063    -1.551757    .0423166
     finance |  -6.499281   8.622511    -0.75   0.451    -23.39909    10.40053
       _cons |   1.051091   .4203889     2.50   0.012     .2271439    1.875038
-------------+----------------------------------------------------------------
    /lnalpha |  -.5407389   .1776553    -3.04   0.002    -.8889369    -.192541
-------------+----------------------------------------------------------------
       alpha |   .5823178   .1034518                      .4110926    .8248605
------------------------------------------------------------------------------
```

Table E- 20. ZINBRM Model 4 output in Stata on burglary

```
Zero-inflated negative binomial regression        Number of obs   =       2602
                                                  Nonzero obs     =        777
                                                  Zero obs        =       1825
Inflation model = logit                           LR chi2(15)     =     139.04
Log likelihood  = -2184.449                       Prob > chi2     =     0.0000
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
bur07        |
        wbur |   1.087446   .1264531     8.60   0.000     .8396021    1.335289
      bstops |   .0557381   .0191954     2.90   0.004     .0181158    .0933605
     sic58xx |   .0267595    .067637     0.40   0.692    -.1058065    .1593255
     sic55xx |   .0262122   .1181604     0.22   0.824    -.2053779    .2578022
     sic54xx |   .2642438   .0775099     3.41   0.001     .1123272    .4161604
     sic75xx |  -.1510147   .0917808    -1.65   0.100    -.3309016    .0288723
     sic73xx |   .2665636   .0718981     3.71   0.000     .1256459    .4074814
     sic72xx |   .0151684    .058454     0.26   0.795    -.0993993    .1297361
     finance |  -.1778873   .1776144    -1.00   0.317    -.5260051    .1702305
    mixeduse |  -.0469803   .1319717    -0.36   0.722    -.3056401    .2116795
   pc1fvland |   .2241027   .0955925     2.34   0.019     .0367448    .4114605
    phousing |   .2282579   .1193509     1.91   0.056    -.0056656    .4621814
    univcoll |   .2290944   .4824258     0.47   0.635    -.7164428    1.174632
     schools |  -.0046305   .1189471    -0.04   0.969    -.2377625    .2285014
   openspace |  -.3727951    .112949    -3.30   0.001    -.5941711   -.1514191
       _cons |  -1.440522   .1382173   -10.42   0.000    -1.711423   -1.169621
-------------+----------------------------------------------------------------
inflate      |
        wbur |  -7.793155   1.976006    -3.94   0.000    -11.66606   -3.920255
      bstops |   .1546213    .058963     2.62   0.009     .0390559    .2701867
     sic58xx |  -.3105542   .5118145    -0.61   0.544    -1.313692    .6925838
     sic55xx |  -1.353326   2.185259    -0.62   0.536    -5.636355    2.929702
     sic54xx |  -.5689332   .9592355    -0.59   0.553       -2.449    1.311134
     sic75xx |  -1.081521   .8597786    -1.26   0.208    -2.766656    .6036136
     sic73xx |   .7777573   .3390896     2.29   0.022      .113154    1.442361
     sic72xx |  -.9096151   .4214808    -2.16   0.031    -1.735702   -.0835278
     finance |  -2.147619   1.493522    -1.44   0.150    -5.074869    .7796299
    mixeduse |  -.1093716   .7231289    -0.15   0.880    -1.526678    1.307935
   pc1fvland |  -.2507968    .478546    -0.52   0.600     -1.18873    .6871361
    phousing |  -.1997765   .8285872    -0.24   0.809    -1.823778    1.424225
    univcoll |   1.198361   .9002061     1.33   0.183    -.5660107    2.962732
     schools |  -.8213806   .8558584    -0.96   0.337    -2.498832     .856071
   openspace |   -3.00905   1.963209    -1.53   0.125    -6.856868    .8387685
       _cons |   1.700029   .5673198     3.00   0.003     .5881022    2.811955
-------------+----------------------------------------------------------------
```

```
    /lnalpha |  -.5470846   .1726968    -3.17   0.002    -.8855641   -.2086051
-------------+----------------------------------------------------------------
       alpha |   .5786343   .0999283                      .4124814    .8117157
------------------------------------------------------------------------------
```

Table E- 21. Model fit comparison of PRM & NBRM model 1 on burglary

```
---------------------------------------------------------
                        Variable |    PRM        NBRM    
---------------------------------+-----------------------
bur07                            |                       
                            WBUR |     4.044       4.801 
                                 |     22.35       15.89 
                          BSTOPS |     1.033       1.039 
                                 |      3.67        2.87 
                        Constant |     0.207       0.187 
                                 |    -30.10      -24.36 
---------------------------------+-----------------------
lnalpha                          |                       
                        Constant |                 0.936 
                                 |                 -0.58 
---------------------------------+-----------------------
Statistics                       |                       
                           alpha |                 0.936 
                               N |  2602.000    2602.000 
                              ll | -2333.613   -2241.090 
                             bic |  4690.817    4513.635 
                             aic |  4673.225    4490.179 
---------------------------------------------------------
                                             legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM        -0.079         1      0.014
NBRM       -0.017         1      0.003

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.701       0.653      0.048     9.210
1        0.181       0.261      0.079    62.807
2        0.064       0.065      0.001     0.029
3        0.022       0.015      0.007     8.143
4        0.006       0.004      0.002     2.724
5        0.004       0.001      0.002    11.941
6        0.001       0.001      0.001     2.095
7        0.001       0.000      0.001     4.067
8        0.000       0.000      0.000     0.212
9        0.000       0.000      0.000    10.208
------------------------------------------------
Sum      0.981       1.000      0.141   111.438

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.701       0.704      0.002     0.020
1        0.181       0.198      0.017     3.641
```

2	0.064	0.060	0.003	0.525
3	0.022	0.021	0.001	0.155
4	0.006	0.008	0.002	1.405
5	0.004	0.004	0.000	0.004
6	0.001	0.002	0.001	0.821
7	0.001	0.001	0.000	0.247
8	0.000	0.001	0.001	1.717
9	0.000	0.000	0.000	0.009
Sum	0.981	0.999	0.028	8.544

Tests and Fit Statistics

PRM	BIC=-15771.403	AIC= 1.796	Prefer	Over	Evidence
vs NBRM	BIC=-15948.586	dif= 177.182	NBRM	PRM	Very strong
	AIC= 1.726	dif= 0.070	NBRM	PRM	
	LRX2= 185.046	prob= 0.000	NBRM	PRM	p=0.000
NBRM	BIC=-15948.586	AIC= 1.726	Prefer	Over	Evidence

Table E- 22. Model fit comparison of PRM & NBRM model 2 on burglary

	Variable	PRM	NBRM
bur07			
	WBUR	3.913	4.601
		21.51	15.64
	SIC58XX	1.052	1.064
		1.10	1.02
	SIC55XX	1.114	1.151
		1.25	1.31
	SIC54XX	1.423	1.395
		6.65	4.27
	SIC75XX	0.945	0.937
		-0.82	-0.78
	SIC73XX	1.085	1.090
		1.98	1.53
	SIC72XX	1.063	1.095
		1.44	1.65
	Constant	0.196	0.178
		-29.91	-24.50
lnalpha			
	Constant		0.851
			-1.34
Statistics			
	alpha		0.851
	N	2602.000	2602.000
	ll	-2302.541	-2222.831
	bic	4667.995	4516.439
	aic	4621.082	4463.663

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|

PRM	-0.076	1	0.013
NBRM	-0.018	1	0.003

PRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.701	0.657	0.045	7.862
1	0.181	0.257	0.076	57.770
2	0.064	0.064	0.000	0.000
3	0.022	0.015	0.007	7.522
4	0.006	0.004	0.002	1.932
5	0.004	0.002	0.002	8.671
6	0.001	0.001	0.000	0.938
7	0.001	0.000	0.000	1.840
8	0.000	0.000	0.000	0.365
9	0.000	0.000	0.000	4.332
Sum	0.981	1.000	0.132	91.231

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.701	0.703	0.002	0.010
1	0.181	0.200	0.018	4.404
2	0.064	0.060	0.004	0.680
3	0.022	0.020	0.002	0.293
4	0.006	0.008	0.002	1.206
5	0.004	0.004	0.000	0.014
6	0.001	0.002	0.001	0.787
7	0.001	0.001	0.000	0.242
8	0.000	0.001	0.001	1.724
9	0.000	0.000	0.000	0.011
Sum	0.981	0.999	0.029	9.372

Tests and Fit Statistics

PRM	BIC=-15794.226	AIC= 1.776	Prefer	Over	Evidence
vs NBRM	BIC=-15945.782	dif= 151.556	NBRM	PRM	Very strong
	AIC= 1.715	dif= 0.060	NBRM	PRM	
	LRX2= 159.420	prob= 0.000	NBRM	PRM	p=0.000
NBRM	BIC=-15945.782	AIC= 1.715	Prefer	Over	Evidence

Table E- 23. Model fit comparison of PRM & NBRM model 3 on burglary

Variable	PRM	NBRM
bur07		
WBUR	3.938	4.628
	21.57	15.70
BSTOPS	1.021	1.022
	2.07	1.60
SIC58XX	1.043	1.050
	0.91	0.79
SIC55XX	1.103	1.143
	1.12	1.25

```
                         SIC54XX |     1.415       1.391   
                                 |      6.53        4.23   
                         SIC75XX |     0.939       0.931   
                                 |     -0.91       -0.85   
                         SIC73XX |     1.084       1.088   
                                 |      1.95        1.48   
                         SIC72XX |     1.055       1.084   
                                 |      1.25        1.45   
                        Constant |     0.193       0.175   
                                 |    -29.80      -24.48   
---------------------------------+------------------------
lnalpha                          |
                        Constant |                 0.846   
                                 |                 -1.38   
---------------------------------+------------------------
Statistics                       |
                           alpha |                 0.846   
                               N |  2602.000    2602.000   
                              ll | -2300.676   -2221.615   
                             bic |  4672.129    4521.870   
                             aic |  4619.353    4463.230   
----------------------------------------------------------
                                                legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM        -0.075         1      0.013
NBRM       -0.018         1      0.003

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.701       0.657      0.044     7.790
1        0.181       0.257      0.075    57.384
2        0.064       0.064      0.000     0.000
3        0.022       0.015      0.007     7.402
4        0.006       0.004      0.002     1.880
5        0.004       0.002      0.002     8.560
6        0.001       0.001      0.000     0.924
7        0.001       0.000      0.000     1.841
8        0.000       0.000      0.000     0.363
9        0.000       0.000      0.000     4.415
------------------------------------------------
Sum      0.981       1.000      0.132    90.560

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.701       0.703      0.002     0.010
1        0.181       0.200      0.018     4.422
2        0.064       0.060      0.004     0.679
3        0.022       0.020      0.002     0.292
4        0.006       0.008      0.002     1.206
5        0.004       0.004      0.000     0.014
6        0.001       0.002      0.001     0.782
7        0.001       0.001      0.000     0.238
8        0.000       0.001      0.001     1.719
9        0.000       0.000      0.000     0.010
------------------------------------------------
```

```
Sum      0.981       0.999      0.029     9.373

Tests and Fit Statistics

PRM            BIC=-15790.092  AIC=     1.775  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM      BIC=-15940.350  dif=   150.259  NBRM    PRM   Very strong
               AIC=     1.715  dif=     0.060  NBRM    PRM
               LRX2=  158.123  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM           BIC=-15940.350  AIC=     1.715  Prefer  Over  Evidence
```

Table E- 24. Model fit comparison of PRM & NBRM model 4 on burglary

Variable	PRM	NBRM
bur07		
WBUR	3.677	4.293
	19.55	14.68
BSTOPS	1.027	1.028
	2.55	2.02
SIC58XX	1.039	1.046
	0.81	0.73
SIC55XX	1.104	1.144
	1.11	1.25
SIC54XX	1.388	1.358
	6.14	3.96
SIC75XX	0.931	0.920
	-1.04	-1.01
SIC73XX	1.103	1.106
	2.31	1.78
SIC72XX	1.057	1.087
	1.29	1.50
MIXEDUSE	0.944	0.985
	-0.60	-0.13
PC1FVLAND	1.285	1.270
	3.68	3.00
PHOUSING	1.310	1.319
	3.18	2.60
UNIVCOLL	0.595	0.680
	-1.84	-1.20
SCHOOLS	1.059	1.010
	0.64	0.09
OPENSPACE	0.826	0.839
	-2.17	-1.68
Constant	0.169	0.154
	-24.92	-21.30
lnalpha		
Constant		0.803
		-1.76
Statistics		
alpha		0.803
N	2602.000	2602.000
ll	-2283.481	-2210.971
bic	4684.922	4547.766
aic	4596.961	4453.941

legend: b/t

```
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
PRM        -0.073         1      0.013
NBRM       -0.019         1      0.003

PRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.701       0.659      0.043     7.154
1        0.181       0.254      0.073    53.964
2        0.064       0.064      0.000     0.004
3        0.022       0.016      0.006     6.357
4        0.006       0.005      0.002     1.472
5        0.004       0.002      0.002     7.817
6        0.001       0.001      0.000     0.842
7        0.001       0.000      0.000     1.827
8        0.000       0.000      0.000     0.355
9        0.000       0.000      0.000     4.735
------------------------------------------------
Sum      0.981       1.000      0.127    84.527

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.701       0.703      0.001     0.006
1        0.181       0.200      0.019     4.672
2        0.064       0.060      0.004     0.676
3        0.022       0.020      0.002     0.292
4        0.006       0.008      0.002     1.202
5        0.004       0.004      0.000     0.017
6        0.001       0.002      0.001     0.749
7        0.001       0.001      0.000     0.212
8        0.000       0.001      0.001     1.670
9        0.000       0.000      0.000     0.004
------------------------------------------------
Sum      0.981       0.999      0.029     9.500

Tests and Fit Statistics

PRM            BIC=-15777.299  AIC=     1.767  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM      BIC=-15914.455  dif=   137.156  NBRM    PRM   Very strong
               AIC=     1.712  dif=     0.055  NBRM    PRM
               LRX2=  145.020  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM           BIC=-15914.455  AIC=     1.712  Prefer  Over  Evidence
```

Table E- 25. Model fit comparison of NBRM & ZINBRM model 1 on burglary

Variable	NBRM	ZINBRM
bur07		
WBUR	4.801	3.002
	15.89	8.38
BSTOPS	1.039	1.071
	2.87	3.65
Constant	0.187	0.297
	-24.36	-9.42
lnalpha		
Constant	0.936	0.625
	-0.58	-2.60
inflate		
WBUR		0.004
		-4.19
BSTOPS		1.064
		1.77
Constant		2.007
		1.95
Statistics		
alpha	0.936	
N	2602.000	2602.000
ll	-2241.090	-2225.742
bic	4513.635	4506.533
aic	4490.179	4465.485

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
NBRM	-0.017	1	0.003
ZINBRM	-0.009	1	0.002

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.701	0.704	0.002	0.020
1	0.181	0.198	0.017	3.641
2	0.064	0.060	0.003	0.525
3	0.022	0.021	0.001	0.155
4	0.006	0.008	0.002	1.405
5	0.004	0.004	0.000	0.004
6	0.001	0.002	0.001	0.821
7	0.001	0.001	0.000	0.247
8	0.000	0.001	0.001	1.717
9	0.000	0.000	0.000	0.009
Sum	0.981	0.999	0.028	8.544

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.701	0.705	0.004	0.057

```
1        0.181       0.190      0.009     0.998
2        0.064       0.066      0.002     0.232
3        0.022       0.023      0.001     0.210
4        0.006       0.009      0.003     1.896
5        0.004       0.003      0.000     0.096
6        0.001       0.002      0.000     0.235
7        0.001       0.001      0.000     0.008
8        0.000       0.000      0.000     0.948
9        0.000       0.000      0.000     0.478
------------------------------------------------
Sum      0.981       1.000      0.020     5.159

Tests and Fit Statistics
-------------------------------------------------------------------------
NBRM           BIC=-15948.586  AIC=     1.726  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-15955.688  dif=     7.102  ZINBRM    NBRM  Strong
               AIC=     1.716  dif=     0.009  ZINBRM    NBRM
               Vuong=   3.358  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table E- 26. Model fit comparison of NBRM & ZINBRM model 2 on burglary

Variable	NBRM	ZINBRM
bur07		
WBUR	4.598	3.109
	15.62	9.77
SIC58XX	1.065	1.118
	1.02	1.61
SIC55XX	1.150	1.015
	1.30	0.14
SIC54XX	1.396	1.319
	4.26	3.50
SIC75XX	0.938	0.843
	-0.77	-2.03
SIC73XX	1.091	1.285
	1.53	3.33
SIC72XX	1.096	1.040
	1.65	0.70
FINANCE	0.976	0.784
	-0.14	-1.39
Constant	0.178	0.269
	-24.46	-11.79
lnalpha		
Constant	0.851	0.621
	-1.34	-3.05
inflate		
WBUR		0.001
		-4.60
SIC58XX		1.361
		1.03
SIC55XX		0.000
		-0.02
SIC54XX		0.457
		-0.82
SIC75XX		0.063
		-1.93
SIC73XX		2.351

		2.44
SIC72XX		0.505
		-1.62
FINANCE		0.002
		-0.84
Constant		3.770
		3.26
Statistics		
alpha	0.851	
N	2602.000	2602.000
ll	-2222.822	-2198.927
bic	4524.284	4547.270
aic	4465.644	4435.854

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
NBRM	-0.018	1	0.003
ZINBRM	-0.013	1	0.002

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.701	0.703	0.002	0.010
1	0.181	0.200	0.018	4.404
2	0.064	0.060	0.004	0.681
3	0.022	0.020	0.002	0.294
4	0.006	0.008	0.002	1.204
5	0.004	0.004	0.000	0.014
6	0.001	0.002	0.001	0.787
7	0.001	0.001	0.000	0.242
8	0.000	0.001	0.001	1.725
9	0.000	0.000	0.000	0.011
Sum	0.981	0.999	0.029	9.373

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.701	0.704	0.002	0.019
1	0.181	0.194	0.013	2.172
2	0.064	0.065	0.001	0.051
3	0.022	0.022	0.000	0.018
4	0.006	0.008	0.002	1.417
5	0.004	0.003	0.000	0.161
6	0.001	0.002	0.000	0.247
7	0.001	0.001	0.000	0.000
8	0.000	0.000	0.000	1.073
9	0.000	0.000	0.000	0.236
Sum	0.981	1.000	0.020	5.393

Tests and Fit Statistics

NBRM	BIC=-15937.937	AIC= 1.716	Prefer	Over	Evidence
vs ZINBRM	BIC=-15914.950	dif= -22.986	NBRM	ZINBRM	Very strong

```
AIC=      1.705  dif=      0.011  ZINBRM    NBRM
Vuong=   4.100  prob=     0.000  ZINBRM    NBRM  p=0.000
```

Table E- 27. Model fit comparison of NBRM & ZINBRM model 3 on burglary

Variable	NBRM	ZINBRM
bur07		
WBUR	4.628	3.122
	15.70	9.66
BSTOPS	1.022	1.034
	1.60	1.22
SIC58XX	1.050	1.053
	0.79	0.75
SIC55XX	1.143	1.008
	1.25	0.07
SIC54XX	1.391	1.295
	4.23	3.46
SIC75XX	0.931	0.837
	-0.85	-2.10
SIC73XX	1.088	1.254
	1.48	3.03
SIC72XX	1.084	1.015
	1.45	0.27
Constant	0.175	0.268
	-24.48	-11.52
lnalpha		
Constant	0.846	0.612
	-1.38	-3.03
inflate		
WBUR		0.001
		-4.41
BSTOPS		1.060
		0.68
SIC58XX		1.078
		0.19
SIC55XX		0.000
		-0.02
SIC54XX		0.202
		-1.29
SIC75XX		0.120
		-1.79
SIC73XX		1.910
		2.46
SIC72XX		0.434
		-1.50
Constant		3.491
		3.06
Statistics		
alpha	0.846	
N	2602.000	2602.000
ll	-2221.615	-2199.907
bic	4521.870	4549.231
aic	4463.230	4437.814

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
NBRM	-0.018	1	0.003
ZINBRM	-0.012	1	0.002

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.701	0.703	0.002	0.010
1	0.181	0.200	0.018	4.422
2	0.064	0.060	0.004	0.679
3	0.022	0.020	0.002	0.292
4	0.006	0.008	0.002	1.206
5	0.004	0.004	0.000	0.014
6	0.001	0.002	0.001	0.782
7	0.001	0.001	0.000	0.238
8	0.000	0.001	0.001	1.719
9	0.000	0.000	0.000	0.010
Sum	0.981	0.999	0.029	9.373

ZINB: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.701	0.704	0.002	0.021
1	0.181	0.194	0.012	2.010
2	0.064	0.065	0.001	0.075
3	0.022	0.022	0.001	0.034
4	0.006	0.008	0.002	1.473
5	0.004	0.003	0.000	0.160
6	0.001	0.002	0.000	0.231
7	0.001	0.001	0.000	0.002
8	0.000	0.000	0.000	1.031
9	0.000	0.000	0.000	0.298
Sum	0.981	1.000	0.020	5.333

Tests and Fit Statistics

NBRM	BIC=-15940.350	AIC= 1.715	Prefer	Over	Evidence
vs ZINBRM	BIC=-15912.990	dif= -27.360	NBRM	ZINBRM	Very strong
	AIC= 1.706	dif= 0.010	ZINBRM	NBRM	
	Vuong= 4.130	prob= 0.000	ZINBRM	NBRM	p=0.000

Table E- 28. Model fit comparison of NBRM & ZINBRM model 4 on burglary

	Variable	NBRM	ZINBRM
bur07			
	WBUR	4.293	2.981
		14.68	8.10
	BSTOPS	1.028	1.055
		2.02	2.77
	SIC58XX	1.046	1.025
		0.73	0.36

SIC55XX	1.144	1.035
	1.25	0.28
SIC54XX	1.358	1.288
	3.96	3.23
SIC75XX	0.920	0.858
	-1.01	-1.63
SIC73XX	1.106	1.289
	1.78	3.47
SIC72XX	1.087	1.014
	1.50	0.23
MIXEDUSE	0.985	0.947
	-0.13	-0.41
PC1FVLAND	1.270	1.234
	3.00	2.19
PHOUSING	1.319	1.258
	2.60	1.88
UNIVCOLL	0.680	1.265
	-1.20	0.48
SCHOOLS	1.010	0.998
	0.09	-0.02
OPENSPACE	0.839	0.685
	-1.68	-3.22
Constant	0.154	0.239
	-21.30	-9.32
lnalpha		
Constant	0.803	0.579
	-1.76	-3.01
inflate		
WBUR		0.001
		-3.62
BSTOPS		1.150
		2.29
SIC58XX		0.755
		-0.56
SIC55XX		0.313
		-0.47
SIC54XX		0.391
		-0.74
SIC75XX		0.360
		-1.05
SIC73XX		1.952
		1.85
SIC72XX		0.446
		-1.96
MIXEDUSE		0.860
		-0.21
PC1FVLAND		0.700
		-0.79
PHOUSING		0.829
		-0.22
UNIVCOLL		3.146
		1.27
SCHOOLS		0.480
		-0.90
OPENSPACE		0.047
		-1.33
Constant		5.261
		3.02
Statistics		
alpha	0.803	

```
                             N |  2602.000    2602.000
                            ll | -2210.971   -2185.984
                           bic |  4547.766    4615.752
                           aic |  4453.941    4433.967
----------------------------------------------------------
                                             legend: b/t
Comparison of Mean Observed and Predicted Count
            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.019         1      0.003
ZINBRM       -0.013         1      0.002

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.701       0.703      0.001     0.006
1        0.181       0.200      0.019     4.672
2        0.064       0.060      0.004     0.676
3        0.022       0.020      0.002     0.292
4        0.006       0.008      0.002     1.202
5        0.004       0.004      0.000     0.017
6        0.001       0.002      0.001     0.749
7        0.001       0.001      0.000     0.212
8        0.000       0.001      0.001     1.670
9        0.000       0.000      0.000     0.004
------------------------------------------------
Sum      0.981       0.999      0.029     9.500

ZINB: Predicted and actual probabilities
Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.701       0.704      0.002     0.019
1        0.181       0.194      0.013     2.233
2        0.064       0.065      0.001     0.023
3        0.022       0.022      0.000     0.011
4        0.006       0.008      0.002     1.467
5        0.004       0.003      0.000     0.119
6        0.001       0.002      0.000     0.302
7        0.001       0.001      0.000     0.002
8        0.000       0.000      0.000     1.121
9        0.000       0.000      0.000     0.193
------------------------------------------------
Sum      0.981       1.000      0.020     5.489

Tests and Fit Statistics
--------------------------------------------------------------------------
NBRM           BIC=-15914.455  AIC=     1.712  Prefer  Over  Evidence
--------------------------------------------------------------------------
  vs ZINBRM      BIC=-15846.469  dif=   -67.986  NBRM    ZINBRM  Very strong
               AIC=     1.704  dif=     0.008  ZINBRM    NBRM
               Vuong=   4.244  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table E- 29. Fit comparisons of Model 1 and Model 2 on burglary

Measures of Fit for nbreg of bur07, Current = M2, Saved = M1

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-2369.888	-2369.888	0.000
Log-Lik Full Model	-2222.822	-2241.090	18.268
D	4445.644(2592)	4482.179(2598)	36.535(6)
LR	294.131(8)	257.596(2)	36.535(6)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.062	0.054	0.008
McFadden's Adj R2	0.058	0.053	0.005
ML (Cox-Snell) R2	0.107	0.094	0.013
Cragg-Uhler(Nagelkerke) R2	0.128	0.112	0.015
AIC	1.716	1.726	-0.009
AIC*n	4465.644	4490.179	-24.535
BIC	-15937.937	-15948.586	10.649
BIC'	-231.219	-241.868	10.649
BIC used by Stata	4524.284	4513.635	10.649
AIC used by Stata	4465.644	4490.179	-24.535

Difference of 10.649 in BIC' provides very strong support for saved model.

Note: p-value for difference in LR is only valid if models are nested.

Table E- 30. Fit comparisons of Model 1 and Model 3 on burglary

Measures of Fit for nbreg of bur07, Current = M3, Saved = M1

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-2369.888	-2369.888	0.000
Log-Lik Full Model	-2221.594	-2241.090	19.496
D	4443.188(2591)	4482.179(2598)	38.991(7)
LR	296.587(9)	257.596(2)	38.991(7)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.063	0.054	0.008
McFadden's Adj R2	0.058	0.053	0.005
ML (Cox-Snell) R2	0.108	0.094	0.013
Cragg-Uhler(Nagelkerke) R2	0.129	0.112	0.016
AIC	1.716	1.726	-0.010
AIC*n	4465.188	4490.179	-24.991
BIC	-15932.529	-15948.586	16.057
BIC'	-225.811	-241.868	16.057
BIC used by Stata	4529.692	4513.635	16.057
AIC used by Stata	4465.188	4490.179	-24.991

Difference of 16.057 in BIC' provides very strong support for saved model.

Note: p-value for difference in LR is only valid if models are nested.

Table E- 31. Fit comparisons of Model 1 and Model 4 on burglary

Measures of Fit for nbreg of bur07, Current = M4, Saved = M1

	Current	Saved	Difference
Model:	nbreg	nbreg	
N:	2602	2602	0
Log-Lik Intercept Only	-2369.888	-2369.888	0.000
Log-Lik Full Model	-2210.965	-2241.090	30.125
D	4421.930(2585)	4482.179(2598)	60.249(13)
LR	317.846(15)	257.596(2)	60.249(13)
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.067	0.054	0.013
McFadden's Adj R2	0.060	0.053	0.007
ML (Cox-Snell) R2	0.115	0.094	0.021
Cragg-Uhler(Nagelkerke) R2	0.137	0.112	0.025
AIC	1.713	1.726	-0.013
AIC*n	4455.930	4490.179	-34.249
BIC	-15906.603	-15948.586	41.983
BIC'	-199.885	-241.868	41.983
BIC used by Stata	4555.618	4513.635	41.983
AIC used by Stata	4455.930	4490.179	-34.249

Difference of 41.983 in BIC' provides very strong support for saved model.
Note: p-value for difference in LR is only valid if models are nested.

Table E- 32. Ad hoc Model PRM on burglary

Poisson regression — Number of obs = 2602; LR chi2(5) = 484.39; Prob > chi2 = 0.0000; Pseudo R2 = 0.0955

Log likelihood = -2293.8158

bur07	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wbur	1.295613	.0650811	19.91	0.000	1.168057	1.42317
bstops	.0271598	.0097398	2.79	0.005	.0080701	.0462494
sic54xx	.3958609	.04786	8.27	0.000	.302057	.4896648
pc1fvland	.2524674	.0674761	3.74	0.000	.1202168	.3847181
phousing	.2544981	.0843908	3.02	0.003	.0890951	.4199011
_cons	-1.786771	.067316	-26.54	0.000	-1.918708	-1.654834

Table E- 33. Ad hoc Model NBRM on burglary

Negative binomial regression — Number of obs = 2602; LR chi2(5) = 303.40; Prob > chi2 = 0.0000; Pseudo R2 = 0.0640

Dispersion = mean
Log likelihood = -2218.1869

bur07	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wbur	1.443092	.0977582	14.76	0.000	1.25149	1.634695
bstops	.0312077	.0131263	2.38	0.017	.0054807	.0569348
sic54xx	.3919845	.0722568	5.42	0.000	.2503638	.5336051
pc1fvland	.2393952	.0788996	3.03	0.002	.0847548	.3940356
phousing	.2578064	.1065159	2.42	0.016	.049039	.4665738
_cons	-1.86384	.0825027	-22.59	0.000	-2.025542	-1.702137

```
-------------+----------------------------------------------------------------
    /lnalpha |  -.1912196   .1234256                     -.4331293    .0506902
-------------+----------------------------------------------------------------
       alpha |   .8259512   .1019435                      .6484766    1.051997
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =  151.26 Prob>=chibar2 = 0.000
------------------------------------------------------------------------------
Log-Lik Intercept Only:      -2369.888   Log-Lik Full Model:           -2218.187
D(2595):                      4436.374   LR(5):                          303.401
                                         Prob > LR:                        0.000
McFadden's R2:                   0.064   McFadden's Adj R2:                0.061
ML (Cox-Snell) R2:               0.110   Cragg-Uhler(Nagelkerke) R2:       0.131
AIC:                             1.710   AIC*n:                         4450.374
BIC:                        -15970.799   BIC':                          -264.081
BIC used by Stata:            4491.422   AIC used by Stata:             4450.374
```

Table E- 34. Ad hoc Model ZINBRM model on burglary

```
Zero-inflated negative binomial regression        Number of obs   =       2602
                                                  Nonzero obs     =        777
                                                  Zero obs        =       1825
Inflation model = logit                           LR chi2(5)      =     117.28
Log likelihood  = -2203.455                       Prob > chi2     =     0.0000
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
bur07        |
        wbur |   .9840571   .1301675     7.56   0.000     .7289336    1.239181
      bstops |   .0572152   .0190325     3.01   0.003     .0199121    .0945182
     sic54xx |   .3446197   .0793085     4.35   0.000     .1891778    .5000616
   pc1fvland |   .2598956   .1077796     2.41   0.016     .0486515    .4711397
     phousing |   .1814364   .1300536     1.40   0.163     -.073464    .4363368
       _cons |  -1.398782   .1504429    -9.30   0.000    -1.693645    -1.10392
-------------+----------------------------------------------------------------
inflate      |
        wbur |  -5.322596   1.344772    -3.96   0.000      -7.9583   -2.686891
      bstops |   .0585773   .0368791     1.59   0.112    -.0137043     .130859
     sic54xx |   -.276905   .4731768    -0.59   0.558    -1.204315    .6505045
   pc1fvland |   .0624422   .4251246     0.15   0.883    -.7707868    .8956711
    phousing |   -.264359   .7211583    -0.37   0.714    -1.677803    1.149085
       _cons |   .6929249   .4420674     1.57   0.117    -.1735114    1.559361
-------------+----------------------------------------------------------------
    /lnalpha |  -.6161378   .2013862    -3.06   0.002    -1.010847   -.2214281
-------------+----------------------------------------------------------------
       alpha |   .5400261   .1087538                      .3639105    .8013735
------------------------------------------------------------------------------
Measures of Fit for ZINBRM of bur07
------------------------------------------------------------------------------
Log-Lik Intercept Only:      -2369.888   Log-Lik Full Model:           -2203.455
D(2589):                      4406.909   LR(10):                         332.866
                                         Prob > LR:                        0.000
McFadden's R2:                   0.070   McFadden's Adj R2:                0.065
ML (Cox-Snell) R2:               0.120   Cragg-Uhler(Nagelkerke) R2:       0.143
AIC:                             1.704   AIC*n:                         4432.909
BIC:                        -15953.079   BIC':                          -254.225
BIC used by Stata:            4509.142   AIC used by Stata:             4432.909
```

Table E- 35. Fit statistics comparison of ad hoc model between PRM and NBRM

Variable	PRM	NBRM
bur07		
WBUR	3.653	4.234
	19.91	14.76
BSTOPS	1.028	1.032
	2.79	2.38
SIC54XX	1.486	1.480
	8.27	5.42
PC1FVLAND	1.287	1.270
	3.74	3.03
PHOUSING	1.290	1.294
	3.02	2.42
Constant	0.168	0.155
	-26.54	-22.59
lnalpha		
Constant		0.826
		-1.55
Statistics		
alpha		0.826
N	2602.000	2602.000
ll	-2293.816	-2218.187
bic	4634.816	4491.422
aic	4599.632	4450.374

legend: b/t

Comparison of Mean Observed and Predicted Count

Model	Maximum Difference	At Value	Mean |Diff|
PRM	-0.074	1	0.013
NBRM	-0.019	1	0.003

PRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.701	0.658	0.044	7.526
1	0.181	0.255	0.074	55.789
2	0.064	0.064	0.000	0.003
3	0.022	0.016	0.006	6.769
4	0.006	0.004	0.002	1.621
5	0.004	0.002	0.002	8.168
6	0.001	0.001	0.000	0.942
7	0.001	0.000	0.000	2.054
8	0.000	0.000	0.000	0.328
9	0.000	0.000	0.000	5.477
Sum	0.981	1.000	0.130	88.676

NBRM: Predicted and actual probabilities

Count	Actual	Predicted	|Diff|	Pearson
0	0.701	0.703	0.001	0.005
1	0.181	0.200	0.019	4.669
2	0.064	0.060	0.004	0.647
3	0.022	0.020	0.002	0.294
4	0.006	0.008	0.002	1.175

```
5        0.004       0.004      0.000     0.021
6        0.001       0.002      0.001     0.737
7        0.001       0.001      0.000     0.210
8        0.000       0.001      0.001     1.673
9        0.000       0.000      0.000     0.005
------------------------------------------------
Sum      0.981       0.999      0.029     9.436

Tests and Fit Statistics

PRM            BIC=-15827.405  AIC=     1.768  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM      BIC=-15970.799  dif=   143.394  NBRM    PRM   Very strong
               AIC=     1.710  dif=     0.057  NBRM    PRM
               LRX2=  151.258  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
NBRM           BIC=-15970.799  AIC=     1.710  Prefer  Over  Evidence
```

Table E- 36. Fit statistics comparison of ad hoc model between NBRM and ZIBN

```
----------------------------------------------------------
                        Variable |   NBRM       ZINBRM
---------------------------------+------------------------
bur07                            |
                            WBUR |    4.234       2.675
                                 |    14.76        7.56
                          BSTOPS |    1.032       1.059
                                 |     2.38        3.01
                         SIC54XX |    1.480       1.411
                                 |     5.42        4.35
                       PC1FVLAND |    1.270       1.297
                                 |     3.03        2.41
                        PHOUSING |    1.294       1.199
                                 |     2.42        1.40
                        Constant |    0.155       0.247
                                 |   -22.59       -9.30
---------------------------------+------------------------
lnalpha                          |
                        Constant |    0.826       0.540
                                 |    -1.55       -3.06
---------------------------------+------------------------
inflate                          |
                            WBUR |                0.005
                                 |                -3.96
                          BSTOPS |                1.060
                                 |                 1.59
                         SIC54XX |                0.758
                                 |                -0.59
                       PC1FVLAND |                1.064
                                 |                 0.15
                        PHOUSING |                0.768
                                 |                -0.37
                        Constant |                2.000
                                 |                 1.57
---------------------------------+------------------------
Statistics                       |
                           alpha |    0.826
                               N | 2602.000    2602.000
                              ll | -2218.187   -2203.455
                             bic | 4491.422    4509.142
                             aic | 4450.374    4432.909
```

```
----------------------------------------------------------
                                               legend: b/t
Comparison of Mean Observed and Predicted Count

            Maximum       At      Mean
Model     Difference    Value    |Diff|
---------------------------------------------
NBRM       -0.019         1      0.003
ZINBRM       -0.010         1      0.002

NBRM: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.701       0.703      0.001     0.005
1        0.181       0.200      0.019     4.669
2        0.064       0.060      0.004     0.647
3        0.022       0.020      0.002     0.294
4        0.006       0.008      0.002     1.175
5        0.004       0.004      0.000     0.021
6        0.001       0.002      0.001     0.737
7        0.001       0.001      0.000     0.210
8        0.000       0.001      0.001     1.673
9        0.000       0.000      0.000     0.005
------------------------------------------------
Sum      0.981       0.999      0.029     9.436

ZINB: Predicted and actual probabilities

Count   Actual    Predicted    |Diff|   Pearson
------------------------------------------------
0        0.701       0.704      0.003     0.036
1        0.181       0.192      0.010     1.418
2        0.064       0.066      0.002     0.200
3        0.022       0.023      0.001     0.112
4        0.006       0.008      0.002     1.639
5        0.004       0.003      0.000     0.149
6        0.001       0.001      0.000     0.206
7        0.001       0.001      0.000     0.010
8        0.000       0.000      0.000     0.957
9        0.000       0.000      0.000     0.435
------------------------------------------------
Sum      0.981       1.000      0.020     5.162

Tests and Fit Statistics
-------------------------------------------------------------------------
NBRM           BIC=-15970.799  AIC=     1.710  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINBRM      BIC=-15953.079  dif=   -17.720  NBRM    ZINBRM  Very strong
               AIC=     1.704  dif=     0.007  ZINBRM    NBRM
               Vuong=   3.394  prob=    0.000  ZINBRM    NBRM  p=0.000
```

Table E- 37. Ad hoc NBRM Model showing discrete changes by predictors

```
 Observed SD: .90081004
----------------------------------------------------------------------
       bur07 |      b         z     P>|z|       %      %StdX      SDofX
-------------+--------------------------------------------------------
        wbur |   1.44309   14.762   0.000    323.4      59.1     0.3217
      bstops |   0.03121    2.377   0.017      3.2       8.4     2.5823
```

```
     sic54xx |   0.39198    5.425  0.000     48.0     16.7     0.3935
   pc1fvland |   0.23940    3.034  0.002     27.0     12.2     0.4811
    phousing |   0.25781    2.420  0.016     29.4      8.2     0.3050
-------------+-------------------------------------------------------
    ln alpha |  -0.19122
       alpha |   0.82595   SE(alpha) = 0.10194
---------------------------------------------------------------------
 LR test of alpha=0: 151.26   Prob>=LRX2 = 0.000
---------------------------------------------------------------------
       b = raw coefficient
       z = z-score for test of b=0
   P>|z| = p-value for z-test
       % = percent change in expected count for unit increase in X
   %StdX = percent change in expected count for SD increase in X
   SDofX = standard deviation of X
```

Table E- 38. Ad hoc ZINBRM Model showing discrete changes by predictors

```
ZINBRM (N=2602): Percentage Change in Expected Count

Observed SD: .90081004

Count Equation: Percentage Change in Expected Count for Those Not Always 0
---------------------------------------------------------------------
       bur07 |      b         z     P>|z|      %      %StdX      SDofX
-------------+-------------------------------------------------------
        wbur |   0.98406    7.560  0.000    167.5     37.2     0.3217
      bstops |   0.05722    3.006  0.003      5.9     15.9     2.5823
     sic54xx |   0.34462    4.345  0.000     41.1     14.5     0.3935
   pc1fvland |   0.25990    2.411  0.016     29.7     13.3     0.4811
    phousing |   0.18144    1.395  0.163     19.9      5.7     0.3050
-------------+-------------------------------------------------------
    ln alpha |  -0.61614
       alpha |   0.54003   SE(alpha) = 0.10875
---------------------------------------------------------------------
Binary Equation: Factor Change in Odds of Always 0
---------------------------------------------------------------------
     Always0 |      b         z     P>|z|      %      %StdX      SDofX
-------------+-------------------------------------------------------
        wbur |  -5.32260   -3.958  0.000    -99.5    -82.0     0.3217
      bstops |   0.05858    1.588  0.112      6.0     16.3     2.5823
     sic54xx |  -0.27691   -0.585  0.558    -24.2    -10.3     0.3935
   pc1fvland |   0.06244    0.147  0.883      6.4      3.0     0.4811
    phousing |  -0.26436   -0.367  0.714    -23.2     -7.7     0.3050
---------------------------------------------------------------------
       b = raw coefficient
       z = z-score for test of b=0
   P>|z| = p-value for z-test
       % = percent change in odds for unit increase in X
   %StdX = percent change in odds for SD increase in X
   SDofX = standard deviation of X
```

Sung-suk Violet Yu

Place of Birth
Seoul, South Korea

Education	**Dates attended**
Visual Media High School, Seoul, South Korea	03/1986—02/1989
Simon Fraser University, Burnaby, BC, Canada - BA in Criminology	09/1999—08/2002
Simon Fraser University, Burnaby, BC, Canada - MA in Criminology -	09/2002—08/2004
Rutgers University, Newark, NJ, USA - MA in Criminal Justice -	09/2004—05/2006
Rutgers University, Newark, NJ, USA - PhD in Criminal Justice	09/2006—05/2009

Employment	**Dates**
Teaching Assistant, Simon Fraser University, Burnaby, BC, Canada	09/2002—12/2003
Instructor, Rutgers University, Newark, NJ, USA	09/2005—07/2008
Research Assistant, Rutgers University, Newark, NJ, USA	06/2005—07/2006
Research Associate, John Jay College of Criminal Justice, New York, NY, USA	10/2006—03/2007
Senior Research Associate, Vera Institute of Justice, New York, NY	03/2009—to present

Publication

Pizarro, Jesenia M.; Corsaro, Nicholas and **Sung-suk Violet Yu** (2007). “Journey to Crime and Victimization: An Application of Routine Activities Theory and Environmental Criminology to Homicide” in Victims & Offenders. Vol. 2(4), pp375-394

Yu, Sung-suk Violet (2008). Review of: Policing Public Transportation: An environmental and procedural evaluation of bus stops, Kooi, Brandon R. (2007). LFB Scholarly Publishing LLC. NY. 188 pages. In Journal of Applied Security Research. Vol. 3 (3)

Mellow, Jeff; Hoge, Steven K.; Lee, Joshua D.; Natarajan, Mangai; **Yu, Sung-suk Violet**; Greifinger, Robert B. and Gary Belkin. (2008) “Mapping the Innovation in Correctional Health Care Service Delivery in New York City”. New York University School of Medicine, Bellevue Hospital Center and Criminal Justice Research and Evaluation Center, John Jay College of Criminal Justice