Spatial Crime Forecasting: Application of Risk Terrain Modeling in a Metropolitan County

Audrey Clubb

Follow this and additional works at: https://scholarworks.gsu.edu/cj_diss

Recommended Citation
doi: https://doi.org/10.57709/10069155
ABSTRACT

SPATIAL CRIME FORECASTING: APPLICATION OF RISK TERRAIN MODELING IN A METROPOLITAN COUNTY

BY

AUDREY C. CLUDD

May 2017

Committee Chair: Dr. Joshua Hinkle

Major Department: Criminal Justice and Criminology

The present study utilizes a novel approach to spatial crime analysis known as Risk Terrain Modeling (RTM) to assess the applicability of this technique in the prediction of predatory violent crime (homicide, aggravated assault, and pedestrian robbery). RTM provides a methodology for analyzing potential correlates of crime in a spatial context to identify specific locations at the highest risk for crime in the future. The intent of this process is to allow police and other community agencies to identify places at the highest risk for crime to allocate additional resources to stave off potential crime problems. The present study applies this technique to a metropolitan county near Atlanta, Georgia. RTM models were generated for each year from 2010 through 2014 based on several area characteristics including physical and social disorder, criminal elements, risky places, socioeconomic conditions, and area economic health. These models are then tested for their predictive validity and discrimination capabilities. The results indicate that the RTM methodology was successful in identifying many areas of risk for future crime, but was limited in the ability to accurately pinpoint areas at the highest risk of crime. The explained variance in the distribution of crime relative to risk was minimal and the
RTM models incorrectly identified many places as high risk that were not subject to future crime while not identifying many areas that did experience crime. While this methodology shows promise on principle, more research is needed to improve performance of the RTM model prior to its use as a potential tool for crime prediction and intervention.
SPATIAL CRIME FORECASTING: APPLICATION OF RISK TERRAIN MODELING IN A METROPOLITAN COUNTY

BY

AUDREY C. CLUBB

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Andrew Young School of Policy Studies of Georgia State University

GEORGIA STATE UNIVERSITY
2017
Acceptance

This dissertation was prepared under the direction of the candidate’s Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Criminal Justice and Criminology in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair:  Dr. Joshua Hinkle
Committee:  Dr. Dean Dabney
            Dr. Timothy Brezina
            Dr. Elizabeth Groff

Electronic Version Approved:

Mary Beth Walker, Dean
Andrew Young School of Policy Studies
Georgia State University
May, 2017
Dedication

I dedicate this dissertation to my amazing family for all their support through this process. To my mom, who knew that sometimes the correct answer to frustrating findings is a hug, a baby cheesecake, and scary movie. To my dad, who, throughout my entire life, never let me believe for a moment that there was something I could not do. To my husband, who had a truly endless amount of patience and support throughout this entire process and who was always there with a video game and a glass of scotch to keep me going. And to my grandmom, who I will never surprise, but will forever strive to amaze. I love you all and cannot thank you enough.
Acknowledgements

I would like to acknowledge all the members of my dissertation committee and the outstanding professors I have had over the years. My sincere thanks go to Dr. Joshua Hinkle, Dr. Dean Dabney, Dr. Timothy Brezina, and Dr. Elizabeth Groff for their patience and guidance in the development of this dissertation. I would also like to extend my thanks to Dr. Joel Caplan and Dr. Leslie Kennedy at Rutgers University for their ongoing training and support in implementing this promising methodology.

I sincerely thank Dr. Joshua Hinkle for serving as my mentor and dissertation chair. He is a brilliant scholar from who I have learned so much. His course on policing innovation set me on a path to a career in the research and evaluation of policing practice, a critically important topic, with an optimistic and objective eye towards improvement. I cannot thank him enough for his endless patience and guidance throughout the dissertation process as well as the other projects we have worked on together. Dr. Hinkle’s positive nature and encouragement were inspiring.

I also want to acknowledge the support of Dr. Dean Dabney who facilitated the data collection process and development of this dissertation. Dr. Dabney was central to building relationships with community agencies to access the data needed for this dissertation, obtaining funding for research, and working through the scope. He has been an excellent mentor and has made me a much stronger researcher.

Finally, I must again thank my husband, Richard Clubb, for his encouragement, patience, and support throughout my graduate career. I am so lucky to have married such a truly amazing person.
# Table of Contents

Dedication ................................................................................................................................. iv

Acknowledgements .................................................................................................................... v

List of Tables ............................................................................................................................... x

List of Figures ............................................................................................................................. xi

Chapter I – Introduction ........................................................................................................... 1

Chapter II – Literature Review ................................................................................................ 6

  Intelligence-Led Policing and the Development of Risk Terrain Modeling. ......................... 6
    Importance of data, information, knowledge, and intelligence ............................................. 7
    Early policing and the growing importance of information. ................................................. 8
    Rapid developments in information calls for policing innovations. ...................................... 10
    Policing Innovation and the Use of Information and Intelligence ........................................ 12
    Summary. ................................................................................................................................. 18

Defining Risk Terrain Modeling ............................................................................................... 19

  Comparison with existing crime modeling techniques. ......................................................... 21

Risk Terrain Modeling Methodology ......................................................................................... 23

Theoretical Foundations of Risk Terrain Modeling ................................................................. 26

  Theory of Risky Places ............................................................................................................ 27
  Broken Windows Theory ......................................................................................................... 28
  Rational Choice Theory, Routine Activities Theory, and Situational Crime Prevention ...... 31
  Social Disorganization Theory ............................................................................................... 32
  Theoretical foundations summary .......................................................................................... 33
Analysis Plan......................................................................................................................................................... 77

Step 1: Identify an outcome event (dependent variable) of interest.................................................. 77

Step 2: Identify area of study..................................................................................................................... 78

Step 3: Identify a time period for study.................................................................................................... 78

Step 4: Identify risk factors (independent variables) related to outcome of interest. ........ 78

Step 5: Obtain data and maps. .................................................................................................................. 78

Step 6: Generate maps for each risk factor and outcome measure................................................... 78

Step 7: Refine model to omit non-significant measures ............................................................... 78

Step 8: Apply variable weighting, as appropriate................................................................................ 79

Step 9: Combine individual risk maps into risk terrain map............................................................. 80

Step 10: Communicate model findings.................................................................................................. 80

Step 11: Examine predictive validity of the RTM Models................................................................. 81

Chapter IV – Results................................................................................................................................. 85

Variable Selection and Weighting........................................................................................................... 85

Descriptive Statistics................................................................................................................................. 92

RTM Model Comparison.............................................................................................................................. 94

Model Evaluation........................................................................................................................................ 100

Model Discrimination................................................................................................................................ 105

Summary................................................................................................................................................... 107

Chapter V – Conclusions .............................................................................................................................. 108

Summary and Discussion of Findings........................................................................................................ 108

Variable selection.................................................................................................................................... 108
List of Tables

Table 3.1 – Demographic Characteristics of Unincorporated DeKalb County GA (2010) ........... 51
Table 3.2 – Predatory Violent Crime .......................................................................................... 54
Table 3.3 – Independent Variables .......................................................................................... 57
Table 4.1 – Bivariate Correlations .......................................................................................... 86
Table 4.2 – Bivariate Correlations with Predatory Violent Crime ............................................. 89
Table 4.3 – Variable Weighting ................................................................................................ 91
Table 4.4 – Variable Descriptive Summary ............................................................................. 93
Table 4.5 – Percent Raster Grid Identified as High-Risk ........................................................... 94
Table 4.6 – Comparison of High-Risk Cell Identification (Percentage) ...................................... 99
Table 4.7 – Logistic Regression Analysis for High-Risk RTM Models .................................... 103
Table 4.8 – AUC Values for RTM High-Risk Models .............................................................. 106
List of Figures

Figure 3.1 – Unincorporated DeKalb County Map................................................................. 50

Figure 4.1 – Risk Terrain Modeling Heat Maps....................................................................... 96

Figure 4.2 – High-Risk Areas ................................................................................................. 98

Figure 4.3 – High-Risk Areas with PVC Overlay ................................................................. 101
Chapter I – Introduction

Increased demands on the criminal justice system, budgetary limitations related to the recent economic recession in the United States, and increased emphasis on evidence-based practice drive the need for police departments to continuously seek more efficient and effective means to address crime problems. While the economic recession has begun to subside, a 2011 report from The Office of Community Oriented Policing Services suggests that police departments have seen decreases in staff and funding that may not recover to 2009 peak resource levels. A means to address these limitations is the use of crime mapping and analysis strategies designed to identify places where crime is likely to occur. These techniques, often deemed “predictive policing” or “intelligence-led policing,” assist police in strategic planning to ensure they are using their limited resources to effectively address the individuals, groups, and places that pose the greatest threat of crime.

Recent and rapid technological advancements in geospatial crime mapping and statistical analysis have the potential to revolutionize our understanding of and response to crime. These innovations take advantage of advanced computing capabilities and data analytics to identify geographic and temporal crime patterns. To date, these techniques focus identifying places where crime clusters using retrospective crime data. Such techniques include hot spot analysis, near repeat analysis, PredPol, and Compstat (see Braga, Papachristos, & Hureau, 2014; Huet, 2015; Short, D’Orsogna, Brantingham, & Tita, 2009; Weisburd, Mastrofski, McNally, Greenspan, & Willis, 2003). Automated techniques and computer software allow police departments to quickly assess crime patterns and allocate resources in those areas where crime concentrates.
While these techniques rely on prior crime patterns to assess risk of future crime, more recent techniques seek to include known correlates of crime to identify places where crime is likely to occur in the future even if it is not already a problem. These techniques have the potential to improve upon traditional policing practices by identifying risk factors associated with persistent crime risk. This process may allow police and other community organizations to address the underlying issues associated with crime in certain areas (e.g., vacant buildings, low-level crime, economic disadvantage) to have a more lasting effect on crime reduction. One of the most recent and promising developments in this field is Risk Terrain Modeling (RTM).

RTM utilizes geospatial and statistical modeling to evaluate potential risk and protective factors and to identify micro-places at the highest risk of future criminal events. This process draws upon suitability modeling techniques and the concept of the “environmental backcloth” of crime (Brantingham & Brantingham, 1981) to take a predictive approach to crime analysis based on the spatial context in which crime occurs. The RTM process is used to generate “heat maps” to direct police to those areas where underlying factors associated with crime generate an environment in which the risk of crime is higher than the surrounding areas (Caplan & Kennedy, 2011). If successfully conducted and implemented, RTM can be used both to direct police to high-crime areas and to direct community organizations to areas where early intervention to address underlying risk factors may stave off more serious crime issues. While RTM offers an intriguing and novel approach to predictive crime analysis, it is a relatively new technique requiring additional research. The present study seeks to build our understanding of the applicability of RTM modeling and its predictive validity.
Prior RTM evaluations have demonstrated a significant correlation between risk predictions developed in the RTM process and crime outcomes (see Daley, et al., 2016; Drawve, 2016; Drawve, Thomas, & Walker, 2016; Caplan, Kennedy, Barnum, & Piza, 2015; Caplan, Kennedy, & Miller, 2011; Caplan, Kennedy, & Piza, 2013; Kennedy, Caplan, & Piza, 2011; Kennedy, Caplan, Piza, & Buccine-Schrader, 2016; Moreto, Piza, & Caplan, 2014; Piza, Feng, Kennedy, & Caplan, 2016). However, given the broad range of variables considered, crime types assessed, and methodologies used, it is difficult to conduct a cross-study review to identify consistency of findings. This is necessary to evaluate the efficacy of RTM in crime prediction and intervention. In addition, prior evaluations of RTM research have included limited discussion of the extent of the predictive validity of these models. The present study seeks to improve understanding of RTM applicability by providing a thorough examination.

The present study applies the RTM methodology outlined by Caplan and Kennedy (2010) to assess risk of predatory violent crime (homicide, aggravated assault, and pedestrian robbery) in Unincorporated DeKalb County, Georgia. RTM models were generated for each year from 2010 through 2014 based on several area characteristics including physical and social disorder, criminal elements, risky places, socioeconomic conditions, and area economic health. Data for these measurements were obtained from several agencies within DeKalb County, the United States Census Bureau, and the DeKalb County Police Department. The predictive validity and discrimination capabilities of these models are then tested utilizing crime data from 2011 through 2015.

Chapter II provides a literature review tracing the historical use of information and intelligence in policing, the primary concepts and techniques of RTM, and a summary of RTM
models in the current empirical literature. Tracing the evolution of the use of information and intelligence in policing, including modern policing innovations, draws attention to the importance of these elements in increasing the effectiveness of police responses to crime. The discussion of RTM briefly summarizes its history, intent, and methodology to provide background on the novel technique for the reader. Finally, while there are few peer-reviewed evaluations of RTM methodology to date, a discussion of major findings and limitations is provided to lay the foundation for the present study.

Chapter III details the methodology used for the present study including the variables included as risk factors and the application of RTM techniques to assess crime in Unincorporated DeKalb County, Georgia. RTM is a complex process used to map potential risk factors, identify significant risk factors, build a weighted RTM model, identify high-risk areas, and determine implications of findings. The present study includes a process for model evaluation not utilized in previous studies. A thorough discussion of this process is also provided as it applies to the present study.

Chapter IV presents the results of each phase of the present study. The purpose of the present study is to demonstrate if RTM methodology is applicable to the context of Unincorporated DeKalb County, to examine the relationship between risk predictions and future crime outcomes, and to evaluate the diagnostic capabilities of the RTM technique. These issues are explored and findings are detailed. The results section includes a discussion of variable selection and weighting used to generate the RTM models, a visual comparison of the models generated for each year (2010 through 2014), a comparison of models and crime outcomes in subsequent years utilizing logistic regression, examinations of Nagelkerke’s $R^2$ to
explore model performance, and receiver operating characteristic analysis to examine
diagnostic capabilities of RTM models.

Chapter V summarizes the findings of the present study, identifies limitations, proposes
directions for future research, and discusses policy implications. Because RTM is nascent
science and because of challenges encountered in the present study, a strong emphasis is
placed on areas of research needed to improve model performance. Important steps are
needed to improve theoretical and methodological guidance to increase the utility of the RTM
technique prior to its application as a predictive analytic tool for practitioners.
Chapter II – Literature Review

Risk Terrain Modeling (RTM) is a novel, data-driven approach to assessing spatial crime risk that builds upon principles of several criminological theories, intelligence-led policing, and recent advancements in computing technology. Because of its novelty and continued development, there are few empirical studies assessing its validity and limitations. However, much can be gleaned from prior research into the theoretical underpinnings of RTM and the history of intelligence use in policing in order to evaluate the utility of RTM and its potential for improvement. This literature review examines intelligence-led policing and policing innovation, the process of RTM, the theoretical foundations of RTM, findings from RTM research to date, and correlates of crime as they apply to the present study.

Intelligence-Led Policing and the Development of Risk Terrain Modeling.

Since the inception of law enforcement and policing, officers have been responsible for utilizing various sources of information to investigate specific crimes and to become more effective in addressing crime problems in their communities. As societies grew and became more complex, law enforcement officers and agencies have developed more advanced techniques for collecting information about crime and coalescing that information to inform strategy and practice. As such, the use of information by police has evolved significantly from the earliest days of policing. RTM is one of the most recent innovations in the use of predictive policing, a form of intelligence-led policing, which uses a variety of factors related to crime to estimate where crime is most likely to occur in the future. This forward-looking approach allows police and other government agencies to more effectively allocate resources to better mitigate crime problems. It is important to recognize the developments that have led to the predictive

6
approach that is growing in importance today. The following is a brief outline of the key advancements in the use of information and intelligence in policing.

**Importance of data, information, knowledge, and intelligence.** Before discussing specific innovations, it is first important to understand the difference in the levels of information available to police. Modern policing initiatives are rife with buzzwords intended to demonstrate the value of such programs. These initiatives include descriptors such as “intelligence-led,” “data-driven,” and “smart initiatives.” These are particularly prevalent in light of the Homeland Security response to the terrorist attacks of September 11, 2001. However, the extent to which information informs these initiatives can vary substantially. Many scholars argue that police departments are using such terms interchangeably by claiming to use more advanced intelligence-led policing practices when little analysis is actually taking place (James, 2013; Ratcliffe, 2002; Ratcliffe, 2011).

Ratcliffe (2011) outlines important distinctions between data, information, knowledge, and intelligence. Each of these exists on a continuum with data representing the lowest level and intelligence representing the highest. Data refers to very general observations and measurements such as crime statistics or facts about individual criminals of crime events (Ratcliffe, 2011). Ratcliffe’s definition of information is more nebulous, but represents data that has been contextualized but not interpreted. Knowledge represents information that has been interpreted and given meaning that is directly relevant for police use (Ratcliffe, 2011). Finally, intelligence goes a step beyond knowledge, indicating data analysis that has resulted into specific recommendations for police response (Ratcliffe, 2011). These delineations may be too
specific for practical assessments of policing, but the concept is essential: Police have the potential to improve their effectiveness through the use of data, but the value of that data is determined by the extent to which it is contextualized, assessed, and utilized in the implementation of policy and practice. Intelligence-led policing practices, including RTM, should involve objective analysis of data related to crime and its correlates to inform police policies and practices.

The history of policing has largely followed this continuum in its development, with modern practices beginning to incorporate crime intelligence. This evolution has been driven by a number of factors, particularly advancements in technology, increasing globalization, and increasing societal complexities. While officers in the earliest days of policing relied on reports of specific crimes and their own intuition, indicative of data and information, modern police departments have sought to incorporate systematic analysis of crime statistics and targeted response, indicative of knowledge and intelligence. RTM is one of the most recent attempts to utilize intelligence-led practices to improve policing.

**Early policing and the growing importance of information.** Prior to the 1600s, information on crime and criminals had an important but unofficial role in policing. Under the Frankpledge system, local residents were expected to share information on known crimes and criminals, but there was no formalized method for cataloging, rapidly sharing, or analyzing these data (Rawlings, 2002). Information was gathered from voluntary reporting by residents or

---

1 While this continuum provides important distinctions to understand the value of analysis and intelligence, it is too specific for the context of this paper. As such, the use of these terms in this text are generalized and should not be considered specific to the level of analysis identified by Ratcliffe (2011), though intelligence can be assumed as the highest level of attainment.
the investigations of the appointed tithing, the ten-man unit assigned to maintain law and order, and the parish constable (Rawlings, 2002). Because communities were smaller at that time, the activities of these contingents may have been effective in solving specific crimes. However, as societies became more compiled and mobile, further steps were needed to address growing crime problems.

Beginning in the late 1600s, the use of information took on a more formalized role. This was bolstered by the creation of the Metropolitan Police Force in London, England by Robert Peel in the early 1800s. Under Peel’s leadership, the focus of policing shifted from response to specific crime instances to crime deterrence and prevention (Rawlings, 2002). Police began to recognize organized criminal activity, which at that time, was considered a crime against the Crown (Rawlings, 2002). While data analysis did not progress significantly during this time, increased organization of the police force laid the foundation for data collection and sharing. Over the next century, several important developments contributed to building these capabilities including the creation of the first detective department (1842), the establishment of the Association of Chiefs of Police (1893), and information sharing through the Police Gazette (1945) to encourage interagency cooperation (Rawlings, 2002). However, deterrence remained the priority at that time, and the importance of data collection and analysis remained limited.

The Professional Era of policing in the United States, circa the early 1900s embraced the concept of information sharing, but not necessarily information analysis. Major technological advances in this time period, including the advent of the telephone, radio, and automobile provided the capacity for a more cognizant and accessible police force. However, these advances had the unfortunate consequence of ushering in an era of reactive policing practices.
in which crime prevention and deterrence was largely abandoned in favor of quick response to
calls for service from members of the public (Schmalleger & Worrall, 2009). The reactive
approach meant that while information was collected, there was little incentive or effort made
to analyze this information into actionable intelligence to address existing serious crime
problems or prevent new crime from developing.

Early eras of policing were slow to recognize the importance of data and information in
improving policing efficacy, but increased organization and communication provided the
groundwork for rapid developments in information use in the second half of the 20th Century.
Peel’s policing model instituted the formal organization needed to facilitate the collection and
sharing of data while the Professional Era in the United States brought the technology needed
to automate this process. Yet, the analysis of this information remained limited, preventing the
development of intelligence-led practice.

**Rapid developments in information calls for policing innovations.** The latter half of the
20th Century ushered in an era of rapid developments in policing practice. Stimulated by
significant social and political turmoil in the United States during the early to middle 1900s,
along with rapid development of computer technology and data collection efforts, the use of
information on crime and criminals became a necessary tool for police. Increased attention to
racial and gender inequality, prominent political events such as the Vietnam War and the Cold
War, and graphic media coverage of police brutality in riot situations called attention to the
need for objective evaluation of police operations and to restore faith in the police in the public
eye. Concurrent with, and immediately following, these conflicts the United States also
experienced a significant and steadily increasing crime rate (Bureau of Justice Statistics, 2017),
calling into question the effectiveness of police operations in addressing crime. This was compounded by a series of studies finding that the standard policing tactics of the era such as random preventive patrol (Kelling, Pate, Dieckman, & Brown, 1974), increasing the size of the police force (Levine, 1975), enhanced resources and training for investigators (Greenwood & Petersilia, 1975) and rapidly responding to calls for service (Spelman & Brown, 1984) were ineffective in reducing crime rates or closing cases.

The political climate, escalating crime rates and “nothing works” studies led to a prevailing view that policing had become too reactive and overly focused on the means of policing (amount of resources at their disposal, number of arrests made, average response time and so on) and had thus lost sight of the end goal of reducing crime and making their communities safer (see Goldstein, 1979). In short, there was tremendous pressure for police to try new and innovative tactics, resulting in the development of paradigms such as problem-oriented policing and community policing which sought to make police more proactive strategies to target root causes for crime prevention rather than merely responding to it after it has been reported. Along with this climate, the rapid development of computing technology and a growing academic interest in crime patterns during this time provided the tools necessary to foster the growth of information use in policing.

Despite the initial disappointing implications for the crime problem in the United States, increasing crime rates and critical research had the positive effect of encouraging more empirical research and encouraging innovation in policing practice. Over the next several decades, policing practice saw the growth of research examining crime in the macro-context by analyzing crime patterns. In addition, policing research began to recognize the influence of
underlying social and economic characteristics that has the potential to influence criminal behavior. This new theoretically- and research-informed approach spawned rapid policing innovation that has continued today. Among the key innovative approaches to policing that influence modern practice are community policing, problem-oriented policing, pulling levers policing, hot spot policing, and Compstat.

**Policing Innovation and the Use of Information and Intelligence.** Since the mid 1900s, police have sought new means to address their limitations, become more efficient and effective, and demonstrate accountability. This has resulted in a number of innovations, each of which incorporate some elements of information collection, information analysis, and intelligence development. These key innovations are summarized in the following sections.

*Community policing.* The Civil Rights Movements of the 1950s and 1960s and social protests to events such as the Vietnam War of the 1960s and 1970s highlighted a strained relationship between the government and the citizens. These events called attention to discriminatory practices, underrepresented groups of potential victims, and increased the need for police to extend their efforts beyond traditional reactive responses to calls for service. Images of violent police response to public protests highlighted the divide between the police and certain groups within the community they were intended to protect. In response, police departments began to shift their focus to improving police-community relations and policing effectiveness, leading to the development of community policing (Bureau of Justice Assistance, 2004).

Community policing is defined as “a philosophy that promotes organizational strategies that support the systematic use of partnerships and problem-solving techniques to proactively
address the immediate conditions that give rise to public safety issues such as crime, social disorder, and fear of crime” (Office of Community Oriented Policing Services, 2014, p. 1). This approach was novel in its use of citizens as a source of information to identify crime problems and underlying factors and to suggest options to addresses these issues (Goldstein, 1987). This shifted the focus of policing from reaction to individually-reported crimes to recognizing the social context in which crime occurs, utilizing a variety of sources of information to build a more comprehensive picture of crime problems, and assessing this information to make better decisions on how to respond. Community policing, or community-oriented policing, continues today as a component of most major police departments.

However, while community policing was an important step in increasing information and intelligence use in policing, research findings challenge the extent to which information is effectively used and the extent to which these practices are effective in addressing crime. Mastrofski (2007) suggests that community policing units within police departments often lack adequate resources and the departments themselves fail to embrace the community policing philosophy. As a result, community policing units are unable to adequately analyze community information, generate actionable intelligence, and respond effectively to recommendations (Mastrofski, 2007). Challenges and inconsistencies in implementation of community-policing efforts make it difficult to summarize its impacts. A recent systematic review suggested that impacts on crime and fear of crime have been limited, but there is evidence of positive impacts on citizen satisfaction, police legitimacy, and perceptions of disorder (Gill, Weisburd, Telep, Vitter & Bennett, 2014). Nonetheless, attempts at community policing highlight an important source of information and the potential for translating this information into usable intelligence.
**Problem-oriented policing.** In a closely related and contemporary innovation to community policing, problem-oriented policing emphasizes the assessment of information about the community to address the issues underlying specific crime problems. Problem-oriented policing was originally proposed by Herman Goldstein in 1979 and encourages the systematic assessment of crime problems and potential solutions (Eck & Spelman, 1987). This emphasis on assessment is indicative of converting information into intelligence.

Evaluations of problem-oriented policing programs suggest significant and positive impacts on reducing crime, but identify limitations in the analysis aspect of the process. In an extensive systematic review of problem-oriented policing studies, Weisburd, Telep, Hinkle, & Eck (2010) found that problem-oriented policing programs were associated with a significant though modest decrease in crime and that studies that were more methodologically sound produced more significantly robust findings. However, Braga & Weisburd (2007) argue that prior implementations of the problem-oriented policing model have been plagued by failure to adequately conduct the analysis phase and an overreliance on traditional policing responses. Despite these limitations, this focus on problem analysis and informed response demonstrates a significant philosophical jump in incorporating information and intelligence into policing practice.

**Pulling levers policing.** Pulling levers policing, also known as focused deterrence, takes problem-oriented policing a step further by building information and intelligence on specific criminal groups or crime types and using a variety of targeted responses to maximize the deterrent effects of police response (Braga, Kennedy, Waring, & Piehl, 2001). Pulling levers approaches target serious or pervasive crime problems by gathering extensive information on
individuals involved, contextual influences, driving mechanisms, and other related factors. These factors are analyzed and one or more targeted responses by police and other agencies.

A systematic review of ten pulling levers interventions found that these strategies had significant moderate effects on crime (Braga & Weisburd, 2012). While not all were successful, many of the interventions included in the systematic review incorporated information analysis and acted on intelligence products. Pulling levers policing contributes to the use of information and intelligence in policing beyond problem-oriented policing by targeting specific problems, emphasizing the gathering of information for specific individuals or groups, and maximizing potential law enforcement and community responses. Pulling levers policing represents a significant step in successfully incorporating analysis and the generation of intelligence.

*Hot spot policing.* Hot spots policing is a data-driven approach that recognizes the consistent finding that crime clusters in a limited number of places. Based on findings by Sherman, Gartin, and Buerger (1989) that crime most often occurs in a limited number of “micro-places,” they proposed that police can increase their efficiency and effectiveness by identifying these micro-places and focusing police resources at these locations. This approach has benefited from technological advancements, including geographic information systems (GIS), to quickly identify locations where crime is currently clustering. However, unlike earlier innovations, hot spots policing generally does not include an assessment of the context of crime and relies on traditional policing responses of targeted patrols and localized “crackdowns.”

In a comprehensive systematic review of 19 hot spot policing evaluations, Braga, Papachristos, and Hureau (2014) found noteworthy crime decreases in 20 out of 25 experimental/quasi-experimental tests and their meta-analysis revealed a small but significant
mean effect size indicating that hot spot policing strategies are effective in reducing crime. Their study revealed that the benefits of such interventions were often diffused to local areas, thus reducing crime around the hot spots, with little evidence of displacement of crime to other local areas. The use of crime data to identify the locations where crime clusters demonstrates a limited but important contribution to the use of information and intelligence in policing, particularly in its use of computing technology.

*Compstat.* Compstat, and similar programs, combine elements of hot spots policing and problem-oriented policing along with organizational changes to increase police accountability (see Weisburd, Mastrofski, McNally, Greenspan, & Willis, 2003; Weisburd, Mastrofski, Willis, & Greenspan, 2007). This model is intended to incorporate crime mapping, interactions with individuals and organizations in the community, and intervention evaluation (Silverman, 2007). Like hot spot policing, Compstat programs take advantage of technological advancements to examine spatial crime patterns and trends. In building upon hot spot policing, Compstat programs use input from individuals and organizations within the community to provide insight into factors underlying crime and potential solutions. As such, Compstat programs demonstrate improved collection of information and utilization of information analysis for the generation of policing intelligence.

However, critics argue that participation of outside individuals and organizations in the planning process is minimal and innovation in response is often stifled in favor of traditional police practices (Weisburd, et al., 2007). Several studies reinforce this criticism, finding that such programs favor management accountability and quick results over innovative and long-term crime reduction strategies (Weisburd, et al., 2003; Willis, Mastrofski, & Weisburd, 2004).
Ideally, Compstat programs incorporate many of the key components of intelligence-led policing, but the implementation appears to fall short.

*Predictive Policing.* This most recent approach to policing, predictive policing, relies heavily of advancements in computing and statistical modeling to help police identify not only where crime has occurred in the past, but where it is most likely to occur in the future (see Perry, McInnis, Price, Smith, & Hollywood, 2013; Sklansky, 2011). Alternatively referred to as crime forecasting, this approach aims to identify future crime trends to assist police with resource allocation and early intervention or prevention. As most predictive policing techniques have been in practice for less than a decade, there is little independent research evaluating their efficacy in crime prediction and subsequent crime prevention. In addition, the specific variables considered and statistical modeling used varies substantially across technique. While predictive policing is in its infancy, it represents an important, and potentially revolutionary, approach to information and intelligence use in policing.

Several forms of predictive policing exist that utilize different measurements and methods in their prediction. For example, kernel density estimation can be used to identify a “heat map” around hot spots of previous crimes to suggest the areas around where crime clusters exist and are thus likely to occur in the future (Perry, et al., 2013). Similarly, near repeat methods predict localized increased risk immediately following an instance of a specific type of crime, particularly residential burglary (Short, et al., 2009). PredPol is a commercially available software program that claims to utilize earthquake modeling algorithms to predict crime based on prior crime events though the actual methodology remains black box (Huet, 2015).
Such predictive policing techniques can provide valuable information to police departments but have limitations. Groff and LaVinge (2002) suggest that these methods rely too heavily on retrospective patterns, when crime forecasting efforts should focus on identifying indicators and early warning signals of future crime. While there is strong theoretical and empirical evidence to support crime hot spot and near repeat patterns (Braga, Papachristos, & Hureau, 2014; Short, et al., 2009), these approaches often fail to evaluate the influence of other factors that may contribute to the desistance of crime in hot spots or generate crime in future hot spots. As evident from prior policing innovations, there are structural, social, economic, and other contextual area characteristics that influence crime. As such, forecasting models should incorporate such measures not only to refine predictive models based on prior crime patterns, but to aid in the identification of places that may not currently have crime problems but are at risk.

RTM is a form of predictive crime modeling that seeks to address this limitation by evaluating contextual area characteristics as potential identifiers of future crime risk. RTM justification, theory, and methodology are thoroughly discussed in subsequent sections. However, it is important to recognize that RTM is a novel approach to predictive crime modeling that warrants extensive evaluation, but has the potential to be an asset to police as a forward-looking tool for crime prevention and intervention.

**Summary.** A number of factors have contributed to innovations in crime analysis and police practice over the last several centuries with almost exponential developments in the last few decades. As technology continues to develop along with big data analytic methods, we can expect continued advancement of information and intelligence use in policing including the use
of predictive policing. While these techniques offer intriguing and potentially revolutionary approaches to understanding and addressing crime, it is critical that such methods be thoroughly evaluated and refined prior to large-scale implementation, particularly at the risk of existing practice. The present study seeks to contribute to our understanding of a specific predictive policing practice, RTM, through a thorough demonstration of its capabilities in predicting crime.

**Defining Risk Terrain Modeling.**

RTM is a recently developed predictive crime analysis methodology that seeks to identify specific locations at the highest risk of future of crime base on a variety of underlying social issues, structural factors, and low-level crime patterns. RTM incorporates two innovative models of crime analysis: geographic modeling and predictive policing. GIS to generate computerized models of crime events and patterns (Caplan & Moreto, 2012). Analysis of these computerized models can be used to identify specific locations where risk factors associated with crime cluster to assist with police resource allocation. Further, GIS modeling can be used in conjunction with statistical analysis to produce “forecasting models” to identify not just where crime is currently clustering, but where crime is likely to cluster in the future. Predictive policing seeks to improve upon the “gut” feeling of patrol officers and reactive responses to crime mapping by identifying and targeting police efforts at places and times at increased risk of crime and deviance (Perry, et al., 2013). Such techniques often use advanced computer analytics and data mining to enhance the predictive accuracy of their models. RTM builds upon geographic modeling and predictive policing by providing a methodological framework for risk assessment.
It is important to note that while RTM is an innovative approach to crime analysis, it builds upon a strong history of geographic modeling methodology. As defined by Malczewski (2004), “land-use suitability analysis aims at identifying the most appropriate spatial pattern for future land uses according to specific requirements, preferences, or predictors of some activity” (p.4). Malczewski (2004) provides a thorough discussion of the history of suitability modeling, with a particular emphasis on the critical role of GIS and its use in a variety of fields including environmental science, urban/regional planning, and geographic and ecological sciences. While the construction of suitability models can vary significantly based on the intent of a given study, suitability modeling includes a process of problem analysis and model building to identify specific locations based on a set of criteria. For example, the Environmental Systems Research Institute (n.d.) prescribes a six-step process from problem identification through implementation. RTM reflects a similar process specified to the context of understanding crime risk.

The development of RTM is attributed to Joel Caplan and Leslie Kennedy with the Rutgers Center for Public Security at Rutgers University. The technique grew from a 2007 study of murder rates in Irvington, New Jersey (Caplan & Kennedy, 2011). According to Caplan and Kennedy (2010), RTM “is an approach to risk assessment that standardizes risk factors to common geographic units over a continuous surface... Risk Terrain maps assist in strategic decision-making and tactical action by showing where conditions are ideal for events to occur in the future” (p. 23). The technique provides a methodology for systematically assessing factors underlying crime to take a forward-looking approach to spatial crime risk prediction. While retrospective analyses of crime hot spots provide strong predictors of crime, RTM focused on
the correlates of crime to identify not only where crime is already a problem, but where it is likely to become a problem in the future. When used in conjunction with other crime analysis techniques, RTM can provide police and other community agencies with the means to seek early intervention before more serious crime problems develop. However, RTM is still developing both in methodology and in examination of its empirical validity. As such, more research is needed to evaluate RTM and its efficacy as a crime-prediction tool.

Comparison with existing crime modeling techniques. Interest in spatial modeling and analysis of crime has grown substantially in recent decades. Groff and LaVigne (2001) suggest that advancements in GIS technology led to rapid increases in the use of crime mapping in police departments from the early 1980s through late 1990s. These early spatial analysis techniques to model patterns of crime clustering to identify “hot spots” where crime occurs. Eck, et al. (2005) identify many of these techniques including standard deviation distance, nearest neighbor clustering, thematic mapping, and kernel density estimation. Regardless of the specific analysis method, the central focus of these techniques is the analysis of retrospective crime data to predict future crime. These methods are based on the assumption that where crime has clustered in the past, it will cluster in the future. Such approaches can be implemented by trained crime analysts using GIS software or through publicly available resources such as CrimeStat.

Several commercially-available spatial crime analysis software programs have been developed to assist police departments and other entities with identifying crime risk based on areas where crime clusters. For example, PredPol utilizes a self-learning algorithm to make short term predictions of small geographic areas of 500’x500’ at the highest risk of crime
based on prior crime patterns (PredPol, 2015). Police officers can then be allocated to these “boxes” to maximize their effectiveness (PredPol, 2015). A similar methodology is employed by CAP Index to aid businesses and government agencies in decision making based on crime risk. CAP Index improves upon PredPol by incorporating demographic characteristics (CAP Index, 2017), but utilizes a proprietary algorithm and does not appear to be theoretically informed. Because the methodology and algorithm are not available for review, it is difficult to determine the process behind this predictive analytic technique.

RTM is unique from other hot spot and predictive analytic techniques in two key ways. First, with the exception of the CAP Index software and individual studies of spatial crime correlates (see Cohen, Gorr, & Olligschlaeger, 2007; Groff & LaVigne, 2001; Groff & LaLavigne, 2002), predictive crime techniques have relied solely on prior crime trends to predict future crime of the same type. RTM omits crime events (other than low-level crime indicators) as predictors and focuses on environmental factors that are expected to be underlying correlates of crime. This allows a unique advantage to RTM modeling in the identification of places where crime may not yet be a problem, but where the environmental context exists that it may be in the future. Second, RTM is intended to draw upon criminological theory to identify the correlates used to predict crime risk. Using the Theory of Risky Places and other existing criminological theories, discussed in subsequent sections, RTM attempts to identify risk factors for crime that have previously been identified in criminological literature. In sum, RTM is distinct from other contemporary crime predictive analytic approaches in its use of correlates of crime rather than past crime as predictor variables.
Risk Terrain Modeling Methodology.

Because of the novelty and complexity of RTM process, this section is devoted to describing the general RTM methodology to familiarize the reader. This methodology is expanded upon in Chapter II, including its application specific to the present study. RTM is a multi-step, systematic process. Caplan and Kennedy (2011) outline 10 steps, each of which is summarized in the following discussion.²

Step 1 - Identify an outcome event (dependent variable) of interest. Caplan and Kennedy (2010) suggest the selection of a single crime outcomes event of interest or the use of multiple crime outcome events that are closely related. The outcome is generally classified as a dichotomous variable to represent the presence or absence of the crime of interest rather than a continuous measures of the number of that event that occur in a given area.

Step 2 - Identify area of study. The selection of the place should be meaningful for operational management (e.g., within the jurisdiction of a single law enforcement entity), that data available for the analysis will be able to cover the extent of the study area, and that the area selected will be applicable to the outcome event of interests (Caplan & Kennedy, 2010). Prior RTM demonstration studies are generally at the city level, though they vary from neighborhood to nation in scope.

Step 3 - Identify a time period for study. Time period refers to the duration of the study and can vary from a week to multiple years. Caplan and Kennedy (2010) provide little guidance

² Note that Caplan and Kennedy provide slightly different steps in their 2010 and 2011 texts. While the intent stays the same, there is some variation in the step number identifiers. This study utilizes the 2011 methodology, but includes guidance from both the 2010 and 2011 texts.
into the appropriate time frame for an RTM analysis, but suggest that it should be selected based on short- and long-term planning intentions.

**Step 4 - Identify risk factors (independent variables) related to outcome of interest.** The selection of risk factors should be as comprehensive as possible and can be informed theoretically, through meta-analysis, and/or through practical logic (Caplan & Kennedy, 2010). Data to be used in the analysis requires spatial identifiers. While they strongly suggest theoretically- and empirically-guided variable selection, they also identify that the final selection of variables to be used in the models is at the discretion of the researcher and should have reasonable justification to maximize credibility and validity of the model (Caplan & Kennedy, 2010). It is also important to note that while risk is the primary focus of RTM, protective or mitigating factors that may deter crime can also be included. Caplan and Kennedy (2011) provide suggestions for risk factors relevant to certain crime types.

**Step 5 - Obtain data and maps.** Geographic files, including shapefiles should be collected that represent the study area of interest.

**Step 6 - Generate maps for each risk factor and outcome measure.** Data must be converted to geographic files showing the distribution of the risk factor and operationalizing it into relative risk. The way in which risk factors are mapped depends on how they are operationalized. For example, buffers may be added to show the influence of a single risk factor over a wider area or density estimation can be used to show amalgamations or clusters of risk factors. Further, these map layers have to be reclassified based on risk. This classification essentially gives a risk “score” that indicates that risk is higher where certain risk factors are present and lower where those risk factors are absent. The means by which map layers are
generated and risk is classified is at the discretion of the researcher. Caplan and Kennedy (2010) offer several potential options for this process.

**Step 7 - Refine model to omit non-significant measures.** Caplan and Kennedy (2010) suggest that only those variables that are statistically significant should be included in the model to maximize model accuracy and validity. Chi-square tests are conducted for each risk factor to identify those that are significantly associated with the outcome measure. Those that are not statistically significant are dropped from the model.

**Step 8 - Apply variable weighting, as appropriate.** At the discretion of the researcher, weights may be added to those risk factors thought to be stronger or more important contributors to determining risk. Caplan and Kennedy (2010) suggest conducting a logistic regression with all risk factors and the outcome variable, and basing the weighting on the resulting coefficient. The coefficients become multipliers for the risk value of each risk factor.

**Step 9 - Combine individual risk maps into risk terrain map.** At this point, each risk factor layer is rastered to create a common grid. The values of each risk layer in the grid are then summed to give a total risk value for each cell. For example, assume there are 5 risk factors each with its own layer of individual cells, and each cell has a value of 0 (risk factor not present) or 1 (risk factor present). Once all of the layers are summed, each cell in the existing map will have a value between 0 and 5 with higher values representing higher risk.

**Step 10 - Communicate model findings.** The final step of the process identified by Caplan and Kennedy (2011) is to translate the risk terrain maps into findings that can be used to inform policy and practice. This can involve the generation of “heat maps” illustrating the areas of highest risk where police can plan to focus future resources. Further analysis of the models can
suggest those risk factors most strongly associated with risk. Using this information, police and other community organizations can seek means to mitigate these risk factors. This final step is the point at which information becomes intelligence.

**Theoretical Foundations of Risk Terrain Modeling.**

RTM is a data-driven approach to assessing crime risk, and as such, does not adhere to a specific criminological or criminal justice theory. While a closely-associated theory, the Theory of Risky Places, provides justification for RTM, it does not provide guidance on the specific factors related to crime or how those factors interact to generate risk. In practice, models are based on theoretical assumptions that crime is influenced by spatial characteristics and relies on existing criminological and criminal justice theories to identify risk factors. RTM is, at its core, an environmental approach to criminology that considers the “environmental backdrop” in which crime can occur (Brantingham & Brantingham, 1981). According to Kennedy, Caplan, and Piza (2012), “[t]he surrounding environment is very much a part of any criminal activity – as the environment emits cues which may or may not affect an offender’s decision-making or daily routines” (p. 13). RTM seeks to understand and evaluate the elements of the environmental backdrop and how they influence crime.

The mechanisms by which these elements influence crime can be theoretically informed by several existing criminological theories. In practice, RTM is largely informed by Broken Windows Theory, which posits that crime develops and flourishes in areas where physical and social disorder indicate a lack of both formal and informal social control (Wilson & Kelling, 1982). However, the approach to RTM can also incorporate a number of other theories, particularly if a targeted response is in mind. Such theoretical foundations may include Rational
Choice Theory, Routine Activities Theory, and Situational Crime Prevention. Each of these theories is summarized in the following as they relate to RTM.

**Theory of Risky Places.** Based on findings from early RTM studies, Caplan and Kennedy (2012) proposed the Theory of Risky Places (TRP). This theory postulates that some places are at a greater risk of crime based on a number of factors that exist in a spatial context (Caplan & Kennedy, 2012). They argue that “[r]isk is a function of threat, vulnerability, and consequence” (Caplan & Kennedy, 2012, p.1). This theory takes an almost mathematical approach to calculating crime risk by location. This is atypical of other criminal justice and criminological theories in that TRP is a theory driven by methodology rather than a theory to be verified by method. Caplan and Kennedy (2012) propose 3 tenets to their theory:

1) All places are risky, but because of the spatial influence of criminogenic features, some places are riskier than others.;

2) Crime emerges at places when there is high vulnerability based on the combined spatial influences of criminogenic features at said places.; and,

3) The overall effect of risky places on crime is a function of differential vulnerability and exposure throughout the landscape. (p.1).

TRP places emphasis on the spatial distribution of crime correlates and vulnerabilities as they relate to crime outcomes. This theory has many similarities to Routine Activities Theory, discussed in a subsequent section, but focuses more on place than individual crimes.
TRP provides an important context to justify the methodology of RTM, but is lacking in specificity as a criminological or criminal justice theory. The tenets of TRP support the additive nature of risk factors coming together at specific places. Characteristics of places can either provide criminogenic influence to encourage criminal behavior or contribute to vulnerability to crime. While this provides contextual insight, the theory has significant limitations in its lack of specificity. TRP does not provide guidance regarding what risk factors are relevant or how they may interact. Further, the theory provides no guidance regarding how some risk factors may be more influential than others.

These criticisms do not negate the value of the theory, but point to the capacity for additional development. More detail is needed to describe the selection of “criminogenic features” and how these converge and interact. This limitation can be addressed in part by drawing upon other criminal justice and criminological theory. The present study is not intended to test TRP or another specific theory. The present study is a demonstration and validation of methodology that draw upon RTP and other existing theories to create valid risk terrain models. The following subsections summarize theoretical guidance that both provides background for TRP and the basis for the present analysis.

**Broken Windows Theory.** In accordance with Broken Windows Theory (BWT), RTM proposes that future crime problems can be predicted based on the presence of underlying elements of physical and social disorder. Postulated by Wilson and Kelling in 1982, BWT suggests that these elements indicate to the community and potential offenders that “no one cares,” thus encouraging continued and escalated deviant and criminal behavior. More specifically, signs of untended disorder signal diminished social control in an area. This in turn
increases fear of crime among residents which leads them to withdraw from the community leaving it more vulnerable to further disorder and criminal invasion as would-be offenders are said to use disorder as cues that they can operate in the area with relative impunity. Thus, unintended disorder starts a cycle of decline (Skogan, 1990) that puts an area at risk for becoming a hot spot of crime.

While RTM can incorporate a number of variables, a key focus is the use of existing low-level disorder issues as a means to predict future crime problems. For example, RTM analyses often utilize measures of physical disorder (e.g., abandoned vehicles, malfunctioning street lighting) and minor crime (e.g., drug use/sales). These variables may be indicative of a lack of both formal and informal social control. Because of its methodological rather than theoretical emphasis, RTM does not adhere to the assumption that low-level disorder is a mechanism that fosters an apathetic or pessimistic attitude towards the areas. However, both BWT and RTM recognize the applicability of low-level disorder as a predictor and potential catalyst for more serious crime problems.

Empirical evaluations of policing responses based on BWT have demonstrated modest results on crime outcomes, but this does not diminish the value of BWT as a theoretical framework for understanding crime and RTM. In a 2015 systematic review and meta-analysis examining thirty studies of disorder-focused policing, they found modest but significant improvements in crime outcomes (Braga, Welsh, & Schnell, 2015). Weisburd, Hinkle, Braga, and Wooditch (2015) critique the findings of this study based on the suggestion that the studies included in the systematic review do not fully demonstrate the mechanisms underlying BWT. They assert that because of the construction of these studies, there is insufficient evidence to
demonstrate the connection between police impact on the relationship between the driver in BWT, physical and social disorder, and the mediating influence, collective efficacy and informal social control (Weisburd, et al., 2015). These findings do not negate the value of BWT for as a means for understanding the influence of disorder on crime, but point to the need for additional evaluation research on how policing responses are implemented in the broken windows’ context.

The present study does not attempt to evaluate a broken windows policing response, but rather seeks to utilize the principles of BWT as a framework for assessing spatial crime risk. As such, existing empirical evaluations are used herein to identify common measures used to identify crime and disorder. Prior BWT research has used a wide array of measures of physical and social disorder. For example, Sampson and Raudenbush (2004) examined thirty potential indicators of physical and social disorder ranging from observations of intoxicated people and fighting/arguing to burned out commercial buildings and the presence of alcohol and tobacco advertising. Physical disorder measures often incorporate the presence of abandoned or damaged buildings, overgrown foliage, litter, and graffiti. Social disorder measures have shown more variation. These measures may include perceived instances of disorderly behavior (e.g., loitering teens, drunken behavior, loud music) identified through survey research or through official reports of disorder crimes (e.g., truant children, prostitution, drug use) reported to police. To develop a comprehensive and useful RTM, the present study will attempt to incorporate a broad range of physical and social disorder measures that pose tangible intervention opportunities for police and public services.

Rational Choice Theory (RCT), Routine Activities Theory (RAT), and Situational Crime Prevention (SCP) are based on the similar premise that offenders are more likely to engage in deviant or criminal behavior when potential targets are more vulnerable and the risk of getting caught is lower. RCT proposes that offenders are rational actors who consciously weigh the potential rewards and risks prior to engaging in criminal behavior (Cornish & Clarke, 1997). As such, crime can be expected to occur where potential victims are vulnerable, possess something of high value, and where there is little chance that the offender will be apprehended. This is closely tied to the concepts of RAT, which posits that crime occurs at the intersection of a motivated offender, a suitable target, and a lack of capable guardianship (Cohen & Felson, 1979). Both RCT and RAT have elements of spatial context that directly apply to RTM. RTM can incorporate elements of RCT and RAT by including measures associated with the intersection of potential victims and offenders (e.g., public transit stops, retail and alcohol outlets) as well as factors indicative of a lack of guardianship and risk of apprehension (e.g., prior crime and disorder, code enforcement issues, absence of physical security).

The spatial context common to RCT, RAT, and RTM are incorporated into the concepts of Situational Crime Prevention (SCP). Proposed by Clarke in 1983, SCP emphasizes the reaction to crime problems by “target hardening.” SCP largely ignores the motivation of individual offenders. SCP suggests that crime can be prevented by identifying aspects of a given location that are vulnerable to crime and incorporating means to make these places more difficult to access by potential offenders (Clarke, 1983). Vulnerabilities may include limited visibility (e.g., limited street lighting, shrubs obscuring line of sight), lack of physical security (e.g., window
bars, door locks), lack of guardianship (e.g., limited police presence, abandoned buildings). In the context of both RCT and RAT, the presence of these vulnerabilities in a given areas indicate to potential offenders that the risk of being apprehended is limited, thus they encourage deviant and criminal behavior. SCP takes this a step further by suggesting means to protect vulnerable places by making the areas more difficult to access and increasing the risk of detection and apprehension. Measures of these vulnerabilities can be incorporated into RTM analyses to aid in the identification of places at risk of future crime and to provide a proactive means of addressing these vulnerabilities.

**Social Disorganization Theory.** As RTM is a methodological approach to criminology, rather than a specific theoretical approach, its development can also be tied to aspects of many criminological theories. While social disorganization theory and collective efficacy are not a functional aspect of RTM, many RTM applications incorporate sociodemographic characteristics into the risk terrain model either as control measures or as potential risk or protective factors. Social disorganization theory (SDT) posits that in densely populated urban areas - characterized by residential instability and concentrated socioeconomic disadvantage – residents will have conflicting values that can lead to antisocial interactions and diminished capacity for informal social control among residents (Kornhauser, 1978; Shaw & McKay, 1942). Further, these conflicting values and lack of social ties also result in diminished collective efficacy, or the ability of the community to come together and take action to address crime and delinquency (Sampson, Raudenbush, & Earls, 1997). Regardless of the mechanisms by which residential instability and concentrated disadvantage are associated with crime, social disorganization
variables that point to potential difficulties for residents to address crime are important contextual considerations in the development of a risk terrain model.

**Theoretical foundations summary.** In sum, while RTM is generally considered an atheoretical method of analysis, the basic principles and selection of variables can be rationally informed by several existing criminological theories outlined above. At the core of RTM is the assumption that the risk of crime can be predicted at micro-locations based on area characteristics and presence of protective and risk factors. This place-based approach is theoretically supported by principles of environmental criminology and the closely associated broken windows theory. In addition, the selection of analysis variables can be theoretically informed by RCT, RAT, SCP, and SDT. The following section summarizes several categories of variables applicable to RTM along with their theoretical and rational justification.

**Empirical Evaluations of Risk Terrain Modeling.**

RTM is a relatively new technique of crime risk analysis with the first area demonstration study published in 2009 based on data from 2007 (Caplan & Kennedy, 2011). As such, there have been few empirical evaluations. Ten studies were identified that were published in peer-reviewed journals between 2011 and 2016 that involved the use of RTM. There was substantial variation in the study location, outcome measure, risk factors considered, study duration, and focus of evaluation. This variation makes cross-study comparison impractical, but findings consistently support the use of RTM as a useful tool forecasting crime for both property crime (Moreto, Piza, & Caplan, 2014; Caplan, et al., 2015; Piza, et al., 2016) and violent crime (Kennedy, Caplan, & Piza, 2011; Kennedy, et al., 2016; Caplan, Kennedy, &
There are several key differences in study scope and implementation that make cross-study comparison difficult. First, RTM evaluations utilize a wide array and combination of predictive indicators. The number of indicators varies substantially from as few as three to more than twenty. The types of measures include, but are not limited to, prior related crimes, residences of individuals with criminal histories, known territories for gangs or drug markets, foreclosures, calls for disorder service (e.g., abandoned vehicles, street light outage), multi-family and at-risk housing, retail centers, transit nodes, gas stations and convenience stores, restaurants (both fast food and sit-down), schools, variety stores, strip and night clubs, schools, tattoo parlors, tobacco stores, pawn shops, on- and off-site liquor outlets, foreclosures, hotels/motels, laundromats, parks, pawn shops, and bars. Only two of the ten identified studies included more traditional neighborhood disadvantage variables such as percent below poverty line and residential instability. Further, the influence of these risk factors can be operationalized in a number of ways including the presence/absence of the risk factor, density (clustering) of adjacent risk factors, and distance from risk factors (Caplan, 2011). Kennedy, Caplan, and Piza (2011) compared RTM models generated for four, three-month time periods and found that the significance of predictors used in those models varied across time along with the predictive validity of the RTM models themselves. Piza et al. (2016) found that the significance and influence of the risk factors in the RTM models varied based on neighborhood contextual variables of disadvantage and residential instability. These findings suggest the importance of carefully considering a range of variables and their significance, but also point to the need for
additional research and theoretical development to guide the determination of model variables.

A second key difference in existing RTM evaluations is study location. RTM evaluation studies have been conducted in several cities including Newark and Irvington, New Jersey; Chicago, Illinois; Little Rock, Arkansas; Fort Worth, Texas; and Colorado Springs, Colorado. Vast differences exist not only in region but in demographic characteristics (United States Census Bureau, n.d.a) across many of these locations. Given the consistent findings demonstrating the predictive capability of RTM models, the application across multiple locations further reinforces the applicability and utility of RTM in a variety of locations. However, variation in the scope of these evaluations studies makes comparison of evaluation findings across place unfeasible.

A third important difference across existing RTM evaluations is the measurement of the outcome crime of interest. Prior RTM evaluations have looked at a variety of crime outcome measures including shootings, assault, burglary, residential burglary, assault, robbery, composite violent crime measures (including combinations of homicide, aggravated assault, weapons violations, shootings, and robbery), motor vehicle crime, and child maltreatment. Studies varied in their treatment of the outcome measure as either dichotomous (presence/absence of a crime) or count (number of crimes per raster cell). Again, the consistent findings that RTM significantly identifies areas at highest risk of crime supports its applicability across multiple crime types. Meta-analytic techniques may be effective in comparing the regression results in a future study. It appears that RTM techniques are effective in identifying risk across this wide range of crimes, but variation in implementation makes it difficult to determine is RTM is more effective in determining risk for one type of crime over another.
Finally, it is important to note another key difference in RTM evaluations: timeframe. Prior studies vary substantially in the time period of variables used in the analysis. For example, Kennedy, Caplan, and Piza (2011) build their RTM models using predictor variables collected over a three-month period with crime data collected for the three months immediately following. In contrast, Drawve, Thomas, and Walker (2016) build their RTM models using predictors variables and crime outcome variables collected over the entirety of 2013. Both show differences in the time span covered, three months versus one year, and predictor crime periods, whether or not the crime outcome data used for modeling was from the same or a subsequent time period. Two studies (Drawve, Thomas, & Walker, 2016; Drawve, 2016) included post-test evaluations to compare the initial RTM models generated to crime data collected for subsequent time periods. This is an important distinction from traditional RTM validation, which measures the predictive validity of the RTM model using the crime data used to generate it. Drawve (2016) found RTM to consistently predict crime outcomes (robberies in Little Rock, Arkansas) across a range of time periods from 2 days to 36 months. Thirty-six months is the longest time span over which RTM models have been examined, as identified for this study. Additional research is needed to optimize the time span over which RTM models are generated and to evaluate their predictive validity over longer periods of time.

The variations identified herein suggest that RTM is applicable over a wide range of contexts, but further research is needed to allow for cross-study comparison. As such, this variation is both an asset and a potential limitation. Nonetheless, there are several important findings from existing RTM evaluations, in addition to strong support for its effectiveness in risk identification, that warrant further consideration. Several studies have found that RTM models
outperform traditional hot-spot analysis techniques (Caplan, Kennedy, and Miller, 2011; Drawve, 2016; Kennedy, Caplan, & Piza, 2013). Drawve (2016) found RTM to be a consistently better predictor of crime (robbery) than four other hot-spot analytic techniques, and was second only to kernel density estimation in accuracy of predictions. Caplan, Kennedy, and Piza (2013) suggest that RTM is not necessarily better than other techniques, but can be used in conjunction with them to better meet the needs of the intended audience.

In addition to the effectiveness of RTM in predicting risk of crime outcomes compared to other analytic technique, early research is also promising for police response to RTM findings. In a 2015 report for the National Institute of Justice, Kennedy, Caplan, and Piza demonstrated the effectiveness of police responses to findings from risk terrain models across five US cities. Their study examined the performance of risk terrain models and subsequent targeted interventions in Chicago, IL, Colorado Springs, CO; Glendale, AZ; Kansas City, MO; and Newark, NJ (Kennedy, Caplan, & Piza, 2015). The analysis process involved identifying a specific crime problem in each city and conducting a risk terrain model to identify problem areas and associated risk factors. A targeted intervention was then implemented in each area based on the identified risk factors. In sum, four of the five evaluations resulted in crime reductions of 12-42% in the target areas compared to the control areas (Kennedy, Caplan, & Piza, 2015). The fifth study, in Chicago, IL, did not yield sufficient data to conduct the evaluation (Kennedy, Caplan, & Piza, 2015). Additional research is needed to understand the effectiveness of RTM and RTM-based interventions, but these early studies suggest promising potential for the use of RTM as a tool to address crime.
Initial evaluations of RTM are certainly promising, but additional research is needed. Additional evaluations increase the capabilities of cross-study comparison and the potential identification of means to enhance the modeling technique. The present study aims to build the evaluation literature for RTM by providing demonstration in a unique place, examining additional predictor variables, and increasing the range of post-test evaluation.

**Correlates of Predatory Street Crime in an RTM Framework.**

As observed in the review of prior RTM demonstration studies, there is substantial variation in the predictor variables used in RTM models. Because the RTM methodology is primarily atheoretical, there is little guidance regarding what variables should be used to generate the most comprehensive and valid RTM models. In their 2011 text, Caplan and Kennedy provide literature reviews for a number of crime types that identify previously identified environmental correlates of crime. However, they also suggest in their 2010 text that variables selected can be informed by any combination of theory, prior empirical research, and practical logic. Each of these were considered in the selection of variables for the current study.

The following subsections briefly describe several categories of predictor variables relevant to the present study’s outcome variable: predatory violent crime. In this study, predatory violent crime includes homicides, aggravated assaults, and pedestrian robberies. Discussed in more detail in Chapter III, these crime types are indicative of violent interpersonal crime. The identified variable categories provided guidance on the selection of variables, though the final list of variables included in the study were based on available data resources and data quality.
**Criminal elements.** In contrast to other hot spot spatial crime analysis methods, RTM does not rely on the occurrence of prior offenses to predict crime of the same type. This is consistent with RTM’s approach to identifying risk in the environmental context and understanding why crime occurs, not just where (Caplan & Kennedy, 2011). Measurements of crime in a spatial context are often used as risk factors in RTM models, but are included as precursors or correlates to the outcome crime of interest. For example, in a study by Caplan, Kennedy, and Miller (2011) the locations of drug arrests and known gang member residences were significant predictors in the RTM model for shootings in Irvington, New Jersey. They point to empirical evidence (Brantingham & Brantingham, 1981; Fagan & Wilkinson, 1998, Klein, 1995; Blumstein, 1995; Lum, 2008) that both gangs and drug markets, which may be expected around both risk factors, are closely tied to violent crime (Caplan, Kennedy, & Miller, 2011). Locations of drug arrests are used as a proxy for police intervention in drug markets while known gang member residences are used as a proxy for gang territory (Caplan, Kennedy, & Miller, 2011). As such, while these variables are related to criminal events, the focus is on the context of the crime, not the crime instance itself.

In an RTM model intended for predatory violent crime, criminal elements would include the location of individuals, groups, or markets associated with criminal activity that may be either an indicator or a contributor to predatory violent crime. The presence of these criminal elements in a specific location is expected to contribute to risk of violent crime by increasing exposure to potential criminals. The connection between illicit markets, gangs, and violent crime is well documented in literature on pulling levers policing (Braga, 2008; Braga, Kennedy, Waring, & Piehl, 2001; Braga & Weisburd, 2012), and recidivism by prior offenders is a
prominent phenomenon in criminology (Durose, Cooper, & Snyder, 2014). Prior RTM evaluations have included several criminal element risk factors related to violent crimes such as gang residences/territories (Caplan, Kennedy, & Miller, 2011; Kennedy, et al., 2016; Caplan, Kennedy, & Piza, 2013) and drug arrests/markets (Caplan, Kennedy, & Miller, 2011). Daley, et al. (2016) used several prior crime measures to predict child maltreatment including aggravated assaults, murders, domestic violence, drug crimes, gang presence, prostitution offenses, and robberies. Kennedy, Caplan, & Piza (2011) also included the address of parolees in their RTM model, but found that it was not a statistically-significant predictor of shootings. Other measures that may be relevant but have not yet been tested in an RTM include addresses of jail releases, addresses of probationers, and locations of other illicit markets (e.g., brothels, gun markets).

**Physical and social disorder.** Consistent with the tenets of BWT, the presence of physical and social disorder is indicative of neighborhood decline, a lack of social control, and the potential for more serious crime to develop. In an RTM model of predatory violent crime, places with existing or increasing levels of physical and social disorder should be at higher risk for future crime. Physical and social disorder have not been included in many RTM evaluations to date, but could be a very important predictor. Kennedy et al. (2016) provide the only violent crime RTM that included physical disorder measures, and found that calls for service related to street light outages and abandoned vehicles were significant predictors of assault. Daley, et al. (2016) is the only study to include a measure of social disorder, prostitution, but found that it was not a significant predictor of child maltreatment. As such there is limited research on the role of physical and social disorder in determining risk.
Despite the limited existing research, physical and social disorder warrant further evaluation as predictors for RTM models. According to Skogan (2012),

Physical disorders include the overt signs of negligence or unchecked decay as well as the visible consequences of malevolent misconduct. These include abandoned, boarded up, or severely dilapidated buildings; abandoned, stripped, and burned out cars; collapsing garages; broken streetlights; junk-filled and unmowed vacant lots; street litter; loose syringes and condoms laying on the pavement; illegal dumping; garbage-strewn alleys; graffiti; and of course, broken windows. (p. 175-176)

This statement identifies many potential measures of physical disorder that can be considered for an RTM analysis. Code enforcement units or agencies within that serve the area of the study are ideal sources of data for these risk factors. Skogan (2012) describes social disorders as “unsettling or potentially threatening and perhaps unlawful public behaviors” (p. 175). This may include truant or loitering youths, panhandling, overt prostitution, overt drug use, or public drunkenness. Many of these can be identified through low-level offenses reported to police. While these measures of physical and social disorder may not directly contribute to violent crime, their potential impact on social control may contribute to the risk for future crime.

**Risky places.** Perhaps the most prominent risk factor in prior RTM evaluations is the presence of “risky places.” RTM studies include the location of a wide range of business entities that are hypothesized to contribute to crime. While the justification for each entity varies across study, the rationale can be summarized into three key categories: alcohol influence, cash
businesses, and group congregation. Other potentially “risky places” exist, but those described in the following are most commonly found in RTM models. Each of these entities is hypothesized to increase risk through increased vulnerability.

Alcohol influence. Alcohol influence refers to entities in which alcohol is served for either on-site or off-site consumption. These may include businesses such as bars, restaurants, night clubs, strip clubs, and liquor stores. In a comprehensive literature review, Snowden (2015) summarizes a litany of empirical studies showing a significant relationship between alcohol consumption and violent crime. In short, alcohol can contribute to individual vulnerability or aggressiveness in addition to situational cues that can encourage violence (Snowden, 2015). As such, violent predatory crime risk should be higher in areas where alcohol is sold or consumed. This was illustrated in Grubesic and Pridemore’s (2011) study of alcohol outlets in Cincinnati, Ohio which found higher levels of violent crime around clusters of alcohol outlets. One exception to the alcohol-crime connection in a spatial context is the presence of liquor stores. While alcohol is sold at these locations, it is intended for off-site consumption. Thus, spatial patterns may be more difficult to determine for alcohol-related businesses targeted at off-site consumption.

Though the measurement varies, alcohol entities are common in several RTM evaluations related to violent crime. These may be included in the model individually with risk map layers for each type of entity (e.g., bars, restaurants, clubs), as composite groups of alcohol entities, or included with other risky places. Despite the measurement type, both on- and off-site alcohol outlets have been identified as significant predictors of violent crime risk in RTM
models (Drawve, Thomas, & Walker, 2016; Caplan, Kennedy, & Miller, 2011; Drawve, 2016; Kennedy, et al., 2016).

Cash businesses. Wright et al. (2014) discusses the strong theoretical connection between cash economies and street crime suggesting that cash-oriented crimes such as robbery are driven by ready access to cash. Cash-based businesses such as pawn shops, check-cashing businesses, variety stores, grocery stores, convenience stores, laundromats, fast food restaurants, and other general retail businesses are commonly found to be significant predictors of violent crime in RTM models (Caplan, Kennedy, & Piza, 2011; Kennedy, et al., 2016; Caplan, Kennedy, & Miller, 2011; Drawve, 2016; Drawve, Thomas, & Walker, 2016). While many of these business entities may be better predictors of violent commercial/business robbery, rather than the violent offenses included in the present study, cash transactions at these locations may drive pedestrian robbery in the immediately surrounding areas as potential offenders are aware that patrons are likely to be entering or leaving with cash.

Group congregation. The third type of risky places represent locations where large numbers of individuals likely congregate. This can include a wide range of entities such as schools, sports arenas, apartment buildings, and hotel/motels. The common characteristic shared between these locations is that groups of individuals come together, perhaps with limited capable guardianship. For example, Murray and Swatt (2013) found that the presence of high schools was significantly associated with increased felonious assault both during and after the school day in the immediately surrounding area. Three RTM evaluations found schools to be significant predictors of violent crime (Caplan, Kennedy, & Piza, 2013; Drawve, Thomas, & Walker, 2016; Kennedy, et al., 2016). Others are consistent with the tenants of RAT, which
posits that crime will occur when potential offenders and victims converge in the absence of capable guardianship (Cohen & Felson, 1979). While there may not be crowds at hotels/motels, there is no guardianship in private rooms, allowing crime to go unchecked. Drawve, Thomas, & Walker (2016) found hotels/motels to be a significant predictor of violent crime risk in their RTM model. At sports arenas, gathering crowds bring many potential offenders and victims together, often exacerbated by the presence of alcohol and the lack of adequate security. These and other locations where groups of individuals congregate, either over a short-term or long-term period warrant consideration as potential risk factors.

**Sociodemographic conditions.** Sociodemographic factors are important independent and control variables throughout criminological literature, but are only found in a few RTM evaluations. This is an interesting omission as area sociodemographic characteristics may not themselves be risk factors for crime, but their context can affect the impact of other risk terrain variables. Using a number of common demographic variables (e.g., percent population below poverty line, geographic mobility, racial heterogeneity), Piza, et al. (2016) found that the influence of other risk factors on their outcome measure, motor vehicle crime, varied by the neighborhood context. While motor vehicle crime is distinctly different from the predatory violent crimes used in the present study, their findings indicate that sociodemographic characteristics warrant further analysis in RTM.

Consistent with the tenets of SDT, measures of socioeconomic disadvantage are associated with crime because they impede social control (Kornhauser, 1978; Shaw & McKay, 1942). As such, one may expect higher RTM risk values in areas where there are indications of socioeconomic disadvantage. Common measures of disadvantaged include median income,
percent unemployed, percent of households with income below poverty line, percent receiving public assistance, percent black or percent non-white, number of single family households, and residential stability. These measures are consistent with SDT. Neither of the two RTM evaluations identified for this study that included sociodemographic variables included them as risk factors. Piza et al. (2016) included a consolidated measure of socioeconomic disadvantage and residential instability as an interaction term for other risk factor variables. Drawve, Thomas, & Walker (2016) included such variables in a secondary analysis to examine their RTM model in a neighborhood context. Further research is needed to examine these sociodemographic variables as controls in the RTM analysis. Other demographic variables may be included as well such as age, sex, and level of education. Sources of this data may include Decennial Census, the American Community Survey, and local agencies overseeing public assistance or public health.

**Summary.** Prior RTM evaluations have examined many potential risk factors, with many more remaining to be tested. The purpose of the present study is not to test all possible risk factors, but to select those best able to predict predatory violent crime. Risk factors for the present study are based on prior RTM research as well as other available theoretically- and practically-informed data sources. The final selection of variables is detailed in Chapter III.

**Present Study.**

The present study seeks to add to the body of literature examining RTM as a crime-risk-prediction tool in several ways. RTM is a relatively new technique and the methodology varies substantially across studies. This study explores aspects of this variation and provides additional insight into the predictive validity of RTM.
First, because of the novelty and complexity of RTM, additional research is needed to further evaluate its applicability in different contexts. To date, RTM has been applied primarily in three cities: Newark and Irvington, New Jersey; Little Rock, Arkansas; and Chicago, Illinois. As previously discussed, there are substantial differences in the characteristics of these areas. The present study examines the applicability of RTM to unincorporated DeKalb County, Georgia. Several aspects of DeKalb County, including a higher than average non-white population and a high poverty rate (United States Census Bureau, n.d) yields a unique environment in which to test RTM. Further, while previous studies included RTM models at the city level, the present study is the first to apply the approach at the county level, which is geographically larger and more varied than units of analysis in past applications of RTM. This study will examine whether RTM retains it predictive validity across this unique landscape.

Second, the present study incorporates several spatial variables not previously included in RTM models. As detailed in Chapter III, this study includes measures of socioeconomic disadvantage and area economic health as potential risk factors for the RTM models. Socioeconomic disadvantage, a key component of many criminological studies, is strikingly underrepresented in RTM research. This study utilizes several measures of socioeconomic disadvantage and demographic variables as both risk factors and controls. This study also incorporates measures of area economic health, which has not appeared in prior RTM evaluations. While socioeconomic measures capture the economic well-being of residents, area economic health measures capture the relative economic conditions related to property valuation and businesses. Though closely related, area economic conditions are intended to measure non-residential economic impacts on environment as it applies to risk. In addition to
tradition RTM variables, the present study places an emphasis on the social and economic context of risk assessment.

Third, this study examines the validity of RTM over a longer period of time than has been seen in previous RTM research. Prior RTM studies have generally produced a single RTM model that spans one year of data or a few models each spanning only a few months. The longest time span covered in a study was Kennedy, Caplan, and Piza’s 2011 evaluation of shootings in Newark, New Jersey which covered five three-month periods from July 2008 through September 2009 (15 months). In addition, few studies have included post-test validation (i.e., comparing RTM models to subsequently collected data that was not used to construct the model). In the only extensive post-test validation, Drawve (2016) covered several follow-up time spans from two days to three years. The present substantially builds upon prior RTM research by comparing five RTM models each spanning one year (2010-2014) and examining post-test validation up to five years out (2011-2015). As neighborhood change is often a slow process and the effects of neighborhood change may take years, this extended post-test validation is an important contribution to understanding the applicability of RTM to long term planning.

The remainder of this text details the analysis and findings. Chapter III provides a detailed discussion of the analysis plan, technique, and variables used in the study. Chapter IV includes the results of each stage of the analysis process. Finally, Chapter V summarizes findings, discusses implications of those findings, and suggests directions for future research.
Chapter III – Methodology

The purpose of the present study is to examine the utility and effectiveness of Risk Terrain Modeling (RTM) as a tool to identify micro-areas at the highest risk of crime, including those areas that may not currently demonstrate a significant crime problem but are at elevated risk for future crime. This study implements the step-by-step evaluative method of predictive model building and analysis methodology of RTM utilizing crime data and risk-factor data from DeKalb County, GA. In addition, this study examines the predictive validity of RTM in successfully identifying those places at the highest risk of future crime. This chapter details data sources, variables, analysis methodology, and evaluation procedures used.

Summary of the RTM Process.

Before proceeding to sample and variable specification, it is important to identify the steps of the RTM process to clarify how data will be used. A detailed description of how these steps were implemented in the present study is provided in subsequent sections. The following provides a general summary of the techniques used in this analysis. The following steps are identified by Caplan and Kennedy (2011).

Step 1: Identify an outcome event (dependent variable) of interest

Step 2: Identify area of study

Step 3: Identify a time period for study

Step 4: Identify risk factors (independent variables) related to outcome of interest

Step 5: Obtain data and maps

Step 6: Generate maps for each risk factor and outcome measure

Step 7: Refine model to omit non-significant measures
Step 8: Apply variable weighting, as appropriate

Step 9: Combine individual risk maps into risk terrain map

Step 10: Communicate model findings

This study adds further step providing model validation.

Step 11: Examine predictive validity and model fit.

As detailed in subsequent sections, the present study includes the development of risk terrain models for predatory violent crime annually from 2010 through 2014. These models are then compared to outcome measures for subsequent years through 2015.

Data and Sample.

Data for this study include spatial measures of social, economic, disorder, zoning, business, and crime characteristics from a variety of sources to provide a thorough foundation to predict crime risk. The following subsections identify general characteristics of the study sample. Additional details related to the source and measurement of variables are included in subsequent sections.

Location. The location for analysis is limited to the unincorporated area of DeKalb County, GA. DeKalb County is located in the northeast metropolitan Atlanta area. DeKalb County is culturally and economically diverse, with a population of over 700,000 (United States Census Bureau, n.d.a). Within DeKalb County, several cities have obtained incorporated status. These cities are primarily located in the northern half of the county and consist of approximately one-third of the population of DeKalb County (Niesse, 2016). Incorporated cities include the following: Dunwoody, Brookhaven, Chamblee, Doraville, Tucker, Clarkston, Stone Mountain, Atlanta, Decatur, Avondale Estates, and Lithonia. Incorporated areas provide many
government services, including police and law enforcement, independently from those provided by the county. It was not within the scope of the present study to collect data from each of the independent agencies within the incorporated cities. As such, the present study is limited to the unincorporated area of DeKalb County, as illustrated in Figure 3.1.

Figure 3.1 – Unincorporated DeKalb County Map
Table 3.1 details the demographic characteristics of the unincorporated DeKalb County area.

Table 3.1 – Demographic Characteristics of Unincorporated DeKalb County GA (2010)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total population</td>
<td>564614</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Population Median</td>
<td>34.9</td>
</tr>
<tr>
<td>% Under 19 Years of Age</td>
<td>34.3</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>% Male</td>
<td>47.4</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>High school diploma or above</td>
<td>86.9</td>
</tr>
<tr>
<td>Bachelor degree or above</td>
<td>34.5</td>
</tr>
<tr>
<td>Income</td>
<td></td>
</tr>
<tr>
<td>Population Mean</td>
<td>52328</td>
</tr>
<tr>
<td>Population Median</td>
<td>48831</td>
</tr>
<tr>
<td>% Below Poverty Line</td>
<td>13.6</td>
</tr>
<tr>
<td>% Unemployed</td>
<td>12.5</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>% White</td>
<td>29.5</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>% Hispanic or Latino</td>
<td>9.0</td>
</tr>
</tbody>
</table>

**Time frame.** The present study involves the generation of five predictive models - one for each year from 2010 to 2014. RTM models are generated by comparing predictor variables for a given year to crime patterns for the same year to identify spatial weights. Tests of statistical validity were conducted with the subsequent years of crime data from the model year through 2015. Each of the individual yearly models are compared side by side to examine consistency of predictions and are also compared to crime data from the year immediately following the model through 2015 to examine post-test accuracy of predictions. Full data sets were generated for 2010, 2011, 2012, 2013, and 2014. Crime data were collected for each of these years as well as for 2015.
**Unit of measurement.** Data for this study were collected at the level of measurement appropriate and available for each variable. Data were measured at the smallest available geographic unit possible, while maintaining standards of confidentiality to prevent the identification of specific individuals or households. Variable data were available at levels of geography ranging from point (address) to Census tract. Additional details on measurement level are provided in the subsequent sections specific to the individual variables.

**Unit of analysis.** The unit of analysis for this study is the raster cell. Raster cells are generated by splitting a geographic area into a grid format of equally-sized squares. These raster cells represent micro-areas. Raster cell size is at the discretion of the researcher, but should be small enough to identify specific areas of risk. Caplan and Kennedy (2011) recommend that raster cell size should be equal to one-half of the average block size. Block size is determined based on designations by the United States Census Bureau and represent the smallest geographic unit of measurement. Census blocks “are statistical areas bounded by visible features, such as streets, roads, streams, and railroad tracks, and by nonvisible boundaries, such as selected property lines and city, township, school district, and county limits and short line-of-sight extensions of streets and roads” (United States Census Bureau, n.d.b, p. 1). The average block length within Unincorporated DeKalb County is equal to 845 feet, which is substantially larger than is seen in other RTM studies. This may be because the present study covers a larger geographic area than incorporated urban, suburban, and rural land areas which leads to wide variance in block length. As such, the median block length equal to 675 feet will be used for the present study as it is closer to those values used in other RTM evaluations. Using the one-half block size methodology proposed by Caplan and Kennedy, 2010), this
produced a 338’x338’ raster cell as the unit of analysis. This is within the raster cell sizes seen in other studies which range from approximately 100’x100’ (see Caplan, Kennedy, & Miller, 2011) to 425’x425’ (see Caplan, et al., 2015). This resulted in a total of 41,229 raster cells.

**Dependent Variables.**

The basic outcome of any risk terrain model is a crime occurrence. The present study examines the application of RTM methodology to explain predatory violent crime. Predatory violent crime in the current study includes homicide, aggravated assault, and pedestrian robbery. This variable is intended to measure interpersonal violent crimes between individuals and intentionally omits violent events against businesses (e.g. business robbery). Simple assaults are omitted to allow the analysis to focus on more serious crimes that are likely to come to the attention of the public. Sexual assaults and kidnapping were omitted because of the specific offender/victim context. Because the total number of these events per year is relatively small (approximately 2500), logistic regression will be used. The dependent variable will be the presence (1) or absence (0) of predatory violent crimes in each raster cell.

It is important to note that RTM models must be tailored to the outcome crime of interest. While there may be some similarities, there are also substantial differences in the risk factors applicable to each crime type. For example, characteristics in residential areas such as vacant properties or damaged street lights may be irrelevant to understanding commercial robbery. The measures of predatory violent crime were selected in the present study because these crime types can be present in a variety of environments. This was deemed a better approach to test the utility of the RTM methodology in the varied context of Unincorporated DeKalb County.
The DeKalb County Police Department provided data on all crimes reported to or identified by police from 2010 through 2015. For each incident, the following data points were used to generate the dependent variables for this study.

- Incident address (street address)
- Incident date (year)
- Crime Type (UCR code and descriptor)

Each incident was geocoded to the street address to create point shapefiles for each year. Approximately 3.7 percent of all reported incidents (all crime types) between 2010 and 2015 could not be located due to improperly recorded addresses provided in the original data.

Because some crime events occurred outside of the bounds of the present study, the resulting geocoded file was clipped to include only those crimes occurring in the Unincorporated DeKalb area. The three crime types used to measure predatory violent crime – homicide, aggravated assault, and pedestrian robbery – were selected to generate a map layer for the dependent variable. Crime types were determined based on the Uniform Crime Report (UCR) code, a standard identifier of criminal offenses.

The total number of Predatory Violent Crime offenses in Unincorporated DeKalb County are summarized in Table 3.2.

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homicide</td>
<td>118</td>
<td>79</td>
<td>76</td>
<td>68</td>
<td>112</td>
<td>102</td>
</tr>
<tr>
<td>Aggravated Assault</td>
<td>845</td>
<td>687</td>
<td>778</td>
<td>693</td>
<td>767</td>
<td>937</td>
</tr>
<tr>
<td>Pedestrian Robbery</td>
<td>1325</td>
<td>1306</td>
<td>1376</td>
<td>1514</td>
<td>1336</td>
<td>1549</td>
</tr>
<tr>
<td>Total Predatory Violent Crime</td>
<td>2288</td>
<td>2072</td>
<td>2230</td>
<td>2275</td>
<td>2215</td>
<td>2588</td>
</tr>
</tbody>
</table>
Once the maps were joined to the grid, the number of predatory violent crime offenses for each cell were identified. Cells with one or more violent crimes were coded with a 1 and cells with no crimes per cell were coded as 0 for the dependent variable.

**Independent Variables.**

The independent variables selected for this study were derived from a comprehensive study conducted by Georgia State University on behalf of the DeKalb County District Attorney. A broad, atheoretical approach was taken for the DeKalb County study to identify as many potential risk factors as possible to build RTM models for a variety of crime types. The independent variables for the present analysis of predatory violent crime were selected from those identified in the DeKalb County study. Empirical literature for each potential risk factor was reviewed to ensure a logical justification was available for those risk factors used in the present study.

The independent variables in the present study measure potential risk factors for predatory violent crime. These risk factors include measures of disorder, criminal elements, “risky places,” socioeconomic conditions, and area economic health. The following subsections detail variable measurement, data sources, and operationalization. While all variables included in this discussion are assessed, only those that are found to be significantly associated with predatory violent crime will be included in the RTM models.

In addition to risk factors, RTM can also incorporate specific protective factors. While protective factors can include specific entities that may deter crime, the absence of risk factors can also be “protective.” For example, while an area with low property values may be indicative of economic disadvantage, an area with high property values is likely to have more economic
resources to deter crime problems either formally or informally. As such, while measures of potential crime correlates are referred to as “risk factors” throughout this study, their absence can be equally seen as “protective.” The present study focuses on identifying areas at highest risk of future crime and does not include factors that are specifically intended to be protective from or mitigating of crime.

It is also important to note that while each of the following variables were considered for analysis, not all were included in the final models. Step 7 of the RTM process involves assessing each potential independent variable for its statistical significance to the model. Those variables that are not significantly related to the outcome are omitted to improve the explanatory value of the model. As such, each of the variables identified in the following subsections are included in the final RTM models only if they were found to be statistically significant predictors of the outcome variable in analyses presented in Chapter IV.

Each of the independent variables are detailed in the following subsections and summarized in Table 3.3, including indication of those variables used in prior RTM models examining violent crime. Values specific to each model are included in Chapter IV.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Measure</th>
<th>Data Level</th>
<th>Modifiers</th>
<th>Previous RTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical and Social Disorder</td>
<td>Code compliance violations</td>
<td>Parcel</td>
<td>1 block radius (675’ buffer)</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Foreclosures (proxy for vacant properties)</td>
<td>Parcel</td>
<td>None</td>
<td>Yes</td>
</tr>
<tr>
<td>Criminal Elements</td>
<td>Probation supervision address</td>
<td>Address (point)</td>
<td>Kernel Density Estimation (High Risk = &gt; 2 SD)</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Parole supervision address</td>
<td>Address (point)</td>
<td>Kernel Density Estimation (High Risk = &gt; 2 SD)</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Narcotics offenses</td>
<td>Address (point)</td>
<td>Kernel Density Estimation (High Risk = &gt; 2 SD)</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Prostitution offenses</td>
<td>Address (point)</td>
<td>Kernel Density Estimation (High Risk = &gt; 2 SD)</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Weapons violation offenses</td>
<td>Address (point)</td>
<td>Kernel Density Estimation (High Risk = &gt; 2 SD)</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Other low-level offenses</td>
<td>Address (point)</td>
<td>Kernel Density Estimation (High Risk = &gt; 2 SD)</td>
<td>Partial</td>
</tr>
<tr>
<td></td>
<td>School disciplinary violations (in- and out-of school suspension)</td>
<td>School District</td>
<td>High Risk = &gt; 1 SD</td>
<td>No</td>
</tr>
<tr>
<td>Risky Places</td>
<td>Cash centered businesses (pawn shops, firearms dealers, tobacco supply vendors, and non-depository lenders)</td>
<td>Address (point)</td>
<td>2 block radius (1350’ buffer)</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>On-site alcohol and adult establishments</td>
<td>Address (point)</td>
<td>2 block radius (1350’ buffer)</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Off-site alcohol</td>
<td>Address (point)</td>
<td>2 block radius (1350’ buffer)</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Hotels/Motels</td>
<td>Address (point)</td>
<td>2 block radius (1350’ buffer)</td>
<td>No</td>
</tr>
<tr>
<td>Socioeconomic Conditions</td>
<td>Males between 15 and 25 years of age</td>
<td>Census Tract</td>
<td>High Risk = &gt; 1 SD</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Non-White Population (racial heterogeneity)</td>
<td>Census Tract</td>
<td>Between 45-55%</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Hispanic Population (ethnic heterogeneity)</td>
<td>Census Tract</td>
<td>High Risk = &gt; 1 SD</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Percent below poverty line</td>
<td>Census Tract</td>
<td>High Risk = &gt; 1 SD</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Unemployment</td>
<td>Census Tract</td>
<td>High Risk = &gt; 1 SD</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Lower than High School Education</td>
<td>Census Tract</td>
<td>High Risk = &gt; 1 SD</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Single Parent Household</td>
<td>Census Tract</td>
<td>High Risk = &gt; 1 SD</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>SNAP</td>
<td>Census Tract</td>
<td>High Risk = &gt; 2 SD</td>
<td>No</td>
</tr>
<tr>
<td>Area Economic Health</td>
<td>Delta in Median Wage</td>
<td>Census Tract</td>
<td>High Risk = Negative Value</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Delta in Number of Employees</td>
<td>Census Tract</td>
<td>High Risk = Negative Value</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Delta in Property Value</td>
<td>Parcel</td>
<td>High Risk = Lost more than 30% value</td>
<td>No</td>
</tr>
</tbody>
</table>
Physical and social disorder. Consistent with the tenets of Broken Windows Theory (BWT), the presence of physical and social disorder should be associated with increased risk of crime. Disorder is thought to promote a sense of disinterest in the well-being of the community diminishing social control and escalating crime and delinquency (Wilson & Kelling, 1982). Untended disorder leads to a cycle of decline and the potential for crime to develop (Skogan, 1990). In addition, parcels with code violations such as overgrown vegetation or improper maintenance may obscure clandestine criminal behavior, consistent with Routine Activities Theory (RAT) and Situational Crime Prevention (SCP). Two measures of disorder were identified for inclusion in the present study: code compliance violations and foreclosures.

Code compliance. Code compliance violations are included in this study both as sources of physical disorder and social and potential places vulnerable to victimization. While many studies exist that empirically evaluating disorder-oriented policing, or broken windows policing, and examining the relationship between disorder and fear of crime, studies examining the relationship between disorder and crime outcomes are limited. In a study of many forms of physical disorder in Seattle, Washington, Yang (2009) found that crime and disorder tend to concentrate at the same place. Kennedy, et al. (2016) found code calls for service related to street light outages and abandoned vehicles to be significant predictors of violent crime.

The DeKalb County Code Compliance Department is responsible for identifying violations of property maintenance, sign posting, and zoning ordinances. The department provided records of all code compliance violations from 2010 through 2015 along with the parcel identifier, reporting date, and offense type. Violations included a range of physical and social disorder issues and code violations such as improperly maintained vegetation, excessive
noise, invalid permitting, and improper sign placement/maintenance. These data were used for the present study.

The present study utilizes a single disorder predictor variable that incorporates several types of physical and social disorder based on code compliance violations. The intent is not to identify the influence of specific types of disorder, but the general condition of disorder. The parcel identifier for each code compliance violation incident was matched to the parcel shapefile to generate a list of all parcels. Between three to five percent of code enforcement violations per year could not be identified because of improperly recorded parcel identifiers. Each parcel was then coded as either possessing one or more code violations in a given year (1) or not possessing a code violation in that year (0).

Foreclosures. Foreclosures are included in the present study as a proxy measure for vacant properties. Foreclosures occur at a property when the property holder fails to meet the requirements of the mortgage contract and the property lender (usually a bank or government agency such as the United States Department of Housing and Urban Development) retakes control of the property and evicts the inhabitants. Upon eviction, regular maintenance of the property becomes the property of the lender, which would not be expected to frequent the property. Per the DeKalb County Code Enforcement Division,

Improperly maintained and unsecure vacant properties can become a hazard to the health and safety of anyone who may come on or near the property and can adversely affect the aesthetic and economic attributes of communities. Difficulties often arise in locating the person responsible for maintenance of
foreclosed properties. DeKalb County finds that there is a substantial need directly related to the public health, safety and welfare to comprehensively address these concerns through the adoption of the Foreclosure Registry Ordinance. (DeKalb County Georgia, n.d.)

Without inhabitants present, the property may not be properly maintained with the potential for physical disorder, consistent with BWT. In addition, vacant properties are unlikely to have adequate oversight, a contributor to crime in RAT.

Several studies have demonstrated a connection between area foreclosures and area violent crime problems. In a study of foreclosures in Pittsburgh, Pennsylvania, Cui & Walsh (2015) found that while foreclosures themselves were not associated with increased crime, vacant foreclosed properties were associated with a 19% increase in violent crime. Immergluck & Smith (2006) found a stronger link in their examination of foreclosures in Chicago, Illinois. The found a small but significant relation between increased foreclosures and an increase in violent crime. Specifically, a 2.8% increase in foreclosures was associated with an increase of 6.7% in violent crime (Immergluck & Smith, 2006). Kennedy et al. (2016) found foreclosures to be a significant predictor of assaults in their RTM of Chicago.

Foreclosure data were provided by the DeKalb County Code Enforcement Division. In DeKalb County, the owner of the property is required to report the property to the DeKalb County Foreclosure Registry within 120 days of foreclosure (DeKalb County Georgia, n.d.). Upon foreclosure, the property generally becomes vacant until sold. The data provided included the address and parcel identifier of foreclosed properties as well as the date of foreclosure. Based
on this information, a shapefile was generated that identifies those properties (parcels) that were foreclosed on within a given year. Each property registered for foreclosure for a given year was coded as a yes (1), while each property not in foreclosure was coded as a no (0). A potential limitation of this data is that it was not possible to identify the duration of which the property was in foreclosure status or if the property remained vacant while in foreclosure. However, the available data still meet the intent of the measure. Less than one percent of data were missing for this variable.

Criminal elements. The work of Sherman, Gartin, and Beurgen (1989) empirically supports the notion that crime tends to concentrate in certain places. While this clustering has been used to inform policing practice in the past (see Braga, Papachristos, & Hureau, 2014), RTM does not utilize the outcome crime of interest as a predictor in assessing risk. Instead, RTM utilizes low-level crime, known offenders, and crime markets (e.g., drug markets, gang territories) to identify the context in which more serious crime occurs. The present study identifies several potential indicators of criminal elements to assess their impact on crime risk.

Probation and parole supervision address. The residence of individuals currently under either probation or parole supervision were included in this study to examine the spatial effects of these individuals with a high propensity to reoffend. According to a study of prisoners released in 2005, “49.7% had either a parole or probation violation or an arrest for a new offense within 3 years that led to imprisonment” (Durose, Cooper, & Snyder, 2014). The presence of individuals under probation or parole supervision is likely to increase the risk for crime in the surrounding area.
Only one RTM study of violent crime has included either probation or parole as a risk factor. Kennedy, Caplan, and Piza (2011) included the presence of parolees as a risk factor in their RTM model but found that it was not a significant predictor of shootings in Newark, New Jersey. Despite this finding, it is important to further consider the potential influence of probation and parole supervised individuals given the strong connection between that status and recidivism.

Data for individuals under probation and parole supervision were provided by the Georgia State Board for Pardons and Paroles. Separate files were provided for both probation and parole. The data in each file included the individuals’ addresses and the years from 2010 and 2014 in which they were under supervision. Because of improperly recorded home addresses, between three and seven percent of data were missing for each year. Addresses were geocoded, and initially a one-half mile buffer was added to each to compensate for the average walking distance of the supervised individual (Yang & Diez-Roux, 2012). However, after reviewing the resulting maps, it was evident that the resulting layer covered almost all of Unincorporated DeKalb County with substantial overlap. As such, the density of individuals under probation and parole supervision was calculated using kernel density estimation. As summarized by Caplan (2010), kernel density estimation is a means of measuring density that weighs those entities that are very close to each other more strongly than those that are near the edge of the search radius. To remain consistent with the study cell size, a cell size of 338’ was used. A search size of one half mile was used in accordance with the average individual walking distance suggested by Yang and Diex-Roux (2012). Consistent with the suggestion of Caplan and Kennedy (2011) those cells with a value of greater than 2 standard deviations were
determined to be high risk and all cells were subsequently coded as either high risk (1) or not high risk (0). This resulted in two risk map layers (one for probation and one for parole) for each year between 2010 and 2014.

*Narcotics offenses.* Narcotics offenses were included as a potential indicator of drug markets. Studies have demonstrated a strong relationship between the presence of drug markets and the occurrence of violent crime (Martinez, Rosenfeld, & Mares, 2008; Reuter, 2009). The relationship may be the result of the presence of crime-prone individuals, competition for market control, or a sense that police are not interested in or not able to address crime in general. Prior RTM evaluations found the presence of drug arrests or drug markets to be significant risk factors for violent crime (Kennedy, Caplan, & Piza, 2011; Caplan, Kennedy, & Miller, 2011) The presence of narcotics offenses, particularly in clusters, may be a potential indicator of the presence of drug markets. Where this occurs, the risk of predatory violent crime may be higher.

Data for all crime types, including narcotics offenses, were provided by the DeKalb County Police Department. The data included all criminal offenses reported to the DeKalb County Police Department along with the date of the offense, address of the offense, and the offense type (UCR code and descriptor). As noted in the description of the dependent variable, approximately 3.7 percent of offenses could not be located because of improperly coded address data. Narcotics offenses were geocoded to identify addresses. Kernel density estimation was used to identify locations where drug crimes cluster to approximate the presence of drug markets. A 338’ cell size and a 1350’ search area were used. This approach was used to improve the measurement of drug crime as a risk factor by focusing on areas
where drugs appear to be a concentrated issue rather than areas where drug offenses may be a rare or one-off occurrence. Those areas with a value more than 2 standard deviations from the mean value were classified as high risk (1) and other areas were coded as not high risk (0). These risk map layers were generated for each year from 2010 through 2014.

Prostitution offenses. Prostitution offenses were included as an indicator of another potential illicit crime market that may underlie violent crime. As with narcotics offenses, the presence of prostitution may be associated with violence through targeting of prostitutes or johns (Farley & Barkan, 2008; Valera, Sawyer & Schiraldi, 2000), market competition, or a general sense that police are unable or uninterested to address crime problems. Visible prostitution can also be a form of social disorder, consistent with BWT (Sampson & Raudenbush., 2004; Wilson & Kelling, 1982). Where prostitution offenses are present, particularly in clusters, the risk of violent crime may be higher. It should be noted that this is a contribution of the present study as prostitution was not included as a risk factor of violent crime in the RTM evaluations reviewed for the present study.

Prostitution offenses were included with the data provided by the DeKalb County Police Department. Addresses were geocoded and kernel density estimation was used to identify areas with high density of prostitution offenses in the same process used for narcotics offenses.

Weapons violation offenses. Weapons offenses are defined by the UCR as “[t]he violation of laws or ordinances prohibiting the manufacture, sale, purchase, transportation, possession, concealment, or use of firearms, cutting instruments, explosives, incendiary devices, or other deadly weapons” (Federal Bureau of Investigation, n.d.). Because weapons are likely to be used in predatory violent crimes, they are included as a risk factor in the present
study. This is another advancement of the present study as weapon violations have not been used in prior RTM evaluations of violent crime. Weapons violation offenses were included with the data provided by the DeKalb County Police Department. Addresses were geocoded and kernel density estimation was used to identify areas with high density of weapons violation offenses in the same process used for narcotics offenses.

*Other low-level offenses.* This study utilizes a measure of “other low-level offenses” that may be related to predatory violent crime. These offenses – counterfeiting, criminal trespass, damage to property, forgery, fraud, peeping tom, shoplifting, simple assault/battery – indicate minor forms of crime may be indicators of or precursors to more serious violent crime. Consistent with BWT, the presence of these variables, particularly in high concentrations, may indicate that there is a lack of formal and informal social control, paving the way for more serious crime. Prior RTM evaluations have used crimes related to the outcome of interest such as the inclusion of gun robberies as a predictor of shootings (Kennedy, Caplan, & Piza, 2011). However, none of the studied identified for review included low-level offenses (other than narcotics offenses). The measure used in the present study seeks to include minor offenses that may not generally come to the attention of the public or be considered serious crime problems. The low-level offenses identified above were included with the data provided by the DeKalb County Police Department. Addresses were geocoded and kernel density estimation was used to identify areas with high density of other low-level offenses in the same process used for narcotics offenses.

*School Discipline.* School discipline was included in the present study as a proxy measure for juvenile delinquency. School discipline was measured as the number of serious disciplinary
actions (in school suspensions and out of school suspensions) in each high school district per year. Because many of the youths in the typical high school age range (13-18) also fall within the typical age-crime curve (National Institute of Justice, 2014), misbehavior in school may be an indicator for the propensity for crime outside of school, particularly when students are not under supervision before and after school.

Data for school disciplinary behavior were provided by the DeKalb County School District. Data included the count of several more serious school disciplinary actions taken in each school including in-school suspension, out-of-school suspension, expulsion, assignment of an alternative education program, and referral to the juvenile justice system. The present study focuses on in-school and out-of-school suspension as these disciplinary actions are serious, but do not necessarily remove the student from the school permanently. Expulsions, referral to alternative education and referral to the juvenile justice system were omitted as those may require the student to leave the present school district.

Disciplinary data were provided by academic school year. For example, one of the five files included data from August 2009 through June 2010, consistent with the academic year. For this study, this 2009/2010 school year was applied to the 2010 RTM model. While the more ideal measure would be a match to the calendar year, use of slightly earlier data was deemed acceptable because students will continue to fit the age-crime curve for the subsequent year.

School attendance zone shape files were obtained for each of the years included in the study. There was a significant shift in school attendance zone boundaries between the 2010/2011 and the 2011/2012 school years. As such, two shapefiles were required to ensure that the data were correctly matched to the geographic area. Counts of disciplinary actions
were matched to the corresponding school attendance zone. Those school attendance zones that exceeded one standard deviation were deemed to be high risk (1) and those less than one standard deviation were coded as not high risk (0). One standard deviation was used as there was limited variability in some years for this measure.

**Risky places.** For the purposes of this study, “risky places” refers to the commercial entities with attributes that may contribute to criminal behavior. Because of activities that occur at these locations, patrons and passersby may be more vulnerable to criminal offense with a lack of capable guardianship. The risky places included in this analysis are categorized into four groups: 1) cash-centered businesses, 2) on-site alcohol and adult entertainment establishments, 3) off-site alcohol, and 4) hotels and motels.

Data for the measurement of risky places were provided by the Planning and Sustainability Department for DeKalb County. The data set provided information on all commercial licenses issued by DeKalb County including the name and address of the entity, the NAISC code and description, the license issue date, and the license termination date. The NAISC codes were determined from a standardized list of different business types known as the North American Industry Classification System (NAICS).

Examination of the data provided suggested that many entities had inaccurate NAICS code identifiers. For example, several restaurants were identified interchangeably as full-service restaurants and late-night drinking establishments. To address this issue, the entire database was visually inspected to verify that the relevant entities were included in the correct groups. First, establishments with NAISC codes that appeared to match the measure of interest were reviewed to ensure that only those matching the definition were included in this study.
Second, the entire dataset was reviewed to ensure that establishments of interest were not misclassified. Those that appeared to be misclassified were reclassified and included in the study data set.

Addresses for each establishment of interest were geocoded and mapped. Between one and four percent of addresses were missing each year due to improperly recorded business addresses. Because crime is also likely to occur in the area immediately around the risky place, a buffer equal to two times the median block length (1350’) was included around each entity.

_Cash-centered businesses_. Cash-centered businesses include commercial entities in which cash is expected to be a primary means of transaction, as opposed to credit and debit. Research by Wright et al. (2014) discussed the strong connection between cash economies and street crime. They demonstrate a strong theoretical connection, but suggest that quantitative research examining this connection is limited (Wright, et al., 2014). As such, this study includes commercial entities in the Unincorporated DeKalb County area in which patrons are likely to enter or leave with cash. These include pawn shops, firearms dealers, tobacco supply vendors, and non-depository lenders (e.g., payday loans, car title loans, check cashing services).

_On site alcohol and adult establishments_. This measure includes commercial entities that feature the sale of alcohol and adult or late-night entertainment. Adult entertainment and late-night establishments include strip clubs and other adult-oriented facilities that cater to after-hour audiences. The purpose of including these establishments is to identify locations where individuals may congregate during late night and early morning, when guardianship may be limited.
Several studies point to the presence of violent offenses in areas where there are high densities of on-site alcohol outlets (Gruenewald, Freisthler, Remer, LaScala, Treno, 2006; Speer, Gorman, Labouvie, & Ontkush, 1998; Grubesic & Pridemore, 2011). Alcohol outlets, including both on- and off-site entities, have been included as significant risk factors in many RTM evaluations of violent crime (Kennedy, et al., 2016; Calplan, Kennedy, & Piza, 2013; Drawve, Thomas, & Walker, 2016; Caplan, Kennedy, Miller, 2011). Because there were few entities whose NAICS codes met the intent of this measure and those that did were widely distributed, it was not possible to measure density of on-site alcohol and adult establishments. However, the locations of these entities were included in the present study as the same mechanisms that lead to violence may be present at the individual entities. Consistent with the tenets of RAT, intoxicated individuals may be vulnerable to victimization and the late-night nature of these establishments may suggest limited guardianship.

It was initially intended that this measure would include bars and restaurants that serve alcohol. However, after reviewing the NAICS coding it was not possible from the code to identify with confidence full-service restaurants which served alcohol. Instead of incorporating potential bias from selecting restaurants that may serve alcohol, it was determined that the better course of action would be to include only those that could be confirmed by NAICS code. Therefore, the present study includes only entities in the following categories identified by the NAICS code descriptors (North American Industry Classification System, n.d.): alcoholic beverage drinking places; bars (i.e., drinking places), alcoholic beverage; cocktail lounges; drinking places (i.e., bars, lounges, taverns), alcoholic; lounges, cocktail; nightclubs, alcoholic beverages; taverns (i.e., drinking places).
Off-site alcohol. Off-site alcohol refers to beer, wine, and liquor stores that sell alcohol for off-site consumption. Because alcohol may not be consumed in the immediate area, there were included in the study separately from on-site alcohol establishments. A one-half mile buffer was used for alcohol establishments to demonstrate the average walking distance of individuals (see Yang & Diex-Roux, 2012). This variable cannot measure the effects of alcohol consumption beyond this buffer, but is intended to measure both illicit consumption within walking distance of the establishment and associated effects of the presence of such entities.

Hotels/Motels. Consistent with tenets of RAT, hotels and motels provide a location with relatively little capable guardianship in which individuals can engage in clandestine criminal behavior. In addition, foreign visitors staying at hotels may be more vulnerable as they are unfamiliar with the local area. Hotels and motels are measured in the present study as the location of hotels and motels in Unincorporated DeKalb County. Few studies examine the spatial relationship between hotels/motels and crime. However, a 1975 study by Engstad that compared automobile and bar crimes (violent and disorderly) in areas with and areas without hotels found higher crime rates in areas with hotels (as cited in Eck & Weisburd, 1995). As such, one may expect increased risk of crime at hotels/motels and the immediately surrounding areas.

Socioeconomic conditions. Only two prior RTM studies could be identified that included measures of socioeconomic conditions, though neither used them as primary risk factors (see Piza, et al., 2016; Drawve, Thomas, & Walker, 2016). In accordance with Social Disorganization Theory (SDT), socioeconomic disadvantage is associated with crime, including violent crime, because they are thought to be indicators of and contributors to a lack of informal social
control (Shaw & McKay, 1942; Kornhauser, 1978). This study seeks to identify whether these measures are applicable as risk factors in the prediction of predatory violent crime.

Data for socioeconomic conditions were obtained from the American Community Survey from 2010 through 2014 using the American Fact Finder tool from the United States Census Bureau. The smallest geographic area available for these data across all years of the present study was the Census tract. Data for each of the socioeconomic measures included in this study – males between 15 and 25 years of age, ethnic heterogeneity, racial heterogeneity, unemployment, education, and single parent households – were matched to a tract shapefile for analysis. In addition, data for the use of Supplemental Nutrition Assistance Program (SNAP) was provided in a similar format by the Fiscal Research Center at Georgia State University. The determination criteria for high risk are identified in each of the subsections below.

*Males between 15 and 25 years of age.* This variable is used to measure the presence of individuals with the highest rate of offending: males and individuals near the peak of the age-crime curve. Males have consistently been responsible for a high rate of arrest for violent crime than females (Federal Bureau of Investigation, 2016). In addition, offending propensity peaks between 15 and 19, then begins to decline in the early 20s (National Institute of Justice, 2014). Therefore, individuals at the nexus of these measures, should be most likely to commit crime. This study tests whether this age-sex nexus as a risk factor for violent crime by examining Census tracts with the highest percentage of individuals meeting these criteria. There was not substantial variation across Census tracts for this variable, so those tracts in which the percent of population that were male between the age of 15 and 25 exceeded one standard deviation
were considered to be high risk. High-risk tracts were coded as 1 while not high-risk tracts were coded as 0.

_Hispanic population._ According to SDT and related theories, the presence of racial or ethnic heterogeneity increases the likelihood of crime because those individuals have a difficult time integrating into the community and building shared beliefs necessary for informal social control (Shaw & McKay, 1942; Kornhauser, 1978). Ethnic heterogeneity is measured in the present study as the percentage of the population, by Census tract, that identifies as Hispanic or Latino. This group made up between eight and nine percent of the population of Unincorporated DeKalb County during the span of this study. Because there was limited variation in this measure, those Census tracts with Hispanic/Latino populations more than one standard deviation above the mean were coded as high risk (1) with all other Census tracts coded as not high risk (0).

_Non-white population._ SDT posits that racial heterogeneity has similar effects on crime as ethnic heterogeneity (Shaw & McKay, 1942; Kornhauser, 1978). Racial heterogeneity is measured in the present study as the percentage of the population, by Census tract, that does not identify as “white only” per the American Community Survey. Thus, this variable measures the percentage of the population that is “non-white.” This group makes up between 68 and 70 percent of the population of Unincorporated DeKalb County on average with substantially higher variation (standard deviation of approximately thirty percent). In this case, quantifying high risk for heterogeneity is slightly different than it was for ethnic heterogeneity. High-risk Census tracts were identified as those with non-white populations between 45 to 55 percent of the population. Given the variation and higher representation of non-white population, this
was deemed to be a better measure of racial heterogeneity. High risk was coded as 1, and not high risk was coded as 0.

*Percent below poverty line.* Economic disadvantage is an important factor in SDT and related theories, as individuals that experience economic disadvantage do not have the opportunities and resources to build strong cohesion with the community, thus inhibiting informal social control (Shaw & McKay, 1942; Kornhauser, 1978). The present study includes the percentage of households, by Census tract, whose family income is below the poverty line for a given year. On average in Unincorporated DeKalb County, between eight and eleven percent of households earned below the poverty line over the course of the present study with limited variation (between five and twelve percent). Given the limited variation, those census tracts with percentages of households in poverty exceeding one standard deviation were classified as high risk (1) and all other tracts were identified as not high risk (0).

*Unemployment.* Unemployment among those in the labor force can be an indicator of economic hardship that may be associated with or independent from long-term poverty. The present study examines the percent of population in the labor force, by Census tract, that are unemployed in a given year. This value varied between seven and twelve percent in Unincorporated DeKalb County across the years of the present study. Because there was limited variation in unemployment for some years, those Census tracts with percentages of unemployment greater than one standard deviation were classified as high risk in the present study. High-risk tracts were coded as 1 and not high-risk tracts were coded as 0.

*Lower than high school education.* Individuals without a high school degree have the lowest median weekly earnings of than those at higher levels of educational attainment
(Bureau of Labor Statistics, 2016). In addition, individuals with no high school diploma have a substantially high rate of poverty (Center for Poverty Research, n.d.; DeNacas-Walt & Proctor, 2015). Therefore, the percentage of individuals with less than a high school diploma (or equivalent), in Unincorporated DeKalb County were included in the present study.

Approximately twelve percent of the population in Unincorporated DeKalb County over the age of 25 has less than a high school education. Those census tracts with percentage of individuals with less than high school education greater than two standard deviations above the mean were classified as high risk. High-risk tracts were coded as 1 and not high-risk tracts were coded as 0.

*Single parent household.* In this study, a single-parent household refers specifically to households with a female adult present with children but no male adult in residence. This is a common measure of disadvantage in SDT literature based on the assumption that the single parent has less time and resources available to provide supervision of youths. Approximately nineteen percent of households meet this criterion. Those Census tracts with percent female-headed households with children exceeding one standard deviation were classified as high risk. High-risk tracts were coded as 1 and not high-risk tracts were coded as 0.

*SNAP.* The Supplemental Nutrition Assistance Program (SNAP) is a federally-funded program that provides financial assistance to low-income individuals and families to help with the purchase of food. This variable was measured as the number of households per Census tract in Unincorporated DeKalb County that received SNAP benefits each year. The data provided included the count of households per month that received SNAP benefits per Census tract, assuming that more than five households in the tract received benefits. The latter
stipulation was to protect anonymity, as required by the Fiscal Research Center. The monthly values for within each tract were averaged to get the yearly household SNAP recipient value for the Census tract. High risk was measured as those census tracts in which the number of households receiving SNAP benefits was more than two standard deviations above the mean. High-risk tracts were coded as 1 and not high-risk tracts were coded as 0.

Area economic health. The variables included in this section are a unique contribution to RTM modeling and are rarely represented in criminological research. The intent of these variables is to identify if area economic health, meaning the economic well-being of the place not necessarily its residents, is risk factor for crime. While many studies have examined the effects of crime on local businesses and property values (Pope & Pope, 2011; Lens, & Meltzer, 2016), few consider the effect in the opposite direction. This study utilizes different measures of area economic health as potential risk factors for violent crime. Declining property values and business revenues may be indicators that crime is present and driving the decline as individuals avoid the area to avoid crime. However, failing business and declining property values may also contribute to vacant properties, diminished patronage, and diminished property desirability. Over time, this could develop into disorder and further develop into crime. The measures discussed in the following subsections are included in this study to explore this relationship. In contrast to typical measures of socioeconomic disadvantage, these measures incorporate a wider range of entities, including commercial and industrial entities. These measures also incorporate economic performance of the area itself to measure the impact of individual entering the area, not just residents.
**Delta in median wage.** Declining employee wages may be an indicator that businesses are closing or reducing the number employees, thus there is potential hardship in the economic health of the area. Data were provided by the Fiscal Research Center at Georgia State University based on data from the Bureau of Labor Statistics. These data included the median wage across all businesses within each Census tract in DeKalb County. The focus of this measure is not the absolute value of the area wages, but rather if wages are increasing or decreasing. This delta was calculated by subtracting the previous year from the model year. For example, the delta in the median wage was calculated by subtracting the median wage in 2009 from the median wage in 2010. This was done for each of the five years of the study. Because it is not possible to control for the types of businesses, which may vary substantially in average wage, any negative value was considered to be a risk factor for the purposes of this study. As such, any tract in which the median property value delta was negative was categorized as high risk (1) and any tract in which the median wage was positive or did not change was categorized as not high risk (0).

**Delta in employees.** Declining numbers of employees can also indicate failing businesses, which may contribute to similar disorder issues discussed in the previous section. Data for the total number of employees by Census tract were similarly provided by the Fiscal Research Center and were coded in the same way. Census tracts that had a negative delta in number of employees were categorized as high risk (1) and any tract in which the median wage was positive or did not change was categorized as not high risk (0).

**Delta in property value.** The final measure of area economic health included in this study is the delta in property values. These data were provided by the DeKalb County Property
Appraisal & Assessment Department and included the address and fair market value for all property parcels in DeKalb County. Because of the financial recession that affected the United States from the late 2000s through the first few years of the 2010s, there was a substantial drop in property values throughout DeKalb County, as evident in the data provided for this study. Because this drop was so widespread, high risk focuses only on those parcels that lost more than 30% of their value for each year of the study. The high-risk properties were coded as 1, and the not high-risk properties were coded as 0.

Analysis Plan

The following discussion outlines the analysis plan for the present study. Steps 1 through 6 were completed in preparation for the study and have been outlined in the above methodology and are briefly summarized here. The process for the remaining steps is detailed in the relevant sections. Steps 6 through 10 were completed 5 times to generate an RTM model for each year between 2010 through 2014. These are described below, with results presented in detail in Chapter IV.

Several analysis tools were necessary for this analysis. Data cleaning and analysis were performed in IBM SPSS 23 and Microsoft Excel. Geographic modeling was performed using ESRI ArcGIS 10.1 and Caliper Maptitude 2016. While ArcGIS provides more advanced tools for analysis, Maptitude is more efficient at performing simple functions such as geocoding and data joins.

Step 1: Identify an outcome event (dependent variable) of interest. The outcome variable of interest is predatory violent crime. This measure includes homicide, aggravated assault, and pedestrian robbery.
Step 2: Identify area of study. The area of study is (Unincorporated) DeKalb County, Georgia.

Step 3: Identify a time period for study. The time period for each RTM model is one year (January through December). Tests of statistical validity utilize crime data for subsequent year.

Step 4: Identify risk factors (independent variables) related to outcome of interest. This study examines 24 measures related to disorder, criminal elements, socioeconomic demographics, risky commercial entities, and area economic health. Variables are summarized in Table 3.1, above.

Step 5: Obtain data and maps. Required map shapefiles were obtained and clipped to the extent of Unincorporated DeKalb County. Maps were obtained from the DeKalb County Department of Planning and Sustainability and Census TIGER/Line®.

Step 6: Generate maps for each risk factor and outcome measure. Data for predictor variables were geocoded or matched to the appropriate geographic level shapefile. As identified in Table 3.1, data were recoded to appropriately reflect distance from, density, and risk level as needed. This results in 24 risk map layers that have been converted into raster files with each raster cell having a value of 1, indicating “high risk,” or 0, indicating “not high risk). To illustrate this process, risk map layers are included for all independent variables for 2010 in Appendix A.

Step 7: Refine model to omit non-significant measures. Once all of the potential predictor variables are identified, each was spatially joined to a grid shapefile clipped to the outline of Unincorporated DeKalb County. Grid cells are equal to 338‘x338’ to align with the
raster measure (one-half block by one-half block). The resulting shapefile was then exported to a .csv file for analysis in SPSS. This resulted in a data file with columns for each predictor variable and the associated year’s output variable, with rows assigned to each cell.

Chi-Square tests were performed in SPSS to identify those predictor variables that are significantly associated with the outcome measure (predatory violent crime). Those that are not found to be significantly associated with predatory violent crime are omitted from the RTM model.

To generate the best predictive models in practice, this process should be repeated for each RTM model year. However, the present study seeks to compare RTM models across the five years of the study. For this reason, the variables identified as significant in the 2010 model generation were used across all five years of models so that these models could be compared over time. While this may result in a slight reduction in the quality of prediction, the changing the variables in the model could adversely affect the ability to compare models. As the key focus of this study is examining the viability of RTM over time, the continued use of the same predictors was deemed the appropriate method for variable selection.

**Step 8: Apply variable weighting, as appropriate.** Variable weighting is an optional step in the RTM process that has been included in the present study. This process is intended to allow those variables that are better predictors of the outcome variable to have a stronger effect on the RTM model. In an unweighted model, the presence of each risk factor contributes one unit to the risk value. In a weighted model, a multiplier is added to each risk factor. While this multiplier can be added arbitrarily or omitted, Caplan and Kennedy (2010) suggest determining the spatial weights based on the odds ratio of a logistic regression model. In this
process, a logistic regression is conducted with all significant predictor variables and the outcome crime variable for the same year. The resulting odds ratios are applied as multipliers in the generation of the combined RTM map.

As indicated in the description of Step 7, a key focus of this study is to examine how the models change over time. As such, the spatial weights identified in the 2010 model are used for each of the subsequent models. This may slightly reduce model quality, but was deemed a better option to allow for model comparison.

**Step 9: Combine individual risk maps into risk terrain map.** Using map algebra, values from each of the individual risk layers are combined to generate a consolidated risk layer. The following example formula suggests how the values for each cell are combined to generate a risk value for each cell in the consolidated risk layer.

\[
\text{Cell } x = [\text{Odds Ratio A} \times \text{Predictor A}] + [\text{Odds Ratio B} \times \text{Predictor B}] + [\text{Odds Ratio C} \times \text{Predictor C}] ...
\]

This process is completed using the ArcGIS Map Algebra tool and results in a “heat map” with a color-coded continuum from highest to lowest risk value.

The resulting maps were then reclassified to identify those areas at highest risk of crime. This designation was made based on those areas with risk values that exceed two standard deviations from the mean risk value. This resulted in a high-risk map for each year from 2010 through 2014.

**Step 10: Communicate model findings.** The final stage in the typical RTM model is to translate findings into an actionable format for the intended audience. In the present study,
this process is detailed in Chapter IV – Results. Suggestions for findings will be elaborated upon in Chapter V – Conclusions.

**Step 11: Examine predictive validity of the RTM Models.** This step is a contribution of the current study to those recommended in the RTM process. While several prior RTM evaluations have utilized logistic regression to examine the ability of the model to predict future crime, this has not been included in all studies. Further, model variance is sometimes reported, but is rarely discussed.

Testing model calibration, or reliability, for the RTM model involves using a logistic regression with the RTM model risk values as the predictors and the presence or absence of the outcome crime of interest as the dependent variable. The consolidated risk terrain map generated in Step 9 is split into a shapefile grid similar to that used in Step 7. Outcome crime data is then spatially joined. A binary logistic regression is used to identify how well the risk measure predicts crime. The odds ratio suggests how much a unit of risk increases the likelihood of the outcome crime, and the associated significance test suggests if this relationship is statistically significant. Further, the Nagelkerke $r^2$ is reported to demonstrate the variance in the outcome measure that was identified by the model.

This process is performed for each of the five RTM models conducted for this study. Logistic regressions are conducted by comparing the risk values from the model using the highest risk maps (independent variable) to predatory violent crime data for the subsequent years (dependent variable). The purpose of this step is to examine how well the risk measure increases the likelihood of crime in the next year and to compare if these results are consistent across multiple years.
While these measures of model performance are an important means to assess the quality of a predictive model, it is also important to examine model discrimination. In this context, model discrimination assesses how accurately the assigned risk class matches with the outcomes. This test is widely used in medical and geographic research (see Bennell, Jones, & Melnyk, 2009; Pontius, & Schneider, 2001; Steyerber, et al., 2010), but is limited in its social science applications and has not been applied to RTM.

As an example of this application to the present study, consider the following. RTM models are used to identify a few micro-places that are at the highest risk of crime to allow police and community agencies to target their limited resources in those areas to have the greatest impact. At the same time, police and community agencies may divert their resources from those areas that show a low risk of crime. However, if police target their efforts in the “cells” with the highest risk level but more crime occurs elsewhere or at concentrations lower than expected, then valuable resources are potentially misallocated. Therefore, it is important to consider not just if there is a correlation between the risk value and the outcome on average, but how often that risk value results in a false positive. This is particularly applicable in the case of crime, which is a relatively rare event.

A parallel example in medical science would be a test for a critical disease. The test may be well calibrated in that it identifies most of the individuals that have a high risk of a critical disease. However, if it also incorrectly identifies a high number of individuals that do not develop that disease, then many individuals may undergo unnecessary, costly, and potentially dangerous treatment. As with RTM models and police resource allocation, it is important that
predictive tests accurately discriminate between risk levels. Additional details of the importance of discrimination can be found in an article by Steyerber, et al. (2010).

The preferred means to test model discrimination with a rare event is the use of Receiver Operating Characteristics (ROC), a statistical method used to compare real and observed values. This technique is particularly valuable in its ability to accommodate continuous predicted values and its accommodation of rare event data, typical of the RTM model prediction. The ROC curve plots the true positive rate ($tp$) relative to the false positive rate ($fp$). The true positive rate refers correctly made classifications in which the event (in this case predatory violent crime) is expected and occurs. The false positive rate refers to incorrectly made classifications in which the event is expected but does not occur.

The area between the ROC curve and the chance line is referred to as the area under the curve, or AUC, which indicates the extent to which the proposed model is better at predicting outcomes than random chance. The further the ROC line deviates from the chance line, the better the model is at predicting outcomes. The AUC is simply a geometric calculation of the proportion of the area of the graph that resides below the curved line. While this is impractical to manually calculate, it can be easily calculated with most statistical software packages. The result will be between 0.5, meaning that the model is no better at predicting outcomes than random guessing, and 1, meaning that the model accurately predicts all outcomes. By comparing the AUC for the ROC curve for the RTM model and the AUC for the ROC for any existing policing strategies, this method can be further used to demonstrate predictive capabilities of the RTM model.
Discrimination testing was done in the present study comparing each model with subsequent yearly data using the ROC methodology described above. Results of this analysis are included in Chapter IV.
Chapter IV – Results

This results section begins with Step 7 in the RTM modeling process. The following sections detail findings from each phase of this study. First, the process of variable selection and weighting for the yearly RTM models are discussed. Second, basic variable descriptive statistics are provided for those variables selected for the study. Third, the yearly RTM models are presented and discussion is provided of the visual comparison. Fourth, results of the model analysis to explore explanatory value are presented. Fifth, results of the model discrimination to explore predictive accuracy are presented. Sixth, and finally, a summary of these findings is provided.

Variable Selection and Weighting.

Before generating the risk terrain models, it is important to determine which variables are statistically related to the outcome measure and to identify any necessary variable weighting. Note that each of the variables had been recoded into dichotomous variables as part of Step 6 with a value of 1 meaning high risk and a value of 0 meaning not high risk. Each risk map layer was spatially joined to a map grid of cells measuring 338’x338’. The number of predatory violent crimes for the same year as the predictor variables were also spatially joined to this composite grid. Because these are rare events, this outcome measure was dichotomized as presence (1) or absence (0) within the cell. A more detailed description of this process was provided in Chapter III.
<table>
<thead>
<tr>
<th>Table 4.1 - Bivariate Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code Compliance Violations (1)</td>
</tr>
<tr>
<td>.061 1 0.05 1 0.05 1 0.05 1 0.05 1</td>
</tr>
<tr>
<td>Foreclosures (2)</td>
</tr>
<tr>
<td>.061 1 0.05 1 0.05 1 0.05 1 0.05 1</td>
</tr>
<tr>
<td>Probation Supervision Addresses (3)</td>
</tr>
<tr>
<td>.151 1 0.05 1 0.05 1 0.05 1 0.05 1</td>
</tr>
<tr>
<td>Parole Supervision Addresses (4)</td>
</tr>
<tr>
<td>0.108 0.05 1 0.05 1 0.05 1 0.05 1 1</td>
</tr>
<tr>
<td>Narcotics Offenses (5)</td>
</tr>
<tr>
<td>0.123 0.04 1 0.04 1 0.04 1 0.04 1 1</td>
</tr>
<tr>
<td>Prostitution Offenses (6)</td>
</tr>
<tr>
<td>0.025 0.05 0.05 1 0.05 1 0.05 1 0.05 1</td>
</tr>
<tr>
<td>Weapons Offenses (7)</td>
</tr>
<tr>
<td>0.105 0.05 0.05 1 0.05 1 0.05 1 0.05 1</td>
</tr>
<tr>
<td>Other Crime (8)</td>
</tr>
<tr>
<td>0.081 0.04 0.04 1 0.04 1 0.04 1 0.04 1</td>
</tr>
<tr>
<td>Schools Disciplinary Violations (9)</td>
</tr>
<tr>
<td>0.021 0.00 0.00 1 0.00 1 0.00 1 0.00 1</td>
</tr>
<tr>
<td>Cash Center Business (10)</td>
</tr>
<tr>
<td>0.051 0.01 0.01 1 0.01 1 0.01 1 0.01 1</td>
</tr>
<tr>
<td>On-Site Alcohol and Adult Establishments (11)</td>
</tr>
<tr>
<td>-0.010 -0.005 -0.005 -0.005 -0.005 -0.005 -0.005 -0.005 -0.005 -0.005</td>
</tr>
<tr>
<td>Off-Site Alcohol (12)</td>
</tr>
<tr>
<td>0.022 0.000 0.000 1 0.000 1 0.000 1 0.000 1</td>
</tr>
<tr>
<td>Hotels/Motels (13)</td>
</tr>
<tr>
<td>0.005 0.014 0.014 1 0.014 1 0.014 1 0.014 1</td>
</tr>
<tr>
<td>Males Between 15 and 25 Years of Age (14)</td>
</tr>
<tr>
<td>-0.011 -0.005 -0.005 -0.005 -0.005 -0.005 -0.005 -0.005 -0.005 -0.005</td>
</tr>
<tr>
<td>Racial Heterogeneity (15)</td>
</tr>
<tr>
<td>-0.037 -0.015 -0.015 -0.015 -0.015 -0.015 -0.015 -0.015 -0.015 -0.015</td>
</tr>
<tr>
<td>Ethnic Heterogeneity (16)</td>
</tr>
<tr>
<td>-0.022 -0.001 -0.001 -0.001 -0.001 -0.001 -0.001 -0.001 -0.001 -0.001</td>
</tr>
<tr>
<td>Unemployment (17)</td>
</tr>
<tr>
<td>0.052 0.021 0.021 0.021 0.021 0.021 0.021 0.021 0.021 0.021</td>
</tr>
<tr>
<td>Lower Than High School Education (18)</td>
</tr>
<tr>
<td>-0.022 -0.003 -0.003 -0.003 -0.003 -0.003 -0.003 -0.003 -0.003 -0.003</td>
</tr>
<tr>
<td>Single Parent Household (19)</td>
</tr>
<tr>
<td>0.081 0.022 0.022 0.022 0.022 0.022 0.022 0.022 0.022 0.022</td>
</tr>
<tr>
<td>Percent Below Poverty Line (20)</td>
</tr>
<tr>
<td>0.034 0.015 0.015 0.015 0.015 0.015 0.015 0.015 0.015 0.015</td>
</tr>
<tr>
<td>SNAP (21)</td>
</tr>
<tr>
<td>-0.019 -0.012 -0.012 -0.012 -0.012 -0.012 -0.012 -0.012 -0.012 -0.012</td>
</tr>
<tr>
<td>Delta in Median Wage (22)</td>
</tr>
<tr>
<td>0.012 0.017 0.017 0.017 0.017 0.017 0.017 0.017 0.017 0.017</td>
</tr>
<tr>
<td>Delta in Number of Employees (23)</td>
</tr>
<tr>
<td>0.052 0.031 0.031 0.031 0.031 0.031 0.031 0.031 0.031 0.031</td>
</tr>
<tr>
<td>Delta in Property Values (24)</td>
</tr>
<tr>
<td>0.086 0.030 0.030 0.030 0.030 0.030 0.030 0.030 0.030 0.030</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).**

*Correlation is significant at the 0.05 level (2-tailed).*
Bivariate correlations were examined comparing each of the predictor variables to the outcome measure of predatory violent crime. The results of these bivariate correlations are reported in Table 4.1. Because of the number of significant bivariate relationships, variance inflation factors and tolerance values were examined. No issues related to multicollinearity were identified in this analysis.

Of the 24 variables originally selected as potential risk factors, 21 were found to be significantly related to predatory violent crime. Those that were not significantly related included foreclosures, racial heterogeneity, and percent of population with less than a high school degree.

Foreclosures were examined both as an indicator of economic disadvantage and as a potential proxy for vacant properties. The lack of a significant relationship with predatory violent crime may be due to a limited time in foreclosure status, continued maintenance of the property, invalid use of foreclosure as a proxy for vacancy, or simply because foreclosed properties are not related to predatory violent crime in the study area. Future research may consider using a true measure of vacant status or tenure in foreclosure status to better explore this variable as a potential risk factor.

It is interesting to find that racial heterogeneity was not a significant predictor of predatory violent crime as it is strongly represented in social disorganization theory (SDT) literature. A potential reason for this finding is the high proportion of non-white population in Unincorporated DeKalb County. Further research could consider measuring racial heterogeneity as a concentration of non-white population rather than true heterogeneity (proximity to 50% white and non-white).
It is possible that the lack of significant relationship between percent of population with less than a high school degree and predatory violent crime is due to the measurement of this variable. Educational attainment is included as a measure of socioeconomic disadvantage in accordance with SDT. This is measured as the percentage of individuals over the age of 25 that have less than a high school degree. However, if some percentage of these individuals are not in the labor market (e.g. full-time parent), the relationship between education and disadvantage may not be strong enough to affect crime outcomes. It may be valuable to further explore the relationship between education and disadvantage in this context.

While these variables were not significant predictors of predatory violent crime risk, 21 of the variables analyzed were. These correlations are summarized in Table 4.1.
While most of the correlations were in the direction expected, it is interesting to note that SNAP was negatively associated with predatory violent crime. This indicates that the presence of a high number of households receiving SNAP is associated with a decrease in the risk of predatory violent crime. This is counter to the initial intent of the measure, which was the inclusion of SNAP as an indicator of economic disadvantage. In contrast, it appears that SNAP may be a protective factor. Further investigation is needed to determine how SNAP is
related to the RTM model, particularly in its potential as a protective factor. As such, it is
omitted from the present analysis.

Per RTM methodology, a logistic regression was then performed to identify variable
weights. These weights are based on the odds-ratio outputs from the logistic regression with
predatory violent crime as the dependent variable and all of the significant predictors identified
above, excluding SNAP, as covariates. The spatial weights identified in Table 4.2 were applied to
all of the RTM models generated for this study. Composite RTM maps were generated for each
study year using map algebra with the variable weights being used as multipliers. Note that in
traditional RTM methodology, spatial weight calculations should be conducted for each RTM
model year. However, as comparison of the model’s predictive accuracy across years is a key
aspect of the present study, the same weights were used across all years to avoid confounding
those findings.
Table 4.3 – Variable Weighting

<table>
<thead>
<tr>
<th>Variable</th>
<th>PVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code compliance violations</td>
<td>1.485</td>
</tr>
<tr>
<td>Probation supervision address</td>
<td>1.719</td>
</tr>
<tr>
<td>Parole supervision address</td>
<td>1.315</td>
</tr>
<tr>
<td>Narcotics offenses</td>
<td>1.745</td>
</tr>
<tr>
<td>Prostitution offenses</td>
<td>1.045</td>
</tr>
<tr>
<td>Weapons violation offenses</td>
<td>1.615</td>
</tr>
<tr>
<td>Other low-level offenses</td>
<td>3.004</td>
</tr>
<tr>
<td>School disciplinary violations</td>
<td>0.923</td>
</tr>
<tr>
<td>Cash centered businesses</td>
<td>1.558</td>
</tr>
<tr>
<td>On-site alcohol and adult establishments</td>
<td>1.600</td>
</tr>
<tr>
<td>Off-site alcohol</td>
<td>1.364</td>
</tr>
<tr>
<td>Hotels/Motels</td>
<td>1.026</td>
</tr>
<tr>
<td>Males between 15 and 25 years of age</td>
<td>0.969</td>
</tr>
<tr>
<td>Hispanic Population (ethnic heterogeneity)</td>
<td>1.875</td>
</tr>
<tr>
<td>Unemployment</td>
<td>1.165</td>
</tr>
<tr>
<td>Single Parent Household</td>
<td>1.325</td>
</tr>
<tr>
<td>Percent below poverty line</td>
<td>1.432</td>
</tr>
<tr>
<td>Delta in Median Wage</td>
<td>1.372</td>
</tr>
<tr>
<td>Delta in Number of Employees</td>
<td>1.234</td>
</tr>
<tr>
<td>Delta in Property Value</td>
<td>1.496</td>
</tr>
</tbody>
</table>

The presence of other low-level offenses (counterfeiting, criminal trespass, damage to property, forgery, fraud, gambling, peeping tom, shoplifting, simple assault/battery) was the strongest predictor variable in the logistic regression model. Ethnic heterogeneity, density of narcotics offenses, and density of individuals under probation supervision were also strong predictors. It is interesting to note that some of the variables that were significantly and positively related to the dependent variable in the bivariate correlations showed a negative correlation with the dependent variable in the logistic regression. These variables were percent of males between 15 and 25 years of age and number of school disciplinary violations. Variance inflation factor and tolerance values were assessed for these variables, but no indication of
multicollinearity was present. Thus, these values may suggest suppression, mediation, or similar effects. Future studies are needed to explore these relationships.

The weights from this logistic regression model were then applied to the RTM model. Specifically, the composite risk maps for each year were based on the following formula:

\[
\text{Composite Risk Value} = (\text{code compliance violations} \times 1.458) + (\text{probation supervision address} \times 1.719) + \ldots + (\text{delta in property value} \times 1.496)
\]

**Descriptive Statistics.**

Descriptive statistics for each of the variables included in the study are included in Table 4.3. These values were constructed prior to dichotomizing variables and are intended to give a summary of the extent to which each risk factor is present in Unincorporated DeKalb County.
Once risk maps were generated for each variable, cells were identified as either high risk or not high risk based on the criteria identified in Table 3.1. A further explanation of this recoding process is provided in Chapter III. The percentage of cells identified as high risk for each variable are summarized in Table 4.4.

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code compliance (percentage parcels in which present)</td>
<td>0.7</td>
<td>0.5</td>
<td>0.7</td>
<td>0.7</td>
<td>1</td>
</tr>
<tr>
<td>Probation supervision address (number)</td>
<td>850</td>
<td>892</td>
<td>959</td>
<td>978</td>
<td>1033</td>
</tr>
<tr>
<td>Parole supervision address (number)</td>
<td>272</td>
<td>343</td>
<td>496</td>
<td>566</td>
<td>649</td>
</tr>
<tr>
<td>Narcotics offenses (number)</td>
<td>2428</td>
<td>2609</td>
<td>2761</td>
<td>2229</td>
<td>1731</td>
</tr>
<tr>
<td>Prostitution offenses (number)</td>
<td>59</td>
<td>108</td>
<td>137</td>
<td>131</td>
<td>116</td>
</tr>
<tr>
<td>Weapons violation offenses (number)</td>
<td>616</td>
<td>108</td>
<td>595</td>
<td>454</td>
<td>496</td>
</tr>
<tr>
<td>Other low-level offenses (number)</td>
<td>11011</td>
<td>11032</td>
<td>11497</td>
<td>11428</td>
<td>12292</td>
</tr>
<tr>
<td>School discipline (number)</td>
<td>10367</td>
<td>10003</td>
<td>9817</td>
<td>10037</td>
<td>10126</td>
</tr>
<tr>
<td>Cash centered businesses (number)</td>
<td>39</td>
<td>39</td>
<td>40</td>
<td>41</td>
<td>42</td>
</tr>
<tr>
<td>On-site alcohol and adult establishments (number)</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Off-site alcohol establishments (number)</td>
<td>4</td>
<td>32</td>
<td>48</td>
<td>67</td>
<td>69</td>
</tr>
<tr>
<td>Hotels/Motels (number)</td>
<td>45</td>
<td>45</td>
<td>43</td>
<td>43</td>
<td>36</td>
</tr>
<tr>
<td>Males between 15 and 25 years of age (percent population)</td>
<td>15.1</td>
<td>15.1</td>
<td>14.8</td>
<td>12.7</td>
<td>14.3</td>
</tr>
<tr>
<td>Hispanic Population (percent population)</td>
<td>8.6</td>
<td>8.9</td>
<td>8.8</td>
<td>8.7</td>
<td>8.2</td>
</tr>
<tr>
<td>Unemployment (percent population)</td>
<td>12.5</td>
<td>13.7</td>
<td>14.5</td>
<td>15.4</td>
<td>14.4</td>
</tr>
<tr>
<td>Single Parent Household (percent population)</td>
<td>19.1</td>
<td>19.5</td>
<td>19.5</td>
<td>19.2</td>
<td>19.17</td>
</tr>
<tr>
<td>Poverty (percent population)</td>
<td>13.4</td>
<td>14.8</td>
<td>16.1</td>
<td>10.6</td>
<td>16.7</td>
</tr>
<tr>
<td>Delta in Median Wage (mean of delta)</td>
<td>675.5</td>
<td>1089.9</td>
<td>704.3</td>
<td>-421.7</td>
<td>734.8</td>
</tr>
<tr>
<td>Delta in Number of Employees (mean of delta)</td>
<td>-146.6</td>
<td>24.9</td>
<td>-50.6</td>
<td>-244.7</td>
<td>197.8</td>
</tr>
<tr>
<td>Delta in Property Value (mean of delta by percent)</td>
<td>-1.9</td>
<td>-27.7</td>
<td>-16.2</td>
<td>1.9</td>
<td>33.7</td>
</tr>
</tbody>
</table>
### Table 4.5 – Percent Raster Grid Identified as High-Risk

<table>
<thead>
<tr>
<th>Category</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code compliance</td>
<td>12.9</td>
<td>34.8</td>
<td>37.6</td>
<td>39.6</td>
<td>8.0</td>
</tr>
<tr>
<td>Probation supervision address</td>
<td>10.6</td>
<td>9.6</td>
<td>38.6</td>
<td>14.5</td>
<td>21.3</td>
</tr>
<tr>
<td>Parole supervision address</td>
<td>23.4</td>
<td>13.4</td>
<td>23.8</td>
<td>16.4</td>
<td>17.2</td>
</tr>
<tr>
<td>Narcotics offenses</td>
<td>8.8</td>
<td>7.2</td>
<td>7.3</td>
<td>6.8</td>
<td>6.9</td>
</tr>
<tr>
<td>Prostitution offenses</td>
<td>2.0</td>
<td>1.6</td>
<td>2.2</td>
<td>2.1</td>
<td>2.2</td>
</tr>
<tr>
<td>Weapons violation offenses</td>
<td>8.1</td>
<td>2.2</td>
<td>5.5</td>
<td>4.9</td>
<td>5.7</td>
</tr>
<tr>
<td>Other low-level offenses</td>
<td>6.9</td>
<td>4.2</td>
<td>4.5</td>
<td>4.0</td>
<td>5.0</td>
</tr>
<tr>
<td>School discipline</td>
<td>27.6</td>
<td>27.5</td>
<td>39.6</td>
<td>34.6</td>
<td>17.9</td>
</tr>
<tr>
<td>Cash centered businesses</td>
<td>4.9</td>
<td>4.9</td>
<td>6.9</td>
<td>7.4</td>
<td>7.6</td>
</tr>
<tr>
<td>On-site alcohol and adult establishments</td>
<td>0.6</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>1.1</td>
</tr>
<tr>
<td>Off-site alcohol establishments</td>
<td>1.9</td>
<td>15</td>
<td>19.7</td>
<td>24.3</td>
<td>24.9</td>
</tr>
<tr>
<td>Hotels/Motels</td>
<td>5.2</td>
<td>5.3</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
</tr>
<tr>
<td>Males between 15 and 25 years of age</td>
<td>7.8</td>
<td>7.6</td>
<td>8.0</td>
<td>18.1</td>
<td>4.6</td>
</tr>
<tr>
<td>Hispanic Population</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
<td>1.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Unemployment</td>
<td>12.6</td>
<td>12.8</td>
<td>20.2</td>
<td>18.3</td>
<td>18.1</td>
</tr>
<tr>
<td>Single Parent Household</td>
<td>28.2</td>
<td>21.0</td>
<td>15.8</td>
<td>16.3</td>
<td>27.7</td>
</tr>
<tr>
<td>Poverty</td>
<td>9.1</td>
<td>9.1</td>
<td>6.7</td>
<td>1.9</td>
<td>2.6</td>
</tr>
<tr>
<td>Delta in Median Wage</td>
<td>32.7</td>
<td>22.5</td>
<td>35.2</td>
<td>76.1</td>
<td>25.0</td>
</tr>
<tr>
<td>Delta in Number of Employees</td>
<td>20.2</td>
<td>58.7</td>
<td>61.6</td>
<td>66.0</td>
<td>20.2</td>
</tr>
<tr>
<td>Delta in Property Value</td>
<td>8.5</td>
<td>46.0</td>
<td>41.7</td>
<td>15.4</td>
<td>6.5</td>
</tr>
</tbody>
</table>

The key observation from Table 4.4 is that there is some fluctuation in the presence of high-risk areas, particularly in the middle years of this analysis. Based on this, higher risk values or a wider geographic span of higher risk values are expected between 2011 and 2013 than in 2010 or 2014.

**RTM Model Comparison.**

A key objective of this analysis is to provide a comparison of RTM models over time to examine if models remain relatively stable or change dramatically over time. This is the first step in determining if RTM modeling is suitable for application to policing by observing how
quickly the targets move. If risky areas are frequently shifting, RTM models must be quickly assessed and reacted to by police. In this case, social and economic interventions may not be particularly valuable as the underlying risk factors are rapidly shifting. In contrast, if risky areas are relatively stable, they may be more amenable to long-term intervention strategies.

RTM Models from 2010 through 2014 are compared in Figure 4.1. Upon visual inspection, it appears that the highest risk areas, indicated by darker red colors, remain relatively stable across all five years of the study. This suggests that risk factors are remaining concentrated in the same areas. Similarly, the areas of lowest risk, indicated by darker green colors, also remain relatively stable. This suggests that targeting the underlying risk factors related to crime may be feasible as it is possible to determine target areas that are stable over time.
Figure 4.1 – Risk Terrain Modeling Heat Maps
It is also important to note that there does appear to be an increase in light green and yellow areas, particularly between 2011 and 2013. It is likely that this broad increased risk level is associated with those risk factors related to property values, unemployment, and business performance. These factors may coincide with the economic recession, which had broad impacts nationwide. However, the highest risk areas appear to remain relatively consistent across all five years of the study suggesting that targeting of highest risk areas may remain a viable strategy.

Figure 4.2 further explores the stability of predictions over time with a comparison of only those areas determined to be high risk. High risk was identified as those raster cells with a risk value greater than two standard deviations above the mean. Yearly RTM models were reclassified to generate new maps with only those areas identified as high risk coded as 1 and all other not high-risk areas coded as 0. This generates a map of dichotomous data that highlights only high risk areas. Between 4.4 and 5.4 percent of raster cells were identified as high risk across each year of the study. These maps are then examined side-by-side. This comparison further suggests that, at least upon visual inspection, high-risk areas remain relatively stable over time. The red areas, indicating high-risk cells, remain concentrated in the central area along the path of Highway 10.
Figure 4.2 – High-Risk Areas
To examine the relationship more closely, the crosstabs were generated to compare how consistent high-risk areas were across years of the study. The outputs presented in Table 4.5 present the percentage of cells identified as high risk that remained consistent across years of the study.

Table 4.6 – Comparison of High-Risk Cell Identification (Percentage)

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>31.3</td>
<td>29.2</td>
<td>29.8</td>
<td>29.4</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>32.3</td>
<td>32.7</td>
<td>32.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>31.2</td>
<td>31.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td>32.7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results in Table 4.5 suggest that approximately 30 percent of high-risk cells remain stable from year to year. These cells that remain stable are likely to be more amenable to intervention to address risk factors than areas where risk levels fluctuate year to year. It is particularly interesting that this similarity remains consistent as much as four years from the initial RTM model. When compared to all four subsequent years of data, approximately 30 percent of cells remained high risk.

Because two-thirds of the areas identified as high risk show fluctuation, the feasibility of this technique for practical application may be limited. Police or community agencies may be chasing a moving target in interventions aimed at high-risk areas. Further, as much as two-thirds of their efforts may be misplaced. A potential means to address this issue would be to isolate those areas that remain stable and target interventions only at those locations. This would be associated with only approximately 1.5% of the total area (or approximately 2.5 square miles) of Unincorporated DeKalb County, which may be further split into different locations. This may make intervention practically and economically infeasible.
In sum, it appears that risk predictions, particularly for high-risk areas, are relatively stable over time. While only 29-33 percent remained the same, visual inspection of the high-risk maps suggests that fluctuation seems to occur in the same general area. However, further analysis is needed to determine the utility of this technique given such a specific and limited area. In addition, the models thus far have only addressed risk. It is important to identify how well this risk predicts crime to understand the validity of RTM as a tool for crime prediction and prevention.

**Model Evaluation.**

The previous section explored stability in risk predictions across the years of this study. This section explores how well the risk models predict future crime. The RTM process is only valuable if the identified high-risk areas correspond with future crime. For police to rely upon such models for resource allocation to high-risk areas, the areas must demonstrate higher levels of crime than non-high-risk areas. This examination begins with a visual assessment of yearly crime data overlaid on the high-risk maps. In this study, each high-risk RTM model was overlaid with predatory violent crime data for the subsequent year. These maps are illustrated in Figure 4.3.
Figure 4.3 – High-Risk Areas with PVC Overlay
It is difficult to ascertain visually, but it does appear that there are more predatory violent crime events at or near areas designated as high risk. However, many events still occur outside of these areas.

To examine this issue further, logistic regression analyses were conducted for each risk terrain model and each subsequent year of predatory violent crime data. These models use a dichotomous value for risk level as the independent variable and a dichotomous value for presence or absence of predatory violent crime as the dependent variable. High-risk maps are used instead of the full risk map as police and community agencies are more likely to allocate additional resources only to those areas at highest risk. The logistic regression model allows for the interpretation of the likelihood of a predatory violent crime occurring based on being in an area identified as high risk.

To conduct this and subsequent analyses, the high-risk maps (Figure 4.2) for each year were spatially joined to a grid with cells with side length equal to 338’x338’ to correspond to the raster cell size. Thus, each cell is assigned a value of 1 if in a high-risk area or 0 if not in a high-risk area. This was completed for each year from 2010 through 2014. The predatory violent crime data for each year is then spatially joined to this file. The predatory violent crime variables for each year, which appear in this file as a count variable, are then dichotomized to indicate the presence (1) or absence (0) of predatory violent crime. This was completed for each year from 2011 through 2015. The results of this analysis are summarized in Table 4.6.
Table 4.7 – Logistic Regression Analysis for High-Risk RTM Models

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b(SE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>2.412 (.070)</td>
<td>2.155 (.069)</td>
<td>2.192 (.070)</td>
<td>2.133 (.071)</td>
<td>2.097 (.066)</td>
<td></td>
</tr>
<tr>
<td>OR</td>
<td>11.159</td>
<td>8.631</td>
<td>8.957</td>
<td>8.44</td>
<td>8.138</td>
<td></td>
</tr>
<tr>
<td>Nage.</td>
<td>0.116</td>
<td>0.088</td>
<td>0.092</td>
<td>0.085</td>
<td>0.084</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>2.185 (.068)</td>
<td>2.251 (.068)</td>
<td>2.190 (.069)</td>
<td>2.002 (.065)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR</td>
<td>8.891</td>
<td>9.495</td>
<td>8.939</td>
<td>7.407</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nage.</td>
<td>0.096</td>
<td>0.104</td>
<td>0.096</td>
<td>0.079</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>2.183 (.070)</td>
<td>2.070 (.072)</td>
<td>2.004 (.067)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR</td>
<td>8.87</td>
<td>7.925</td>
<td>7.418</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nage.</td>
<td>0.089</td>
<td>0.077</td>
<td>0.073</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>2.179 (.070)</td>
<td>2.058 (.066)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR</td>
<td>8.84</td>
<td>7.829</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nage.</td>
<td>0.092</td>
<td>0.082</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>2.197 (.064)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR</td>
<td>9.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nage.</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: All models were significant (p < .001)

b = log-odds
SE = standard error
OR = odds ratio
Nage. = Nagelkerke $r^2$

In each logistic regression performed in this phase of the analysis, the identification of an area as high risk was significantly associated with the presence of predatory violent crime in subsequent years. High-risk areas had between seven and twelve times higher odds of experiencing a predatory violent crime in subsequent years than areas identified as not high risk. This suggests that the high-risk RTM models constructed in this study were successful in identifying places at a higher risk of predatory violent crime. Based on these findings, it is
reasonable that police or community agencies should allocate more resources to these high-risk areas.

Examining changes in the odds ratios over time, it appears that the predictive capabilities of the RTM models are diminished over time. Note that, with one exception, the odds ratio generally decreases over time. This suggests that these high-risk RTM models are better at predicting risk of predatory violent crime in the year(s) immediately following, but are less effective as time goes on. This finding suggests that RTM models may be better for short-term planning than long-term planning. This is not unexpected as social conditions may change and police interventions may occur between the time of the RTM model development and the time of the crime data comparison that are not controlled in the logistic regression. However, it is important to remember that the odds ratios do remain high and somewhat similar over time. This suggests that while there may be some fall-off in the value of the models over time, they do remain a potentially useful tool for long-term planning.

It is also important to consider how well the models identify variation in the predatory violent crime outcomes. RTM methodology suggests the use of Nagelkerke’s $r^2$ for this purpose. Nagelkerke’s $r^2$ is a pseudo $r^2$ statistic that attempts to represent the explained variance in a logistic regression model. Caplan, Kennedy, & Piza (2012) utilize this value for logistic regression in the same manner as would be used in an OLS regression. Future research into RTM models should certainly consider better measures of rare event model calibration and model fit to better understand how well the model performs.

While this method has some limitations, the use of Nagelkerke’s $r^2$ values as identified by Caplan, Kennedy, and Piza (2012) can provide insight into these high-risk RTM models. Note
that the Nagelkerke’s $r^2$ values range from 0.073 to 0.116. This suggests that the high-risk RTM models explain between 7.3 and 11.6 percent of the variance in the presence of predatory violent crime. This is a relatively low value that indicates that the RTM models are not particularly effective at explaining where crime occurs. The overwhelming majority of variation is explained by factors not considered in these models. Because predatory violent crimes are rare, localized events (4.4 to 5.4 percent of cells across years) while risk factors are more stable and span larger geographic areas, this finding is not unexpected. However, further research would certainly be valuable to identify better risk factors and measures to improve the explanatory value of these models.

**Model Discrimination.**

As detailed in Chapter III, model discrimination provides a means to assess how well the predictive model “diagnoses” crime. If police are intended to concentrate resources in those areas identified as high risk, it is important to identify how accurate these predictions are. The present study utilizes receiver operating characteristics (ROC) curves and assessment of the area under the curve (AUC) to assess model discrimination. The results of this analysis are presented in Table 4.7.
Table 4.8 – AUC Values for RTM High-Risk Models

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Predatory Violent Crime</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>0.563</td>
<td>0.557</td>
<td>0.558</td>
<td>0.553</td>
<td>0.561</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>0.555</td>
<td>0.557</td>
<td>0.552</td>
<td>0.554</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>0.557</td>
<td>0.550</td>
<td>0.556</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>0.553</td>
<td>0.557</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>0.661</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.7 displays the AUC values for each RTM model compared to the presence (1) or absence (0) of predatory violent crime in the subsequent years. The AUC values range from 0.550 to 0.661. AUC values generally range between 0.5, meaning that predictions are not better at identifying the outcome measure than random chance, and 1, meaning that the predictors perfectly identify occurrence of an event. The results of this analysis can be interpreted to mean that the high-risk RTM models are between 10 and 32 percent better at predicting predatory violent crime than random chance. These values are similar across all years of the study and appear to remain stable over time.

This is an adverse finding for the applicability of RTM as a tool for resource allocation to address predatory violent crime. These findings suggest that the high-risk RTM models are not particularly useful in diagnosing where violent crime will occur. This is consistent with the relatively low Nagelkerke’s $r^2$ values identified in the previous section. These findings are likely the result of the substantial proportion of predatory violent crimes that occurred outside of the identified high-risk areas. Based on these results, the current RTM models would need to be drastically improved to be useful for resource allocation. Alternatively, a broader definition of...
high risk could be considered such as risk values of one standard deviation above the mean.

However, this would result in larger geographic areas for resource allocation. These issues will be discussed further in the final chapter.

**Summary.**

In sum, it appears that the risk of predatory violent crime is substantially and significantly higher in areas identified as high risk, but that the models are not particularly accurate in identifying all areas where crime is likely to occur. This suggests that certain places do have a higher risk of crime, but that the variables used in the current risk terrain models or the methods used to generate them frequently misclassify places as high risk or not high risk.
Chapter V – Conclusions

The present study sought to develop comprehensive RTM models based on RTM methodology and to evaluate those models as a potential tool for application in crime prediction and prevention. Successful implementation of RTM methodology could prove to be an important tool for police resource allocation and addressing embedded risk factors that underlie major crime problems. RTM is a nascent technique, and as evidenced from its limited body of research and the findings from the present study, additional development is needed. With its proactive and prosocial focus, RTM remains an innovative and promising approach for the future of policing. However, several challenges and limitations need to be addressed prior to practical implementation. The following sections summarize key findings from the present study as well as limitations, identify several directions for future research, and discuss potential implications for policy and practice.

Summary and Discussion of Findings.

The following subsections identify and discuss key findings from the present study. Specifically, this section addresses each stage of the RTM process including aspects of variable selection, RTM model development, identification of high-risk areas, and model variance and discrimination.

Variable selection. This study considered 24 potential risk factors for predatory violent crime with the aim of generating a comprehensive RTM model. Measures were included for physical and social disorder, criminal elements, risky places, socioeconomic conditions and areas economic health. Upon examination of bivariate correlations, the majority of the risk factors identified were significantly and positively correlated with predatory violent crime.
These relationships were relatively weak, but the significant correlations made them applicable to the RTM model.

**Physical and social disorder.** Physical and social disorder were included in the analysis both in the context of Broken Windows Theory (BWT) and Routine Activities Theory (RAT). Where signs of physical and social disorder are prevalent, BWT posits that they signal disinterest in the well-being of the community allowing further disorder and, eventually, crime to occur. RAT posits that crime will occur in areas where motivated offenders, suitable targets, and lack of capable guardianship converge, which one may expect where physical conditions obscure visual observation. This application of RAT to the physical environment is reinforced in Situational Crime Prevention (SCP).

One of the two physical and social disorder variables were found the be significantly correlated with predatory violent crime. The presence of code compliance violations (e.g., calls for abandoned vehicles, excessive noise, overgrown vegetation) was associated with the presence of predatory violent crime. The presence of foreclosures, in contrast, was not significantly associated with predatory violent crime. Foreclosures were included as a proxy for vacant or abandoned properties where a lack of capable guardianship may allow crime to flourish. It is possible that foreclosure was not an adequate proxy for vacant property. Perhaps the entity taking possession of the property ensured proper maintenance or the properties did not remain vacant for an extended period of time. This may reflect research by Cui & Walsh (2015) which found that vacant properties were associated with violent crime but foreclosures were not. Because foreclosure was found not to be significantly associated with predatory violent crime, it was omitted from the RTM model.
Criminal elements. All of the measures of criminal elements included in this study were found to be significantly and positively correlated with predatory violent crime. Criminal elements were included as potential measures of offenses that are precursors to or otherwise associated with more serious violent crime. For example, high density of reported narcotics offenses could be indicative of an underlying drug market. Given empirical support for the association between drug markets and violent crime (Martínez, Rosenfeld, & Mares, 2008; Reuter, 2009), it reasonable to hypothesize that potential drug markets may be related to predatory violent crime.

Criminal elements were measured by the presence of other, less serious forms of crime and delinquency that may be associated with predatory violent crime. The strongest correlations were for narcotics offenses, weapons violations, and other low-level offenses (counterfeiting, criminal trespass, damage to property, forgery, fraud, peeping tom, shoplifting, simple assault/battery). Prostitution and school disciplinary violations were also significantly and positively associated with the presence of predatory violent crime, but to a lesser extent. While none of the correlations were particularly strong, they were significant and thus included in the RTM model.

Risky places. Risky places, for the purposes of this study, were identify as commercial entities whose function and purpose may increase the risk of victimization. These included the presence of cash-centered businesses (e.g., pawn shops, payday loans), on-site alcohol and adult establishments, off-site alcohol establishments, and hotels or motels. Consistent with RAT, these places are likely to have suitable targets and a lack of capable guardianship. For
example, individuals leaving a bar may be intoxicated leaving them vulnerable to potential offenders seeking an easy target for street robbery.

The location of each of these places was significantly and positively associated with the presence of predatory violent crime at and in the immediately surrounding area. Though the correlations were relatively weak, all were included in the RTM models because the relationships were statistically significant.

Socioeconomic characteristics. This study provides the unique contribution of including socioeconomic indicator as potential risk factors for the RTM model. Socioeconomic characteristics, largely drawn from typical measures of Social Disorganization Theory (SDT) and related theories, reflect the social and economic characteristics of the area with a particular focus on factors related to disadvantage. Where factors associated with disadvantage are concentrated, the ability to develop and enforce social norms may be diminished, thus allowing crime to flourish.

Measures of socioeconomic characteristics in the present study included males between 15 and 25 years of age (a high-level offending group), racial heterogeneity, ethnic heterogeneity, unemployment, individuals without a high school diploma or equivalent, single-parent (mother) households, households with income below poverty line, and households receiving SNAP (e.g., food stamps). Six of the eight measures were significantly associated with predatory violent crime. Five of these – males between 15 and 25 years of age, ethnic heterogeneity, unemployment, single-parent households, and households with income below poverty line – were included as risk factors in the RTM models.
Three socioeconomic risk factors were excluded from the model. Racial heterogeneity and individuals with less than a high school education were not significantly associated with predatory violent crime. It is difficult to hypothesize why these variables may not have been significantly associated with the dependent variable. Further, it was interesting to find that the percent of households receiving SNAP was not positively associated with the dependent variable. It was posited that households receiving SNAP would be a potential indicator of economic disadvantage as households have to demonstrate several forms of economic hardship to qualify for this public assistance. However, it does not appear that this measure worked in the proposed direction. In fact, the use of SNAP may be acting as a protective factor by alleviating some of the economic hardship. This is a thought-provoking finding that warrants further evaluation in future research. As this is a complex issue, this variable was omitted from the present study to maintain focus on the performance of the RTM modeling process.

*Area economic health.* The inclusion of area economic health measures is another unique contribution of this study. Area economic health measures are intended to examine the influence of larger economic conditions at the property value and commercial level that may be unrelated to socioeconomic measures of residents. This study included places where median wage, number of employees, and property value were decreasing. Where wages and employees are decreasing, businesses may be failing and potentially contributing to problems such as abandoned properties and decreased tax revenue. Areas losing more than thirty percent of their property value also suggest decreasing economic stability in the area. All three area economic health measures were significantly and positively associated with the presence of predatory violent crime and were included in the RTM models.
**Development of RTM models.** RTM models were developed in this study for each year from 2010 through 2014 using the RTM methodology identified by Caplan and Kennedy (2011, 2012) and an array of social, economic, crime, and contextual variables. In contrast to other predictive analytic techniques such as hot spot analysis, the present study sought to identify where future crime will occur based on the concentration of area risk factors.

**Model variable weighting.** Variables in the RTM model were weighted in accordance with RTM methodology. This process allows variables that are stronger predictors of the outcome variable to have a stronger impact on risk values than those variables that are less strongly related. A logistic regression with each of the variables to be used in the RTM models, identified above, and a dichotomous (presence/absence) predatory violent crime dependent variable was used to identify this variable weighting. For a full discussion of this process, reference Chapter III.

The results of this logistic regression revealed that the risk factors included in the RTM models have different effects on risk-value outcomes. The strongest predictor of predatory violent crime was other low-level offenses, in which areas with the presence of other low-level offenses were three times more likely to experience predatory violent crime than places without low-level offenses. The presence of individuals under probation supervision, narcotics offenses, cash-centered businesses, and on-site alcohol and adult establishments were also among stronger predictors of predatory violent crime.

In the logistic regression, two variables that were significantly and positively associated with predatory violent crime at the bivariate level were found to be negatively associated with predatory violent crime at the multivariate level. These included school disciplinary violations.
and males between 15 and 25 years of age. These effects of these variables warrant further
analysis, though initial reviews of variance inflation factors and tolerance did not suggest an
issue with multicollinearity. There may be moderation or suppression effects in place that
diminish the influence of these variables. Despite this question, these variables were included
as potential risk factors with the appropriate weighting as suggested by RTM methodology.

Visual examination of RTM models. Visual inspection suggested that the process was
successful in developing “heat maps” indicating the continuum of crime risk based on the
variables selected. Crime risk appeared to remain relatively stable, clustering near the central
and the northern parts of Unincorporated DeKalb. While total risk values fluctuated over time,
the highest and lowest risk areas appeared to remain relatively stable. These maps suggest that
crime risk, as calculated with the variables used in this study, appears to concentrate in certain
areas indicating that crime risk factors are entrenched. This is a valuable finding as it may assist
in the targeting of risk factors in micro-areas to address crime problems.

Identifying high-risk areas. The next step in this analysis was the identification of high-
risk areas based on the composite RTM models. Those areas with a risk value exceeding 2
standard deviations from the mean risk value were considered to be high risk. While this is
somewhat of an arbitrary determination, it is consistent with the methodology proposed by

Upon mapping those areas identified as high risk, it was evident that crime risk
concentrates in certain areas and appears to be relatively stable over time. Approximately five
percent of cells were identified to be high risk across each year of the study. Visual inspection
confirmed that high risk areas appeared to stay in the same general area throughout the study.
An examination of cells identified as high risk indicated that approximately thirty percent remained consistent from year to year. These findings further confirm that RTM may be useful in identifying micro-areas at highest risk of crime in order to target underlying risk factors for intervention. This would allow concentration of intervention in the 1.5 percent of areas with the consistently highest risk of crime.

**Predictive risk assessment.** In the first step of predictive risk assessment, predatory violent crime data were overlaid on maps indicating high-risk areas for the preceding year. Based on visual inspection, it appears that crime does cluster around those areas identified as high risk. However, many instances of violent crime also appear to be distributed outside of these areas.

Logistic regression was used to compare the identification of high-risk areas in the model to the presence or absence of predatory violent crime in subsequent years. The logistic regression analyses showed that the odds that a predatory violent crime would occur in a given area (raster cell) were between 8.8 and 11.2 times higher in those areas identified as high risk compared to those areas that were not high risk. These significant findings suggest that areas classified as high risk are indeed more likely to experience crime than other areas.

The odds ratio fell only slightly as many as five years from the RTM model year. As such, RTM models appear to be valuable in predicting risk for years beyond their initial construction. However, they are most applicable in the time immediately following their development. For practical use, RTM models may be more effective if they are generated and responded to frequently making them better for shorter term planning.
Model variance. While the initial phases of these analyses showed promising results for RTM modeling in crime prediction, examination of model variance indicated potential limitations. Utilizing Caplan, Kennedy, and Piza’s (2012) methodology for interpreting Nagelkerke’s $r^2$, the high-risk RTM models generated in this study explain between 7 and 12 percent of the variation in predatory violent crime. This low explained variance suggests that as much as 93 percent of variation in predatory violent crime can be explained by factors other than the areas being identified as high risk. These results are similar to those seen in other RTM model evaluations (Caplan, 2011; Caplan, Kennedy, & Miller, 2011; Caplan, Kennedy, & Piza, 2012).

While this interpretation indicates that the RTM models are not particularly adept in explaining variation violent crime outcomes, they explained variance is not inconsistent with that of other criminological research. In a review of articles published in Criminology between 1968 to 2005, Weisburd and Piquero (2008) found that the explained variance based on $r^2$ for criminology-related articles were on average approximately 39 percent. Further, they found that approximately 25 percent of studies had an explained variance of less than twenty percent (Weisburd & Piquero, 2008). This reflects the exceedingly complex nature of crime and difficulty in including the many factors that influence crime. As such, while the explained variance in the RTM models generated in this study appear low, they may not be inconsistent with other research in this field. Nonetheless, the value of these and other RTM models warrant further development to improve their explained variance through a better understanding of the variables used to construct the models.
**Model discrimination.** This study provides the unique contribution of analyzing RTM model discrimination. Model discrimination is a means to test how often the diagnosis of “high risk” is correct. Model discrimination was measured using the AUC from the ROC analysis. The results suggest that the RTM models generated in this study were only between 10 and 32 percent better than random chance at identifying where high-risk places will occur. As such, the RTM models appear to correctly predict where crime occurs, but also likely have a high false positive rate. This means that there may be a high number of areas that are assigned as high risk, but no crime actually occurs in those locations. Alternatively, there may be a large number of predatory violent crime instances that are occurring outside of the high-risk areas, as visual inspection of the maps suggested. This could be problematic if additional resources and intervention efforts are targeted at areas where risk may not result in crime outcomes. While this result initially appears to be an adverse finding for the applicability of RTM modeling, it is important to consider the context of the analysis. Crime risk factors considered in this analysis, particularly those measured at the tract level or with large buffers, can span many cells resulting in a high-risk geographic area covering many cells. In contrast, crime is an isolated incident occurring in only one cell. Further, because crime is a rare event, it was measured as a dichotomous (presence/absence) variable in the present study. Numerically, more crimes may be occurring in these high-risk areas than in not high-risk areas. This is something that needs to be explored further.

**Summary.** In sum, RTM methodology as applied in the present study appears to be effective in identifying a relationship between risk and predatory violent crime, but important questions remain regarding its effectiveness in pinpointing specific locations where crime will
occur. Strategies targeting high-risk areas, particularly those addressing the underlying risk factors, may be successful in localized crime intervention. However, more research and development is needed in RTM methodology before it is applied in a broader approach to policing if the intent is to replace existing techniques. Many potential directions for future research and theoretical development are discussed in the subsequent sections.

**Study Limitations and Observations.**

Before proceeding to research implications and directions for future research, it is important to address limitations specific to the present study. This study sought to generate comprehensive RTM models based on a wide variety of variables to predict predatory violent crime. While the study was successful in generating these models in accordance with RTM methodology, the low explained variance calls to question the extent to which these models were **comprehensive**. Further, the results of the ROC analyses suggest that these models only provide a slight improvement over random chance. While the risk value and correlation findings were consistent with other RTM evaluations, some limitations of the present study may have contributed to the low explained variance and AUC values. Note that the following identified limitations are specific to the data and approach of the present study, not to RTM in general. Additional discussions of the RTM methodology, including limitations and directions for future research, are provided in subsequent sections.

**Variable identification.** In preparation for this study, a wide range of potential risk factors were considered and a wide range of data sources were identified. While these variables were theoretically, empirically, and/or rationally informed, it is not reasonable to assume that all possible risk factors were identified. Crime is an exceedingly complex
phenomenon with a nearly infinite number of potential risk factors. This study identified 24 that were reasonable and had accessible data. Other data that were sought but were not in a format that was applicable to this study included gang residences, residences of recent jail releases, alcohol licenses, and census data for small geographic units. Additional research is needed to identify and assess the many other potential risk factors to improve the predictive capabilities of these and other RTM models.

**Data quality and precision.** Several issues with data quality and precision were encountered during this study. Improperly recorded addresses for businesses and crime incidents resulted in missing data and posed challenges in identification of correct addresses. Based on the NAICS codes provided with the business-license data, many businesses appeared to be misclassified requiring interpretation by the researcher to identify the correct entities to include in the analysis. Foreclosure data was used as a proxy for vacancies because vacancy data was not available for the years covered in the current study. School disciplinary data was measured in academic year rather than calendar year, thus requiring the data be applied to the model year following the first several months of data. While none of these issues invalidate the use of these measures, a certain amount of measurement error must be assumed in the outcome of the study.

Several of the measures used in this study were at larger geographic levels than desired for this type of analysis. RTM is a micro-area approach to analyzing crime. Unfortunately, several measures of socioeconomic characteristics and area economic characteristics were only available at the Census tract level geography. This may spread the influence of risk over a wider area than is appropriate. For example, if only a few neighborhoods were highly disadvantaged,
they may artificially inflate the level of risk across the tract. While ideal measures would be chosen at a smaller geographic level (e.g., Census block or block group), two considerations allowed for the inclusion of these variables at the larger geographic unit. First, these variables included only a few of the many variables included in this study and only those Census tracts with the highest values were identified as high risk. It is unlikely that a few neighborhoods could drive the entire Census tract into the highest risk category. Second, variable weighting was used to increase the level of influence of variables in the composite risk model such that those variables that were not strong predictors were diminished in their influence. This should mitigate some of the overestimation of risk that may occur by including the entire Census tract in the model. Nonetheless, these are important considerations in the interpretation of model findings.

**Model weighting.** An issue was encountered in the model weighting process that warrants further evaluation. Two of the variables that were initially positively associated with predatory violent crime – school disciplinary violations and males between 15 and 25 years of age – became negatively associated with predatory violent crime in the multivariate model. While this did not appear to be an issue with multicollinearity, it raises questions regarding potential interaction effects that warrant further consideration. As many of the variables included in an RTM model are related, it is important to consider potential interaction effects (e.g., moderation, mediation, suppression).

**Summary.** Challenges encountered during the study present important considerations in the interpretation of results, but also present an opportunity to pursue better measures and techniques in future research. Building upon these challenges and other observations from the
development and testing of RTM models in this study, the following sections identify areas for future research and improvement of the RTM process.

**Directions for Future Research and Theoretical Development.**

This study explored many aspects of RTM modeling process and identified several directions for future research as well as theoretical and methodological development. RTM is a relatively new technique in crime analysis and prediction, and as such, there are many considerations to further evaluate and improve upon this process. Further, issues identified in the testing of model validity and discrimination point to the need for additional development.

**Improved theoretical guidance for variable selection.** RTM is a methodological approach to crime-risk prediction, not a theoretical approach to understanding how or why crime occurs. However, the process could benefit greatly from additional theoretical guidance. The Theory of Risky Places (TRP) posits that some places are at higher risk of crime than other based on spatial factors that increase the threat of or vulnerability to crime (Caplan & Kennedy, 2012). TRP was developed to complement RTM methodology. TRP is a unique and methodologically-driven approach to understanding crime. However, the theory is vague in identifying what constitutes a risk factor and how those elements interact. This ambiguity made it difficult to select risk factors that would be suitable for crime prevention. Two suggestions are proposed to address this limitation.

First, TRP should be further developed to identify key types of risk factors that should be considered in RTM model development. It is not necessary, nor feasible, to identify all possible risk factors, but guidance on key considerations such as those categorized in the present study (e.g. criminal elements, socioeconomic characteristics) could be incorporated to guide future
research. For example, RAT identifies three components – suitable target, motivated offender, and capable guardian (Cohen & Felson, 1979). This allows for theory testing that focuses on specific, measureable concepts. Adding this specificity encourages replicability and uniformity in RTM models to facilitate cross-study comparison in future research. This could involve theory integration bringing together principles from existing criminological and criminal justice theory and/or assessment of previous RTM evaluations to identify patterns in risk factors correlated with outcome measures. In addition, RTM methodology references the influence of potential protective factors, but does elaborate on how they should be included in RTM models.

Second, this process of theory development may benefit from meta-analytic techniques to identify and assess the influence of risk factors on risk model outcomes. A wide range of variables and measures were used across the studies examined in preparation for the current study. This study added to the list of potential risk factors to be considered with the addition of socioeconomic characteristics and area economic health measures. While it is valuable to consider a range of variables that may help to improve explained variance, it becomes difficult to perform cross-study comparison. This is an essential component to improving our understanding of the connection between area risk and crime outcomes and to improve the efficacy of the RTM modeling process. It is perhaps more important to build a strong foundation of key variables from which to build. A meta-analytic or systematic review of existing RTM evaluations could be used to identify consistently used variables and variables with the strongest influence on outcomes. These key variables can then be incorporated into TRP and RTM methodology.
Each RTM model should be adapted to fit the unique context of the environment, crime outcome of interest, and available data. However, additional theoretical guidance, particularly in variable selection and modeling could lead to substantial improvements in the explanatory value of the RTM model. Further, this guidance may improve the propensity for cross-study comparison to further our understanding of nuances in spatial crime research.

**Improved methodological guidance for analysis.** Because the application of RTM is a new technique in crime analysis, the process is evolving and changing to incorporate improvements and new findings. This innovation is valuable to the continued development of the process. However, guidance provided in the texts (Caplan & Kennedy, 2011, 2012) and the online training program offered by the Rutgers Center for Public Security leaves many questions. Some of these questions and issues were encountered in the present study. While RTM has been applied largely in the academic field where those constructing the models are expected to have some statistical expertise, such issues could limit its applicability among practitioners that are relying upon a rigid methodology without the ability to address such unexpected findings. Four important areas that need further guidance are discussed here: variable interaction, variable weighting, criteria for determining risk level, and model evaluation.

The predictor variables selected for the present study, and those used in previous studies, often measure similar concepts, presenting the opportunity for multicollinearity, interaction effects, and similar unexpected relationships between independent variables. While tests of multicollinearity did not indicate an issue in the present study, this verification was performed outside of the guidance of RTM methodology. RTM methodology should be
modified to include a step, perhaps between current Steps 6 and 7, to test for this important consideration. Further, solutions may be offered such as scaling variables or incorporating interaction terms for closely-related measures. The application of this step will be unique to each RTM model, but should be recognized in the RTM methodology as it has important implications for model fit and validation.

An issue was encountered in the variable weighting process of the RTM analyses in this study in which variables that were significantly and positively associated with predatory violent crime at the bivariate level were significantly and negatively associated with predatory violent crime at the multivariate level. RTM methodology identifies the logistic regression process used in this study as the correct means to apply variable weighting, but provides little insight into the reasoning behind this process or how to address issues such as those encountered in the present study. As RTM is an additive approach, perhaps a better approach would be to apply variable weighting based on the bivariate correlations. Other weighting methods should also be considered including controls of variation in measurement (e.g., addresses versus tract data). Better theoretical and methodological guidance may help to address this issue by addressing the justification for the weighting process. Further research is needed to determine the most accurate and useful variable weighting process.

Another aspect of the RTM modeling process that warrants further research and guidance is the criteria for determining risk level. Risk level determinations are made at two steps in the RTM process. First, risk values are assigned to individual risk factors prior to constructing the composite RTM model. Second, the composite risk model can be reclassified into high-risk areas to identify those areas for targeting. RTM methodology provided by Caplan
and Kennedy (2011) suggests a standard of two standard deviations as a cut point for determining “high risk.” However, this did not apply to all variables included in the present study. For example, minimal variation in some risk factors used in this study meant that there were no areas that exceeded two standard deviations from the mean risk value. The identification of risk level should certainly be adapted to meet the needs of the study, but additional research would help to determine available identification options. Criteria such as quantiles or Jenks breaks can also be considered. Future research may seek to determine if these improve the predictive ability of the RTM models.

Finally, additional research is needed to evaluate RTM models. This study sought to improve understanding of the quality of RTM models by examining Nagelkerke’s $r^2$ and AUC characteristics. Only three studies reviewed in preparation for the present study reported Nagelkerke’s $r^2$, and only one discussed its meaning. No prior studies have included an assessment of AUC. If RTM is to become a tool for policing and community intervention in crime problems, it is important to understand how accurately and precisely these models diagnose crime risk. Bayesian Information Criterion (BIC) is a probabilistic model fit technique used in some RTM evaluations for similar purposes, but those studies provide little discussion of its meaning. This technique was considered beyond the scope of the present analysis, but may present another option for model evaluation.

A number of evaluation methods are available to determine how well the RTM technique predicts crime, but more research is needed to determine the best method for evaluation and to compare across studies. For example, Bayesian Information Criterion (BIC) is a probabilistic model fit technique used in some RTM evaluations for similar purposes, but
those studies provide little discussion of its meaning. This technique was considered beyond the scope of the present analysis, but may present another option for model evaluation. If findings are consistent with those found in the present study, RTM models may have limited utility as a diagnostic tool. It is important that RTM models not only recognize significant correlations between risk measures and crime outcomes, but that those predictions are accurate. This is particularly important as police and other agencies may allocate valuable resources based on those predictions. Additional research is needed to better identify how areas labeled as high risk correspond with crime instances and how often these predictions are incorrect.

**Integration of temporal controls.** It is important to remember that correlation does not equal causation. However, predictive modeling assumes a certain amount of causation, particularly in interventions targeted at addressing risk factors associated with future crime. RTM evaluations compare RTM models to crime events in subsequent time periods. However, these studies cannot identify whether risk factors drive future crime or are simply indicators that mechanisms are in place that are driving both the risk factors and crime. It is also possible in this methodology that crime is driving the presence of risk factors or that a reciprocal relationship exists.

While it is not feasible to directly test causation in the relationship between risk factors and crime outcomes, the incorporation of longitudinal analysis can bring research a step closer. Longitudinal techniques are intended to measure correlations over time. This may include time-series design, growth curve analysis, growth mixture modeling, or similar methodologies. Such research methods consider changes in the same set of subjects, or area cells in this case, using
repeated measurement over an extended period. A thorough discussion of these methodologies are provided by Singer and Willett (2003). Future research should consider the incorporation of longitudinal techniques to better understand the potential predictive relationship between risk factors and crime.

Controls for ongoing interventions. In studies conducted over time, such as the present study, unexpected factors can affect findings in different periods. In RTM models, ongoing intervention efforts by police or community organizations may affect crime outcomes in ways that are not measured in the RTM model. For example, many police departments utilize hot spot or near repeat methodologies to “crack down” on high-crime areas with targeted interventions and preventative patrols. Because these techniques are often successful, at least in the short term, they may decrease crime within the area and time of the RTM study. If these actions correspond with areas that would be deemed as high risk, it can diminish the relationship between high-risk areas and crime outcomes in the RTM model. This is difficult using retrospective data collection, but the use of Compstat maps or patrol data may help to control for police action. It may also be beneficial to communicate with local community organizations to identify when and where they are engaging in outreach efforts as these may have similar effects on the RTM model.

Comparison of RTM and other predictive technique discrimination. Many police departments already utilize hot spot, near repeat, Compstat, and Predpol and other predictive techniques for resource allocation. Drawve (2016) performed such a comparison of “spatial and temporal analysis of crime, nearest neighbor hierarchical, kernel density estimation, and risk terrain modeling” techniques (p.1). The study found that RTM was the second best predictor of
crime after kernel density estimation (Drawve, 2016). Drawve’s study focused on robberies as the outcome measure. Further research is needed to compare these analytic techniques for other crime types. Perhaps RTM modeling works well for burglary prediction compared to most other hot spot techniques, but the question remains if it performs as well for other crime types such as homicide or narcotics use.

**Direction for translating into policing intelligence.** Chapter II of this study included a discussion of the importance of translating information into intelligence. RTM evaluations to date have focused on proving if RTM is effective in predicting crime, but have not sought to suggest interventions based on these findings. As such, RTM information is being analyzed, but is not being translated into actionable intelligence. Once high-risk areas are identified, what should police or community organizations do in these areas? What are the main risk factors driving crime outcomes? Should efforts be made to address those risk factors or should police simply allocate more officers to those areas? Answering these questions is an important next step for the present study and future RTM research.

**Implications for Policy and Practice.**

It is clear from the above discussion that more research is needed to better understand the potential utility of RTM as a predictive analytic technique to address crime, particularly if RTM techniques are intended for practical implementation and the replacement of other police resource allocation tools. The results of this study suggest that while concentrations of high-risk areas remain somewhat stable over time, there is also a substantial amount of variation in both long-term high-risk areas and correlation with crime outcomes that are not accounted for in the RTM models. Further, the low explained variance and limited diagnostic capability challenge the
efficiency of using RTM as a predictive analytic tool. While RTM is promising in principle, the evaluation of the models points to a need for additional research to better understand model performance; to improve model performance, if possible; to determine if these models outperform other available techniques such as hot spot analysis; and to assess if the limited improvements are worth the time and resources needed to conduct such analyses. That is not to say that RTM cannot be an effective tool, but that more research is need to evaluate its utility prior to practical implementation.

RTM modeling is a labor-intensive, time-consuming process compared to many existing hot spot analysis techniques. If RTM does not add significant value beyond existing techniques, it may not be worth the investment by police departments. RTM requires advanced software, access to large amounts of data from other entities, and researchers capable of effectively modeling data and interpreting outputs. This may be beyond the resources available to many police departments. Based on the findings from the present study, it does not appear that there is currently enough evidence to recommend the use of RTM as a replacement for current policing techniques. Additional research is needed to compare the diagnostic capabilities and cost-benefit ratio of RTM relative to other hot spot techniques.

However, it is important to remember that RTM models can be beneficial in addressing problems related to crime that are outside the scope of the police role. The variables used as risk factors are often within the purview of other government or community agencies. For example, business entities identified as “risky places” in the present study must obtain business licenses from DeKalb County. The county government may choose to limit or deny future permits on those businesses strongly associated with predatory violent crime. Alternatively,
community organizations such as non-profits, need-based programs, churches, and public works departments may seek to provide financial resources and increased job opportunities in areas where socioeconomic disadvantage are major risk factors for predatory violent crime. RTM analyses may be beneficial in performing non-police interventions to address crime problems. Nonetheless, caution should be taken in practical implementation until model performance is refined.

Related to the present study, code compliance violations, presence of individuals under probation supervision, cash-centered businesses, and ethnic heterogeneity were among the strongest non-crime predictors of predatory violent crime. Organizations within DeKalb County may consider several approaches to address crime problems by targeting these risk factors. Additional resources can be allocated to the DeKalb County Code Compliance Department to better enforce regulations on property maintenance. State probation officers may increase frequency of meetings with those under their supervision residing in high-risk areas. The Planning and Sustainability Department of DeKalb County may opt to limit future licenses for cash-centered businesses. Finally, community organizations can be used to provide support in large Hispanic communities to address underlying issues that may be related to crime. Each of these relationships need to be thoroughly explored to determine the best intervention options, resources available, and potential adverse effects of intervention. Yet such approaches can be used to address entrenched issues associated with crime to have a more lasting impact.

It is also important to remember that caution should be taken in implementing interventions intended to prevent crime. Even efforts to implement prosocial interventions in specific areas can draw attention to the area as being “high crime.” This can adversely affect
property desirability and value, further exacerbating underlying issues. This can be even more problematic with the police “crack down” approach. Ferguson (2012) and Joh (2014) further caution that areas designated as high crime can lead to unconstitutional profiling, discrimination, and violations of the Fourth Amendment. Others have suggested negative outcomes such as increased fear of crime and reduced police legitimacy (Hinkle & Weisburd, 2008; Rosenbaum, 2007). As such, potential adverse consequences of proactive intervention based on predictive modeling should be considered prior to program implementation.

Conclusion.

The results of this study suggest that while RTM is a promising technique for crime prediction at a conceptual level, more research and development is needed to improve its accuracy, precision, and translation into intelligence-led policing. Examination of area risk factors appears to be a somewhat successful method for predicting future crime, but there are important challenges to the validity of the RTM models. While the relationship between risk values and predatory violent crime outcomes were statistically significant, critical issues related to explanatory value and diagnostic capability were identified. Further, additional instruction is needed in the selection, testing, and weighting of variables used in the development of RTM models. These issues, along with improvements in theoretical and methodological guidance, are needed prior to practical implementation. Once this process is more firmly established, the next key step will be translating findings from RTM models into actionable intelligence.

RTM presents a unique opportunity to develop a proactive and prosocial approach to crime prevention. Innovative approaches like RTM are crucial to our continued understanding of crime problems and improvement of criminal justice response. RTM offers a novel and
promising approach to crime analysis and prevention. However, to ensure RTM is consistent with evidence-based practice, further research, development, and analysis of this technique is needed prior to practical implementation.
Appendix A - 2010 High-Risk Independent Variable Maps

- Code Compliance Violations
- Foreclosures
- Probation (Density)
- Parole (Density)
Narcotics (Density)

Prostitution (Density)

Weapons (Density)

Other Low-Level Offenses (Density)
School Discipline (Tract)

Cash Centered Businesses

On Site Alcohol / Adult Ent.

Off-Site Alcohol
Poverty (Tract)  Unemployment (Tract)  

Less Than High School (Tract)  Single Parent Home (Tract)
References


extensions/spatial-analyst/solving-problems/using-the-conceptual-model-to-create-suitability.htm


Vita

Audrey C. Clubb was born in 1984 in Marietta, Georgia. She completed her Bachelor of Science degree in Criminal Justice in 2006 with a concentration in Forensic and Behavioral Science. During and after her undergraduate education, she worked in engineering, specializing in environmental and security designs for nuclear power stations. Audrey obtained her Master of Science degree in Criminal Justice and Criminology at Georgia State University in 2012. Her thesis was entitled “Defending the Castle: Applying Protection Motivation Theory to Explain the Use of Home Guardianship.” She then pursued her Ph.D. in Criminal Justice and Criminology at Georgia State University, graduating in the Spring of 2017.

Audrey’s primary areas of research include geospatial crime modeling, policing innovation, policy and program evaluation, evidence based policy and practice, and predictive analytics. While at Georgia State University, Audrey conducted research on a variety of topics including fear of crime measures in survey research, neighborhood context of crime and bystander intervention, effects of commercial entities on public health and safety outcomes, and geospatial crime modeling for crime prevention. She also facilitated training for the Atlanta Police Leadership Institute and served as an expert witness for DeKalb County, Georgia. Her work has been published in Criminal Justice Studies, Journal of Financial Crime, Criminal Justice Review, and The Oxford Handbook of Criminological Theory. Audrey was the recipient of the Graduate Academic Achievement Award at Georgia State University. She received the Andrew Young School Fellowship at Georgia State University for the duration of her Ph.D. studies.

Email: audrey.c.cclubb@gmail.com