

JANUARY 3, 2018



THE
CENTER
FOR
STATE
AND
LOCAL
FINANCE

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

Audrey C. Clubb
Dean A. Dabney
Joshua C. Hinkle

Table of Contents

Introduction	2
Study Purpose	2
Methodology	
Data and Sample	3
Dependent Variables	6
Independent Variables	6
Analysis	19
Results	22
Risk Terrain Models	24
Logistic Regression Results	28
Model Discrimination	30
Conclusion	31
Study Limitations and Observations	32
Directions for Future Research and Theoretical Development	34
Implications for Policy and Practice	38
Conclusion	39
References	40
About the Authors	44
About the Center for State and Local Finance	45

Introduction

This paper is provided as a final report on research conducted on behalf of the Center for State and Local Finance. This project involved extensive examination of area characteristics to generate and assess a statistical model to aid in the prediction of future crime problems in DeKalb County, GA. The analysis utilized risk terrain modeling (RTM), a geospatial statistical modeling technique that provides a heat map illustrating the relative risk future criminal behavior in a given area. The present study considered a number of local variables such as existing crime patterns, foreclosures, school performance, and economic conditions as potential protective or risk factors for crime. The resulting heat map was used to identify areas where crime is likely to occur in the future to allow for early intervention to deter more serious crime problems from developing. An assessment of model discrimination was also performed to examine how accurately the model predicted crime in future years. The results suggest that the RTM methodology is very promising, but also needs a great deal of future research and refinement to be an effective tool for targeting crime prevention efforts.

Study Purpose

The purpose of the current study was to employ the process of RTM to identify micro-areas within DeKalb County at the highest risk of future crime problems as well as the underlying factors that may be amenable to intervention. This study is intended to demonstrate and evaluate the capabilities of RTM as a potential tool for crime prediction with a focus on assessing the accuracy of the method in predicting crime in subsequent years. Accurate early detection of areas trending towards crime and subsequent early intervention efforts have the potential to stave off more serious crime problems. The present study sought to accomplish the first step in this process by demonstrating the analysis process to identify specific areas that may warrant intervention by police and other government and community agencies.

RTM utilizes geographic information system (GIS) modeling and logistic regression to determine an area's relative risk based on location-specific characteristics. This is a relatively new technique that requires additional evaluation; however, initial studies are promising. RTM was first demonstrated in 2009 based on data from 2007 (Caplan & Kennedy, 2011). While the risk factors considered in the RTMs vary substantially across studies, these models have been found to be statistically significant positive predictors of both violent (Drawve, Thomas, & Walker, 2014; Drawve, Moak, & Berthelot, 2016; Kennedy, Caplan, & Piza, 2011; Kennedy, Caplan, Piza, & Buccine Schraeder, 2015; Caplan, 2011;) and property crimes (Caplan, Kennedy, Barnum, & Piza, 2015; and Moreto, Piza, & Caplan, 2014). Further, studies suggest that RTM provides more precise outcomes in the prediction of gun crimes than traditional hot spots techniques (Kennedy, Caplan, & Piza, 2011; and Drawve, Moak, & Berthelot, 2016). Additional case studies and comparative research are needed to confirm the efficacy of RTM as a crime forecasting tool, but these initial studies demonstrate promising results.

In addition to the effectiveness of RTM in predicting risk of crime outcomes, 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. This study examined the performance of risk terrain models and subsequent targeted interventions in Chicago, IL; Colorado Springs, IL; 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 building 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.

Early indications of successful predictive capabilities of RTM in these studies suggests that it may be a valuable tool for use in crime prevention. As such, the intent of the present study is to demonstrate the capabilities of RTM for DeKalb County, GA.

Methodology

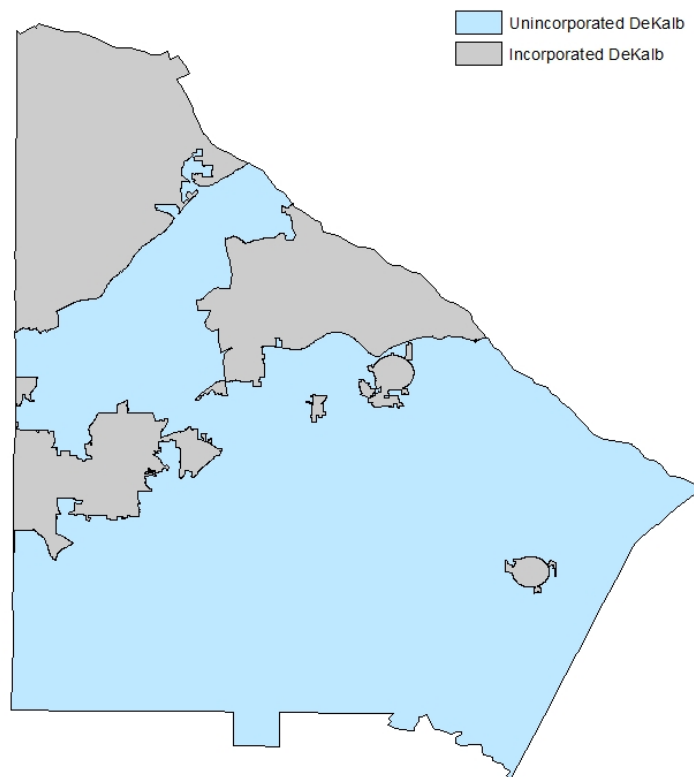
Utilizing RTM methodology, this study generated a risk terrain model across multiple crime types in DeKalb County, GA. The purpose of this phase was to explore the potential explanatory power of RTM for various crime types. By replicating the RTM technique in this new environment, this study sought to provide further evaluation research on both the effectiveness of RTM and the identification of risk and protective factors. A number of risk and protective factors were considered for the analysis, which were subsequently reduced to include only those directly applicable to the crime types for each analysis.

DATA AND SAMPLE

Data for this study were collected from resources available from DeKalb County government and the Fiscal Research Center at Georgia State University. Approximately one-third of the population of DeKalb County lives in Incorporated DeKalb, which consists of those areas that have obtained cityhood (Niesse, 2016). In contrast to Unincorporated DeKalb, many of the local government services provided by the county, including police response, are instead provided by the individual cities. During the data collection process, it became evident that data were not available from DeKalb County entities for these incorporated areas. As such, the scope of this analysis was limited to Unincorporated DeKalb County. Despite this limitation, we believe sufficient data was available for analysis without impacting the integrity of findings. A map showing the incorporated and unincorporated areas of DeKalb County is provided in Figure 1, while Table 1 illustrates demographics from 2010.¹

¹ Several areas of unincorporated DeKalb County have been incorporated into municipalities since then.

Figure 1. Unincorporated Areas of DeKalb County, GA



DeKalb County is located to the northeast of Atlanta, GA, and spans urban, suburban, and rural areas. DeKalb County has a diverse population as summarized in Table 1. RTM models were generated based on 2013 and 2014 for each dependent variable, with variable weightings in both models generated using 2013 baseline data. The 2013 model was compared to 2014 and 2015 outcome data, while the 2014 model was compared to 2015 outcome data, to test their predictive accuracy.

Table 1. Demographic Characteristics of Unincorporated DeKalb County, GA (2010)

Total Population	564,614
AGE	
Population Median	34.9
% Under 19 Years of Age	34.3
SEX	
% Male	47.4
EDUCATION	
High School Diploma or Above	86.9
Bachelor Degree or Above	34.5
INCOME	
Population Mean	52,328
Population Median	48,831
% Below Poverty Line	13.6
% Unemployed	12.5
RACE	
% White	29.5
ETHNICITY	
% Hispanic or Latino	9.0

Based on requirements of confidentiality, many data sets used for this were aggregated from address-level to larger geographic areas such as the block, block group, or census tract. RTM can accommodate these different levels of geographic measurement. The smallest unit available that does not compromise confidentiality or identify specific individuals or addresses were used, unless the data were publicly available (e.g., property value). The unit of analysis for this study was raster cells, which are representative of micro-places. Micro-places were identified by rasterizing the map of DeKalb County. Rasterizing is the process by which a geographic area is divided into a grid in which each cell is of a uniform size. Raster cells must contain sufficient information (e.g., number of crime incidents) to be meaningful for the analysis, but smaller cell size will produce a smoother and more precise outcome map. For current analyses, Raster cells were set to equal to one-half the median block length, which resulted in 41,229 cells of 338'x338'.

Data collection was limited to years 2010 through 2015 in the broader study, with 2013 through 2015 data being used in the current analyses, as these were readily available at most county departments. The following subsections detail the variables used, their origin, and measurement issues encountered to date.

DEPENDENT VARIABLES

Three dependent variables measuring different types of crime were assessed: residential burglary, business robbery and burglary, and predatory violent crime. Burglary refers to “unlawful entry of a structure to commit a felony or theft” (Federal Bureau of Investigation, n.d.). Residential burglary was selected to examine the impact of local factors on crimes targeting houses, apartments, condominiums, and other residential areas. Because of the complex nature of interpersonal relationships on violent crime in the home, the present study limits the scope of crime against the home to property offenses represented by residential burglary. The second dependent variable of interest incorporates both burglary and robbery offenses against commercial entities. Robbery refers to “the taking or attempting to take anything of value from the care, custody, or control of a person or persons by force or threat of force or violence and/or putting the victim in fear.” Because crime against business may be more financially motivated, both robbery and burglary offenses were considered. The third dependent variable, predatory violent crime, includes homicide, aggravated assault, and pedestrian robbery. These offenses are used to examine serious violent crime between individuals. Sexual assault and kidnapping, while violent offenses, are omitted as they often involve a specific offender/victim context.

Each of these dependent variables were measured as the presence or absence of one or more of these offenses within a raster cell in each year. Because the number of these events is relatively small, especially given the use of small rasters, the presence (1) or absence (0) of an offense was dichotomized and logistic regression was used for the analysis. These data were provided by the DeKalb County Police Department and included the following for all reported crimes: incident address (street address), incident date (year), and crime type (UCR code and description). These data were geocoded to include only those occurring in the bounds of Unincorporated DeKalb County.

INDEPENDENT VARIABLES

Multiple independent variables were used to capture contextual factors that could contribute to the presence of these dependent variables. The selection of variables was both theoretically and logically informed. These risk factors include measures of disorder, criminal elements, “risky places,” socioeconomic conditions, and area economic health. The independent variables included in these analyses are summarized in Table 2 and described in detail in the following pages.

Table 2. Independent Variables

VARIABLE	MEASURE	DATA LEVEL	MODIFIERS
PHYSICAL AND SOCIAL DISORDER	Code Compliance Violations	Parcel	1 block radius (675' buffer)
	Foreclosures	Parcel	None
CRIMINAL ELEMENTS	Probation Supervision Address	Address (point)	Kernel Density Estimation (High Risk \geq 2 SD)
	Parole Supervision Address	Address (point)	Kernel Density Estimation (High Risk \geq 2 SD)
	Narcotics Offenses	Address (point)	Kernel Density Estimation (High Risk \geq 2 SD)
	Prostitution Offenses	Address (point)	Kernel Density Estimation (High Risk \geq 2 SD)
	Weapons Violation Offenses	Address (point)	Kernel Density Estimation (High Risk \geq 2 SD)
	Other Low-Level Offenses	Address (point)	Kernel Density Estimation (High Risk \geq 2 SD)
	School Disciplinary Violations	School District	High Risk \geq 1SD
	Sheriff Dispatch Addresses	Address (point)	Kernel Density Estimation (High Risk \geq 2 SD)
	Marshall Dispatch Addresses	Address (point)	Kernel Density Estimation (High Risk \geq 2 SD)
RISKY PLACES	Cash Centered Business	Address (point)	2 block radius (1350' buffer)
	On Site Alcohol and Adult Entertainment	Address (point)	2 block radius (1350' buffer)
	Off-Site Alcohol (Liquor Stores)	Address (point)	2 block radius (1350' buffer)
	Hotels/Motels	Address (point)	2 block radius (1350' buffer)
SOCIOECONOMIC CONDITIONS	Males Between 15 and 25 Years of Age	Census Block Group	High Risk \geq 1SD
	Non-White Population	Census Block Group	Between 45%-55%
	Hispanic Population	Census Block Group	High Risk \geq 1SD
	Percent Below Poverty Line	Census Block Group	High Risk \geq 1SD
	Unemployment	Census Block Group	High Risk \geq 1SD
	Less than High School Education	Census Block Group	High Risk \geq 1SD
	Single Parent Household	Census Block Group	High Risk \geq 1SD
	SNAP	Census Block Group	High Risk \geq 2SD
AREA ECONOMIC HEALTH	Delta in Median Wage	Census Tract	High Risk = Negative Value
	Delta in Number of Employees	Census Tract	High Risk = Negative Value
	Delta in Property Value	Parcel	High Risk = Lost more than 30% value

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) (Cohen & Felson, 1979) and Situational Crime Prevention (SCP) (Clarke, 1995). 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. (2015) 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 (with 2013-2015 data used in current analyses) 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. They 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. (2015) found foreclosures to be a significant predictor of assaults in their RTM of Chicago.

Foreclosure data for the study period 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 Beurger (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 to 2014 (2013 and 2014 data are used in current models) 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 2014 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 and the other offense types discussed below, were provided by the DeKalb County Police Department. The data included all criminal offenses reported to the DeKalb County Police Department from 2010-2015 (2013-2015 were used in the current study) 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 2013 through 2014 in the current study.

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 used in the current models) 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, and 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 studies 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 files included data from August 2012 through June 2013, consistent with the academic year. For this study, this 2012/2013 school year was applied to the 2013 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. 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.

Finally, data on Marshall and Sheriff dispatches in 2013 and 2014 were included as another indicator of criminal and antisocial behavior. This study used data provided by DeKalb County's 911 department to identify addresses to which Sheriff and Marshall personnel were dispatched. Because of the relatively high number of dispatches, kernel density estimation was used to identify areas where Sheriff and Marshall actions were concentrated.

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 during the study period 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; Grubestic & 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., 2015; Caplan, 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 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 (2013 and 2014 data were used for current analyses) 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; Warner, 2014). 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.; DeNavas-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 (Warner, 2014) 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 (with 2013 and 2014 used in current models). 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 from 2010-2015 (with 2013 and 2014 used in the current study). 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 2012 from the median wage in 2013. This was done for each of the 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

This study will follow the process for RTM described in the *Risk Terrain Modeling Manual* (2010) by Joel Caplan and Leslie Kennedy and provided through training with the Rutgers Center for Public Safety (RCPS). These resources provide a detailed process for designing the risk terrain model, weighting data, performing the statistical analysis, and mapping results. 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 above and are briefly summarized here. The process for the remaining steps is detailed in the relevant sections below. Steps 6 through 10 were completed twice to generate an RTM model for 2013 and 2014.

Several 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. The existing RTM utility, RTMdx, was not used to allow for greater control of the analysis processes required for this study. The 11 steps of the process are outlined below.

- **Step 1: Identify an outcome event (dependent variable) of interest.** The outcome variables of interest are a) residential burglary, b) commercial robbery and commercial burglary, and c) predatory violent crime. The predatory violent crime 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 was one year (January through December). Tests of statistical validity utilize crime data for the subsequent year.
- **Step 4: Identify risk factors (independent variables) related to outcome of interest.** This study examined 24 measures related to disorder, criminal elements, socioeconomic demographics, risky commercial entities, and area economic health. Variables are summarized in Table 2 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 resulted 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.”
- **Step 7: Refine model to omit non-significant measures.** Once all of the potential predictor variables were 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 were not found to be significantly associated with predatory violent crime were 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 sought to compare RTM models across the three years of the study. For this reason, the variables identified as significant in the 2013 model generation were used for the 2014 model as well so that these models could be compared over time. While this may result in a slight reduction in the quality of prediction, changing the variables in the model could adversely affect the ability to compare models. As the key focus of this study was 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 was 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 based on the strength of the relationship between that variable and the outcome. 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 variables 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 was to examine how the models change over time. As such, the spatial weights identified in the 2013 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 were combined to generate a consolidated risk layer. The following example formula illustrates how the values for each cell were combined to generate a risk value for each cell in the consolidated risk layer.

$$\text{Cell } x = [\text{Odds Ratio A} * \text{Predictor A}] + [\text{Odds Ratio B} * \text{Predictor B}] + [\text{Odds Ratio C} * \text{Predictor C}] \dots$$

This process was completed using the ArcGIS Map Algebra tool and resulted 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 2013 and 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. For the current project, these include this report, Audrey Clubb's dissertation, a report provided to Dekalb County and other publications in progress including peer-reviewed journal articles.
- **Step 11: Examine predictive validity of the RTM Models.** This step is an added 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 explained is sometimes reported, but is rarely discussed and model discrimination tests have not been applied to RTM in past studies. This contribution is discussed in more detail below.

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 was performed for each of the three RTM models conducted for this report. Logistic regressions were conducted by comparing the risk values from the model using the highest risk maps (independent variable) to data on each of the three outcome variables 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; Steyerberg, 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 Steyerberg, 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 by comparing each model with subsequent yearly data using the ROC methodology described above. Results of this analysis are included in below

RESULTS

Variable Selection and Weighting

The analyses in the present study examined multiple independent variables that may be associated with crime outcomes. To ensure the best fitting model, those predictors that were not significantly related to the outcome measures were dropped prior to generating the risk terrain models. Table 3, Table 4, and Table 5, below show those variables that were included in the model for each dependent variable. This baseline model used 2013 data to develop weights to be used for both the 2013 (to be compared to 2014 and 2015 outcomes) and 2014 models (compared to 2015 outcomes) in order to test the accuracy of the model over time. The listed values refer to the odds ratio of the logistic regression models, which are then applied as weight multipliers in the map algebra to generate the risk terrain models.

Table 3. Variable Weighting for Residential Burglary

Code Compliance Violations	1.713
Probation Supervision Addresses	1.223
Parole Supervision Addresses	1.365
Weapons Offenses	1.307
Foreclosures	1.739
Marshall Dispatch Address	2.413
Sheriff Dispatch Address	1.840
Percent Below Poverty Line	1.371
Delta in Property Value	1.280

The strongest predictors of residential burglary outcomes were addresses to which Marshalls have been dispatched, addresses to which Sheriffs have been dispatched, presence of foreclosures, and presence of code compliance violations. These variables are consistent with persistent property issues, which may lead to some vulnerability of the property.

Table 4. Variable Weighting for Business Robbery and Burglary

Narcotics Offenses	2.172
Prostitution Offenses	1.432
Weapons Offenses	1.451
Other Crime	4.242
Cash Centered Businesses	2.336
Off-Site Alcohol (Liquor Stores)	1.582
Percent Unemployed	1.879
Percent Below Poverty Line	1.598
Percent Non-White	2.260

The strongest predictors of business robbery and burglary were the presence of other minor crime types, presence of narcotics offenses, presence of cash centered businesses, and the presence of racial heterogeneity. The presence of cash centered businesses as one of the strongest predictors is consistent with suitable targets for robbery/burglary.

**Table 5. Variable Weighting for
Predatory Violent Crime**

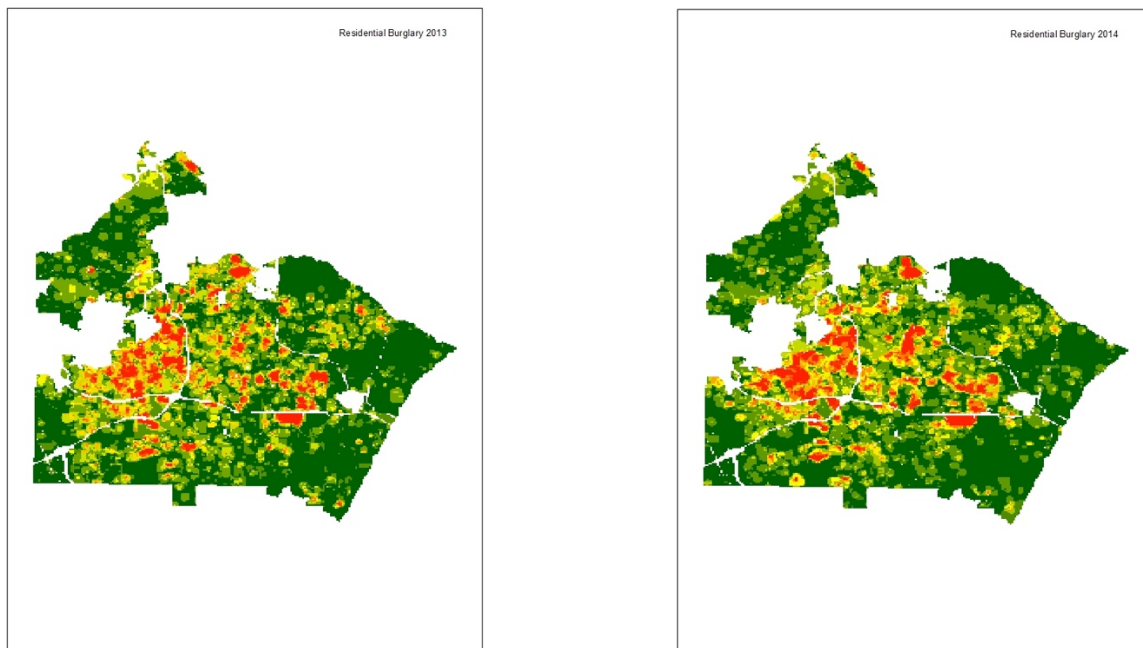
Code Compliance Violations	1.501
Parole Supervision Addresses	1.447
Narcotics Offenses	1.813
Prostitution Offenses	1.376
Weapons Offenses	1.352
Other Crime	2.334
Foreclosures	1.443
Cash Centered Businesses	1.333
Marshall Dispatch Address	2.076
Sheriff Dispatch Address	1.526
Percent Unemployed	1.905
Less than High School Education	1.495
Delta in Median Wages	1.228
Delta in Number of Employees	1.184
Delta in Property Value	1.240

The strongest predictors of predatory violent crime were the presence of other low level crime, addresses to which Marshalls have been dispatched, percent unemployed, and the presence of narcotics offense. These indicators do not necessarily indicate a clear pattern of factors that contribute to a criminogenic environment, but the number of predictors points to the complexity of violent crime.

RISK TERRAIN MODELS

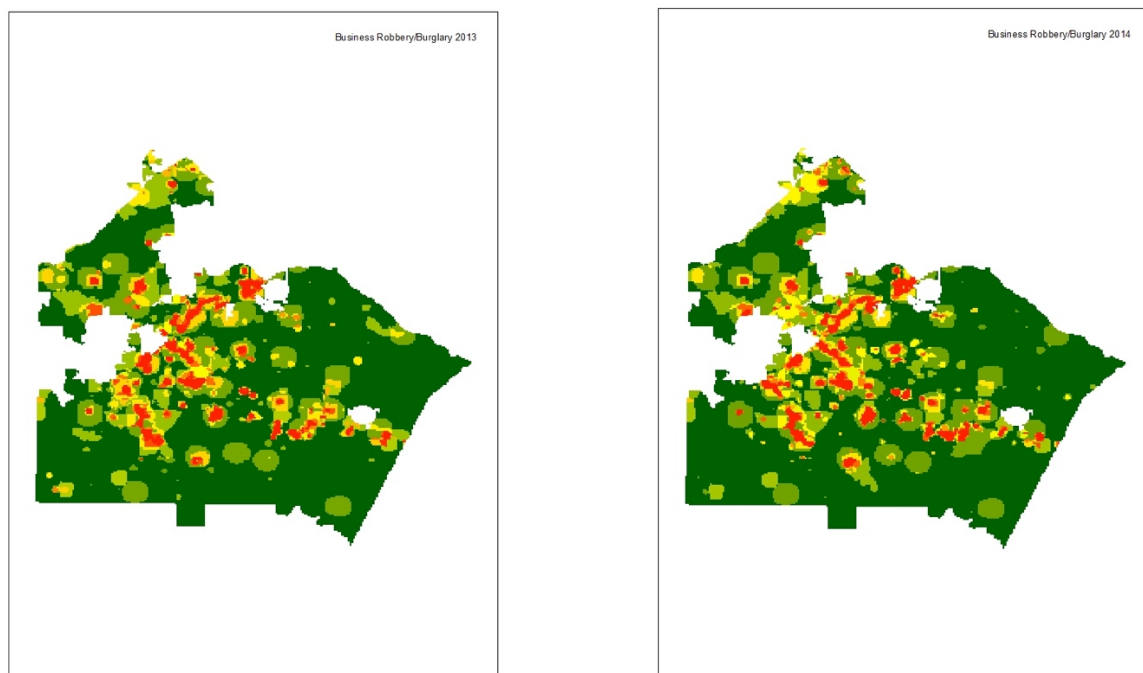
The composite risk terrain models are illustrated in the “heat maps” below. Areas with the highest concentration of risk factors appear in red while areas with the lowest concentration of risk factors are shown in green.

Figure 1. Residential Burglary 2013 and 2014 Risk Terrain Models



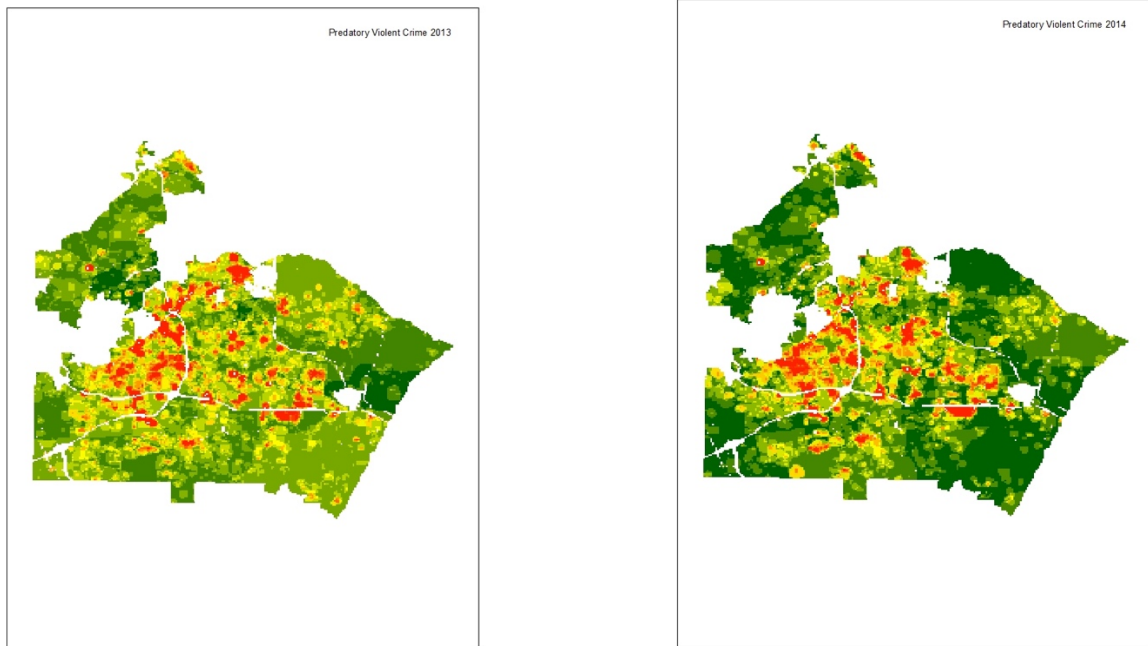
The distribution for residential burglary risk appears to be relatively consistent from 2013 and 2014. The highest concentrations are in the central areas of Unincorporated DeKalb County.

Figure 2. Business Robbery/Burglary 2013 and 2014 Risk Terrain Models



Business robbery and burglary risk factors are similarly concentrated in the central areas of Unincorporated DeKalb County with little change between 2013 and 2014.

Figure 3. Predatory Violent Crime 2013 and 2014 Risk Terrain Models



The highest concentration of predatory violent crime risk is in the central areas of Unincorporated DeKalb County. The distribution of yellow and light green areas in 2013 change to darker green in 2014 indicating decreased moderate level risk in the county overall.

The key observation from the risk terrain models above is that the distribution of risk remains relatively stable from 2013 to 2014. In addition, there appears to be overlap in higher risk areas in the central Unincorporated DeKalb County, suggesting that many predictors of crime are shared among the three crime types examined in the present study. To examine these patterns in further detail, the following figures present heat maps indicating only the highest risk areas in red. High risk areas are denoted in red, and indicate those areas with a risk value greater than two standard deviations above the mean.

Figure 4. High Risk Residential Burglary 2013 and 2014

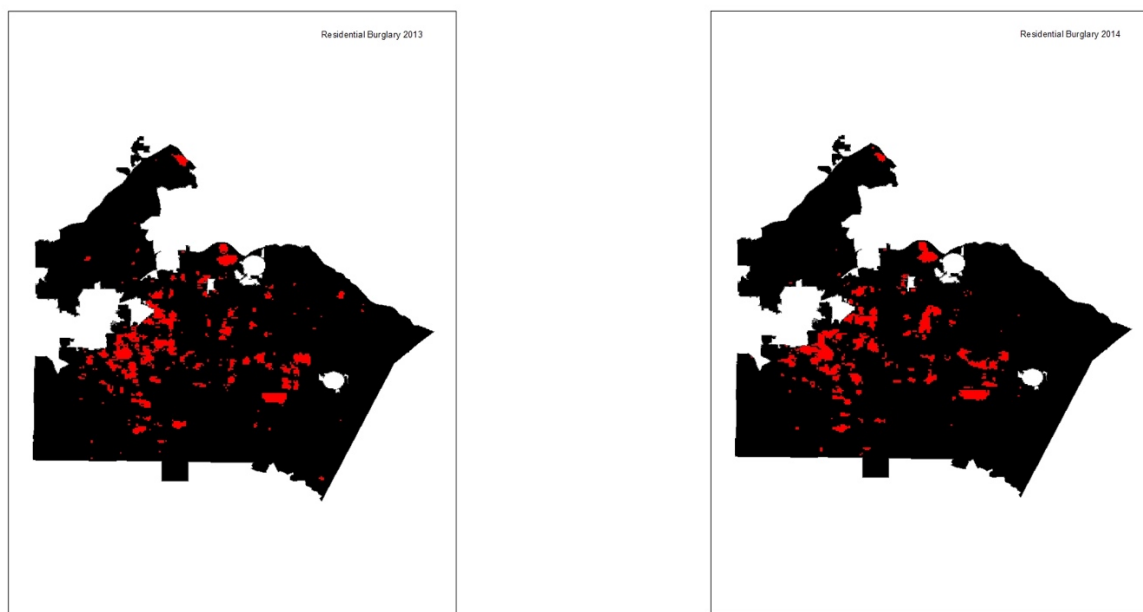


Figure 5. High Risk Business Robbery/Burglary 2013 and 2014

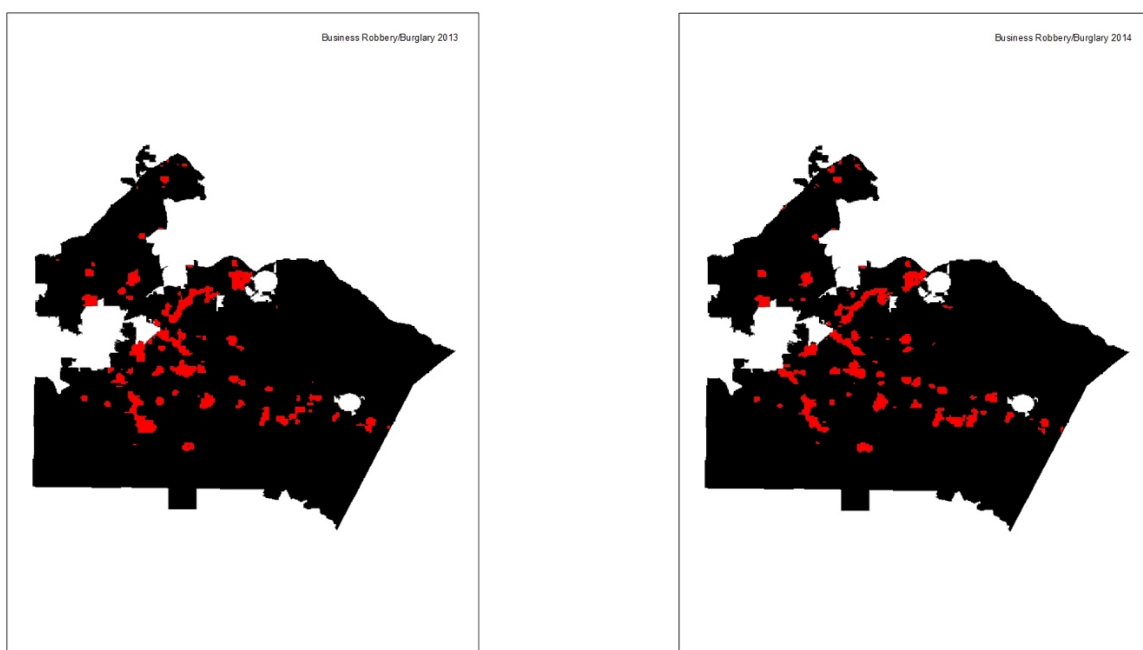
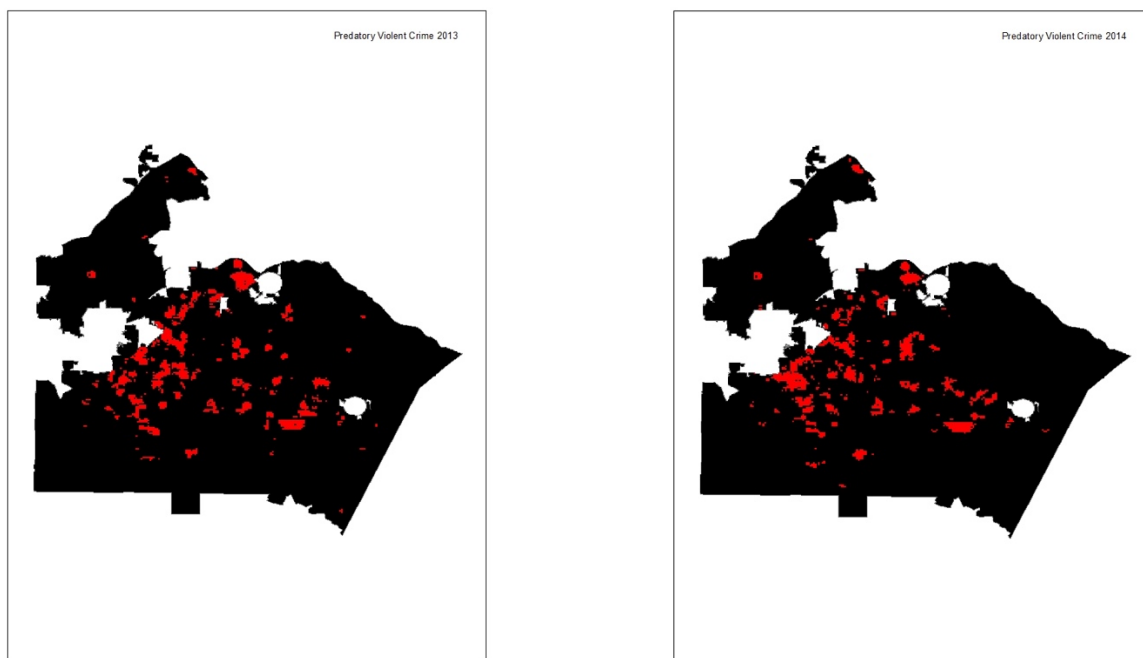


Figure 6. High Risk Predatory Violent Crime 2013 and 2014



The high risk maps reveal a similar pattern of relatively stable and clustered areas likely to experience crime in the future based on the presence of predictor variables. The red areas, indicating those raster cells with a composite risk score greater than two standard deviations above the mean are similar from 2013 to 2014 across each of the dependent variables. The stability of high risk areas is promising for the application of police and community response as they provide consistent targets for efforts to address factors related to crime. These findings are also consistent with other research showing stability in crime hot spots over time (Weisburd et al., 2004; Weisburd, Groff & Yang, 2012).

LOGISTIC REGRESSION RESULTS

To further examine the accuracy of the risk terrain model predictions, logistic regression models were conducted to examine the relationship between predicted high risk areas and the occurrence of crime in subsequent years. Each raster had a value of 1 as a predictor score if its composite risk score was greater than two standard deviations above the mean, and 0 if it was below. Each raster cell also had a value of 1 for an outcome score if the outcome crime of interest occurred in that cell in the subsequent year. A logistic regression was then used to determine if this relationship was statistically significant. This process was conducted to compare the 2013 RTM models to 2014 and 2015 outcomes, and to compare the 2014 RTM models to 2015 outcomes. The results of these analyses are presented in Table 6, Table 7, and Table 8.

Table 6. Logistic Regression for Residential Burglary

			RESIDENTIAL BURGLARY	
			2014	2015
HIGH-RISK RTM MODELS	2013	b(SE)	1.624 (.047)	1.625 (.040)
		OR	5.075	5.079
		Nag.	.083	.092
	2014	b(SE)		1.658 (.040)
		OR		5.247
		Nag.		.092

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

B=log-odds

SE = standard error

OR = odds ratio

Nag. Nagelkerke r^2

The results of the logistic regression for residential burglary suggest that the RTM process generated models in which the high-risk areas were significantly associated with crime outcomes in subsequent years. The odds of a residential burglary occurring was approximately five times higher in those areas identified as high risk as compared to areas identified as not high risk. Interestingly, this value did not change substantially in the comparison of the 2013 model to both 2014 and 2015 outcomes, suggesting that there was no substantial change in predictive performance over a longer period of time.

Table 7. Logistic Regression for Business Robbery/Burglary

			BUSINESS ROBBERY/BURGLARY	
			2014	2015
HIGH-RISK RTM MODELS	2013	b(SE)	2.435 (.114)	2.113 (.086)
		OR	11.417	8.270
		Nag.	.123	.103
	2014	b(SE)		2.207 (.087)
		OR		9.086
		Nag.		.112

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

B=log-odds

SE = standard error

OR = odds ratio

Nag. Nagelkerke r^2

The difference in odds ratio across years and models for business robbery/burglary suggests greater variation in RTM model performance. The 2013 RTM model was associated with eleven times higher odds of business robbery/burglary in 2014 and eight times higher odds of business robbery/burglary in 2015. This indicates diminished model performance over time. However, both appear to be strong predictors.

The 2014 RTM model was associated with nine times higher odds of business robbery/burglary in 2015 showing a similar model performance to that of the 2013-2015 model.

Table 8. Logistic Regression for Predatory Violent Crime

			PREDATORY VIOLENT CRIME	
			2014	2015
HIGH-RISK RTM MODELS	2013	b(SE)	2.119 (.069)	2.043 (.062)
		OR	8.326	7.715
		Nag.	.117	.114
	2014	b(SE)		2.032 (.062)
		OR		7.628
		Nag.		.114

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

B=log-odds

SE = standard error

OR = odds ratio

Nag. Nagelkerke r^2

The results of the logistic regressions for predatory violent crime shows similar patterns. Across all models, high-risk areas have approximately eight times higher odds of experiencing a predatory violent crime in the subsequent years. The model performance appears to decline slightly for the 2013 RTM model from 2014 to 2015.

In sum, it appears that the RTM models generated in the present study provide statistically significant predictions of those areas likely to experience crime in subsequent years. This finding holds across all crime types, though RTM models for business robbery/burglary show higher odds. The consistency in these models may facilitate long-term planning and intervention efforts to address the risk factors for crime.

MODEL DISCRIMINATION

While the initial findings from the visual inspection and logistic regression analysis are positive, further analysis is needed to determine the predictive validity of these models. Examining the Nagelkerke r^2 statistics from the logistic regression models suggests that while the high-risk areas are significant predictors of future crime, they explain only a small amount of variance in crime outcomes. The risk terrain models in the present analysis explain only 8-11% of the variance in the outcome measures. This means that 89-92% of crime in subsequent years was explained by factors not accounted for in the present analysis. Though the risk terrain models may be good indicators of high risk areas, they do not appear to perform well in identifying all of the key factors that contribute to future crime in those areas.

Another means to examine model performance is to examine how well the predictor “diagnoses” crime using receiver operating characteristic curves. The area under the curve (AUC) indicates the extent to

which the prediction performs better than random chance. The AUC values for each of the models is shown in Table 9.

Table 9. AUC Values

		2014	2015
Residential Burglary	2013	.669	.666
	2014		.659
Business Robbery/Burglary	2013	.761	.723
	2014		.731
Predatory Violent Crime	2013	.727	.717
	2014		.720

AUC values for these analyses vary between .659 and .761. A value of 0.5 indicates that the prediction performs no better than random chance while a value of 1 indicates perfect model prediction. The AUC values in this study indicate that the models perform approximately 32% to 52% better than random chance. The models for business robbery/burglary appear and predatory violent crime appear to perform slightly better than for residential burglary. These models perform moderately well, but further development is needed to improve performance prior to implementation.

CONCLUSION

The present study sought to develop comprehensive RTM models 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 briefly summarize key findings from the present study, outline limitations of the current study and data, identify several directions for future research, and discuss potential implications for policy and practice.

First, analyses showed several significant predictors of our crime outcomes and that strongest predictors varied across crime types. For residential burglary, addresses to which Marshalls and Sheriffs had been dispatched, the presence of foreclosures, and code compliance violations were the strongest predictors. The latter two, in particular, make theoretical sense as they can be viewed as indicators of lack of guardianship at properties and thus potentially be deemed low-risk targets by burglars (Cohen & Felson, 1979). Turning to business robbery and burglary, the strongest predictors were the presence of low-level criminal offenses, narcotics offenses, cash-centered businesses, and racial heterogeneity. These again make theoretical sense as cash-centered businesses are suitable targets for robbers and burglars due to the presence of large amounts of cash, while the presence of other types of crimes indicates a local pool

of motivated offenders and social disorganization (Cohen & Felson, 1979). Finally the strongest predictors of predatory violent crime were the presence of low-level crime, addresses to which Marshall's had been dispatched, percent unemployed, and the presence of narcotics offenses, all of which are similarly supported by past research.

Second, the RTM models themselves produced important findings which have implications for crime prevention. Most importantly, the heat maps show: 1) that crime risk is tightly clustered in places—particularly in the central portion of Unincorporated Dekalb County; 2) this clustering was similar across the three different crime outcomes examined, suggesting that these crime types have shared risk factors; 3) the clustering was remarkably stable from 2013 to 2014. Further, the logistic regression analyses showed that the RTM models significantly predicted crime in subsequent years. These findings add support for recent research showing that crime hot spots tend to be stable over time (Weisburd et al., 2004; Weisburd, Groff & Yang, 2012) and tend to share risk factors. As such, this is supportive of the utility for RTM to help police and other agencies identify both areas to focus their limited resources on and which risk factors to target for the largest potential impacts on crime.

Third, a central contribution of the current study was performing a detailed assessment of the discriminant abilities of the RTM models to effectively explain and accurately predict crime risk. This was done by examining Nagelkerke R^2 statistics for the logistic regression models and examining the AUC of the ROC curves for the models. Results from these analysis suggest that while the RTM methods shows much promise, further additions and refinement are required. The R^2 statistics suggest the models only explain 8-11% of the variance in the outcome measures, meaning that 89-92% of crime in subsequent years was explained by factors not accounted for in the current study. The AUC numbers indicate that the RTM models perform approximately 32% to 52% better than random chance at predicting crime. These are not meaningless given effect sizes in the social sciences tend to be small; however, they do show the importance of continuing to refine RTM methodology if we are to more accurately inform policymakers on how to most efficiently and effectively prevent crime. The following sections outline some limitations of the current study and data and make recommendations for future research to improve RTM methodology.

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 modest 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 an 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

A potential issue was encountered in the model weighting process that warrants further evaluation. In generating the logistic regression models necessary to determine model weighting, variables that were significant at the bivariate level were not significant in the multivariate model. This may indicate some level of mediating/moderating effects. Analyses did not indicate that multicollinearity was the cause of these trends, but future research should consider potential interaction effects (e.g. mediation, moderation, 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 others based on spatial factors that increase the threat of or vulnerability to crime (Kennedy & Caplan, 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; Kennedy & Caplan, 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

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 needed 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. 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 all three 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.

References

- Bennell, C., Jones, N.J., & Melnyk, T. (2009). Addressing problems with traditional crime linking methods using receiver operating characteristic analysis. *Legal and Criminological Psychology, 14*, 293-310.
- Braga, A. A., Papachristos, A. V., & Hureau, D. M. (2014). The effect of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly, 31*, 633-663.
- Caplan, J. M. (2010). *GIS for public safety: An annotated guide to ArcGIS tools and procedures*. Newark, N.J.: Rutgers Center on Public Security.
- Caplan, J. M. (2011). Mapping the spatial influence of crime correlates: A comparison of operationalization schemes and implications for crime analysis and criminal justice practice. *Cityscape, 13*(3), 57-83.
- Caplan, J. M. & Kennedy, L. W. (2010). *Risk terrain modeling manual*. Newark, NJ: Rutgers Center on Public Security.
- Caplan, J. M. & Kennedy, L. W. (2011). *Risk terrain modeling compendium*. Newark, NJ: Rutgers Center on Public Security.
- Caplan, J. M., Kennedy, L. W., Barnum, J. D., & Piza, E. L. (2015). Risk terrain modeling for spatial risk assessment. *Cityscape, 17*, 7-16.
- Caplan, J. M., Kennedy, L. W., Miller, J. (2011). Risk terrain modeling: Brokering criminological theory and GIS methods for crime forecasting. *Justice Quarterly, 28*, 360-381.
- Caplan, J. M., Kennedy, L. W., & Piza, E. L. (2013). Joint utility of event-dependent and environmental crime analysis techniques for violent crime forecasting. *Crime & Delinquency, 59*, 243-270.
- Center for Poverty Research. (n.d.). How does level of education relate to poverty?. Retrieved from <http://poverty.ucdavis.edu/faq/how-does-level-education-relate-poverty>.
- Clarke, R. V. (1995). Situational crime prevention. In M. Tonry & D. P. Farrington (Eds.), *Building a safer society: Strategic approaches to crime prevention. Crime and Justice 19* (91–150). Chicago: Univ. of Chicago Press.
- Cohen, L. E. & Felson, M. (1979). Social change and crime rate trends: A routine activities approach. *American Sociological Review, 44*, 588-608.

- Cui, L. & Walsh, R. (2015). Foreclosure, vacancy and crime. *Journal of Urban Economics*, 87, 72-84.
- DeKalb County Georgia. (n.d.). Foreclosure/vacant property registry. Retrieved from <http://www.dekalbcountyga.gov/beautification/foreclosurevacant-property-registry>.
- DeNavas-Walt, C. & Proctor, B. D. (2015). Income and poverty in the United States: 2014. Retrieved from <https://www.census.gov/content/dam/Census/library/publications/2015/demo/p60-252.pdf>.
- Drawve, G. (2016). A metric comparison of predictive hot spot techniques and RTM. *Justice Quarterly*, 33, 369-397.
- Drawve, G., Moak, S. C., & Berthelot, E. R. (2016). Predictability of gun crimes: a comparison of hot spot and risk terrain modelling techniques. *Policing and Society*, 26(3), 312-331.
- Drawve, G., Thomas, S. A., & Walker, J. T. (2014). The likelihood of arrest: A routine activity theory approach. *American Journal of Criminal Justice*, 39, 450-470.
- Drawve, G., Thomas, S. A., & Walker, J. T. (2016). Bringing the physical environment back into neighborhood research: The utility of RTM for developing an aggregate neighborhood risk of crime measure. *Journal of Criminal Justice*, 44, 21-29.
- Durose, M. R., Cooper, A. D., & Snyder, H. N. (2014). Recidivism of prisoners released in 30 states in 2005: Patterns from 2005 to 2010. Retrieved from <https://www.bjs.gov/content/pub/pdf/rprts05p0510.pdf>.
- Engstad, P. A. (1975). Environmental opportunities and the ecology of crime. In D. Cornish and R. V. Clarke (Eds.), *Crime in Canadian Society*. Toronto: Butterworths.
- Farley, M. & Barken, H. (2008). Prostitution, violence, and posttraumatic stress disorder. *Women & Health*, 27(3), 37-49.
- Federal Bureau of Investigation. (n.d.). Offense definitions. Retrieved from <https://ucr.fbi.gov/crime-in-the-u.s/2015/crime-in-the-u.s.-2015/resource-pages/offense-definitions>.
- Federal Bureau of Investigation. (2016). Arrests. Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2015/crime-in-the-u.s.-2015/persons-arrested/arrestmain_final.pdf.
- Ferguson, A. G. (2012). Predictive policing and reasonable suspicion. *Emory Law Journal*, 62, 259-325.
- Grubestic, T. H. & Pridemore, W. A. (2011). Alcohol outlets and clusters of violence. *International Journal of Health Geographics*, 10(30), 1-12.
- Gruenewald, P. J., Freisthler, B., Remer, L., Lascala, E. A., & Treno, A. (2006). Ecological models of alcohol outlets and violent assaults: Crime potentials and geospatial analysis. *Addiction*, 101, 666-677.
- Hinkle, J. C. & Weisburd, D. (2008). The irony of broken windows policing: A micro-place study of the relationship between disorder, focused police crackdowns and fear of crime. *Journal of Criminal Justice*, 36, 503-512.
- Immergluck, D., & Smith, G. (2006). The external costs of foreclosure: The impact of single-family mortgage foreclosures on property values. *Housing Policy Debate*, 17(1), 57-79.
- Joh, E. E. (2014). Policing by the numbers: Big data and the Fourth Amendment. *Washington Law Journal*, 89, 35-68.
- Kennedy, L. W., & Caplan, J. M. (2012). A Theory of Risky Places. *Rutgers Center on Public Security*. Newark, NJ.

- Kennedy, L. W., Caplan, J. M., & Piza, E. L. (2011). Risk clusters, hotspots, and spatial intelligence Risk terrain modeling as an algorithm of police resource allocation strategies. *Journal of Quantitative Criminology*, 27(3), 339-362.
- Kennedy, L., Caplan, J., & Piza, E. (2015). A multi-jurisdictional test of risk terrain modeling and place-based evaluation of environmental risk-based patrol deployment strategies. Retrieved from http://www.rutgerscps.org/uploads/2/7/3/7/27370595/nij6city_resultsexecsum_final.pdf.
- Kennedy, L. W., Caplan, J. M., Piza, E. L., & Buccine-Schraeder, H. (2015). Vulnerability and exposure to crime: Applying risk terrain modeling to the study of assault in Chicago. *Applied Spatial Analysis and Policy*, 1-20.
- Kornhauser, R. (1978). *Social sources of delinquency*. Chicago: University of Chicago Press.
- Lens, M. C. & Meltzer, R. (2016). Is crime bad for business? Crime and commercial property values in New York City. *Journal of Regional Science*, 56, 442-470.
- Martinez, R., Rosenfeld, R., & Mares, D. (2008). Social disorganization, drug market activity, and neighborhood violent crime. *Urban Affairs Review*, 43, 846-874.
- Moreto, W. D., Piza, E. L., & Caplan, J. M. (2014). "A plague on both your houses?": Risk, repeats, and reconsiderations of urban residential burglary. *Justice Quarterly*, 31(6), 237-264.
- National Institute of Justice. (2014). From juvenile delinquency to young adult offending. Retrieved from <https://www.nij.gov/topics/crime/Pages/delinquency-to-adult-offending.aspx>.
- Niesse, M. (2016). Cities outnumbered by unincorporated DeKalb. Atlanta Journal and Constitution. Retrieved from <http://www.ajc.com/news/local-govt--politics/cities-outnumbered-unincorporated-dekalb/Tn5A9ZZ1hEQ2WObtXd6YL/>.
- Piza, E., Feng, S., Kennedy, L., & Caplan, J. (2016). Place-based correlates of motor vehicle theft and recovery: Measuring spatial influence across neighborhood context. *Urban Studies*, 1-24.
- Pontius, R. G., & Schneider, L. C. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment*, 85, 239-248.
- Pope, D. G. & Pope, J. C. (2012). Crime and property values: Evidence form the 1990s crime drop. *Regional Science and Urban Economics*, 42, 177-188.
- Reuter, P. (2009). Systemic violence in drug markets. *Crime, Law and Social Change*, 52, 275-284.
- Rosenbaum, D. P. (2007) The limits of hot spots policing. In D. Weisburd & A. A. Braga (Eds.), *Police innovation: Contrasting perspectives* (245-266). Cambridge: Cambridge University Press.
- Sampson, R. J., & Raudenbush, S. W. (2004). Seeing disorder: Neighborhood stigma and the social constriction of "broken windows." *Social Psychology Quarterly*, 67, 319-342.
- Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas*. Chicago, IL: University of Chicago Press.
- Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27, 27-55.
- Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford: Oxford university press.
- Skogan, W. G. (1990). *Disorder and decline: Crime and the spiral of decay in American cities*. New York: Free Press.

- Speer, P. W., Gorman, D. M., Labouvie, E. W., & Ontkush, M. J. (1998). Violent crime and alcohol availability: Relationships in an urban community. *Journal of Public Health, 19*, 303-318.
- Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M. J., & Kattan, M. W. (2010). Assessing the performance of prediction models: A framework for some traditional and novel features. *Epidemiology, 21*(1), 128-138.
- Valera, R. J., Sawyer, R. G., & Schiraldi, G. R. (2000). Violence and post traumatic stress disorder in a sample of inner city street prostitutes. *American Journal of Health Behavior, 25*, 149-155.
- Warner, B.D. (2014). Neighborhood factors related to the likelihood of successful informal social control efforts. *Journal of Criminal Justice, 42*, 421-430.
- Weisburd, D., Bushway, S., Lum, C., & Yang, S. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of Seattle. *Criminology, 42*(2), 283-322.
- Weisburd, D., Groff, E. R., & Yang, S. (2012). *The criminology of place: Street segments and our understanding of the crime problem*. Oxford: Oxford University Press.
- Wilson, J. Q. & Kelling, G. L (1982). Broken windows: The police and neighborhood safety. *The Atlantic Monthly, March*, 29-38.
- Yang, S. (2009). Assessing the spatial-temporal relationship between disorder and violence. *Journal of Quantitative Criminology, 26*, 139-163.
- Yang, Y. & Diez-Roux, A. V. (2012). Walking distance by trip purpose and population subgroups. *American Journal of Preventative Medicine, 43*(1), 11-19.

About the Authors

Audrey Clubb is a research associate with Resource Development Associates in Oakland, California. Clubb's research interests include justice program evaluation, evidence-based practice, crime in the social context, and policing innovation. She currently conducts research throughout the state of California including evaluations of Assembly Bill 109 implementation, Proposition 47 implementation, Los Angeles County Probation Department, and Whole Person Care implementation. Her research has been published in *Criminal Justice Studies*, *Journal of Financial Crime*, *Criminal Justice Review*, and the *Oxford Handbook of Criminological Theory*. She received her Ph.D. from the Andrew Young School of Policy Studies at Georgia State University.

Dean Dabney is a professor of criminal justice and criminology at Georgia State University. His scholarly interests include the organizational culture within law enforcement agencies, forms of deviance and/or criminal behaviors that occur in organizational settings, and qualitative research methods. He has published four books, including the recently released *Speaking Truth to Power: Confidential Informants and Police Investigations* and several dozen articles appearing in such outlets as *British Journal of Criminology*, *Criminal Justice & Behavior*, and *Justice Quarterly*. He received his Ph.D. in sociology from the University of Florida.

Joshua Hinkle is an associate professor in the Department of Criminal Justice and Criminology at Georgia State University. Hinkle's research interests include evidence-based policing, the disorder-crime nexus, fear of crime and experimental methods. His research has been funded by the National Institute of Justice and the National Science Foundation and appears in journals such as *Criminology*, *Criminology & Public Policy*, the *Journal of Research in Crime and Delinquency* and the *Journal of Experimental Criminology*. He received his Ph.D. from the University of Maryland, College Park.

About the Center for State and Local Finance

The Center for State and Local Finance's (CSLF) mission is to develop the people and ideas for next generation public finance by bringing together the Andrew Young School's nationally-ranked faculty and the broader public finance community. CSLF conducts innovative, nonpartisan research on tax policy and reform, budget and financial management, education finance, and economic development and urban policy. Additionally, it provides premier executive education in public finance for state and local finance officials and works with local and state partners on technical assistance projects on fiscal and economic policy.

CSLF maintains a position of neutrality on public policy issues. However, in order to protect their academic freedom, authors may express a wide range of viewpoints in CSLF's publications. The research, interpretations or conclusions in CSLF publications should be understood to be solely those of the author(s).

For more information on the Center for State and Local Finance, visit our website at: cslf.gsu.edu.