

2019

## Boundaries and Policing: Space, Jurisdictions, and Roles in the Collection of Official Crime Data

Cynthia Suzanne Barnett-Ryan  
West Virginia University, [csbarnett@mix.wvu.edu](mailto:csbarnett@mix.wvu.edu)

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>



Part of the [Geographic Information Sciences Commons](#)

---

### Recommended Citation

Barnett-Ryan, Cynthia Suzanne, "Boundaries and Policing: Space, Jurisdictions, and Roles in the Collection of Official Crime Data" (2019). *Graduate Theses, Dissertations, and Problem Reports*. 3865.  
<https://researchrepository.wvu.edu/etd/3865>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact [researchrepository@mail.wvu.edu](mailto:researchrepository@mail.wvu.edu).

# Boundaries and Policing

Space, Jurisdictions, and roles in the collection of official crime data

Cynthia Barnett-Ryan

Dissertation submitted

To the Eberly College of Arts and Sciences

at West Virginia University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Geography

Gregory Elmes, Ph.D., Chair

Jamison Conley, Ph.D.

Trevor Harris, Ph.D.

James J. Nolan, Ph.D.

Emily H. Griffith, Ph.D.

Department of Geology and Geography

Morgantown, West Virginia

2019

Keywords: Uniform Crime Reporting Program; Policing; Uncertainty; Place

Copyright 2019: Cynthia Barnett-Ryan

## ABSTRACT

# Boundaries and Policing

Space, Jurisdictions, and roles in the collection of official crime data

Cynthia Barnett-Ryan

The Uniform Crime Reporting (UCR) Program is a law enforcement statistical system open to unreported information due to its voluntary nature. As such, there needs to be a valid and accepted means to estimate official reports of crime for those different levels of geography where reporting may be incomplete. Current methods of imputing and modeling UCR data, which have not been updated since the 1960s, are based upon conceptualizations of law enforcement agencies that may no longer be valid. These older models do not appropriately represent the law enforcement assessment of space and place and its effects on discretionary recording behavior. The number of specialized agencies that share jurisdiction and population with primary law enforcement agencies has increased since early data modeling techniques were developed around the 1960s. This study explores the connection between the policing and the collection of crime data to advance our understanding of how differences among types of law enforcement may impact the discretionary decision to record data. To explore this topic, I have divided this study into three papers touching on differing dimensions of place, scale, and uncertainty connected to the recording of law enforcement data. The data for these papers includes national UCR Program data, as well as calls for service and recorded incident data from two law enforcement agencies in the mid-South—Knoxville Police Department and the University of Tennessee Police Department.

Firstly, this research explores the influence of agency attributes to assess their possible impact on the treatment of missing data. The coefficient of variation (CV) is used to measure the internal variation of reported crime within various groups of agencies. The average CVs calculated with and without specialized agencies are compared using a Jackknifing technique to test whether the presence of specialized agencies increases the internal variation within the group or not. The comparison demonstrates that eliminating specialized agencies from the strata has a statistically-significant effect on reducing internal variation for property crimes. For violent offenses, however, the results are more modest. While the average CV for violent crime does decrease with the elimination of specialized agencies, the improvements are not statistically significant. The results from this research point to a greater need to address the changing circumstances to incorporate the diversity of law enforcement agency type.

Secondly, although there is an interest by researchers to use calls for service (CFS) as a useful proxy for recorded incident information by law enforcement as more of this type of data is made available in open data initiatives, the assumption that CFS could serve as a proxy for incident information in spatial

analysis is not supported by the evidence. Instead, there is some indication that law enforcement activities are mediated by the agency's goals for its data, such as intelligence-led policing or fulfillment of Clery Act reporting, thus affecting the recording of incident information. Using data from two different types of law enforcement agency within the same community, CFS and incident reports for property crimes in April 2014 were tested for spatial association using both the Cross-K function and the Co-location Quotient. Findings from this study show there is a modest amount of detectable clustering of CFS for the agency that fits a model of traditional municipal law enforcement. However, the law enforcement agencies serving a large university campus did not show any detectable spatial association for these events. The findings suggest that in the movement towards using open data researchers will need to take greater care in the selection of data to understand if underlying spatial assumptions about the data can be supported.

Thirdly, an increasing quantity of data is currently being made available by law enforcement agencies, but frequently that data is not a consistent level of areal aggregation and scale. Factors such as the Modifiable Areal Unit Problem (MAUP) and the Uncertain Geographic Context Problem (UGCoP) make rectifying differing scales problematic. Central to this problem are the dynamics of recording crime data and whether law enforcement activity—specifically the concept of the patrol officer in a boundary role—is a key influence that should be accounted for in crime data models. With data from a midsized, southern municipal police department, two dasymetric allocation techniques using street networks and street networks weighted by calls for service are used to test potential improvements on the scale and aggregation problem through the introduction of law enforcement activity into allocation models for recorded crime data. Results demonstrate that the introduction of law enforcement activity—especially officer-initiated activity—improves the overall fit of the allocation of recorded crime into smaller subjurisdictional units. In addition, there is modest evidence to advocate for the use of law enforcement-generated subjurisdictional units (such as a precinct or beat) as opposed to population-based Census Tracts. These findings suggest that the production of crime statistics is subject to influences originating from law enforcement agency policy and the recording behavior of its officers.

The findings of the three studies inform important discussions in the geographic community on the heterogeneous nature of law enforcement. More explicitly, combining the conclusions of these three papers contributes to an evolving understanding of the representations of place by geographic information science (GISc) and criminology, and the construction of place through the roles and behaviors of individuals, and the increasing use of “Big Geodata”. Future research with data collected from official police activities should consider the degree of uncertainty introduced by the nature of the activities themselves — especially considering the growing use, influence and reliance on georeferenced data produced by individuals not particularly informed about the nuances of geography.

Uncertainty in geographic representation arises because, of necessity, almost all representations of the world are incomplete. (Longley, et al. 2011, p. 147)

## ACKNOWLEDGMENTS

The saying goes, “it takes a village.” While I know that it’s an overused saying, I feel it is the most apt description of my journey over the past several years. Though, in fact, it might be more correct to thank all my “villages” that made my studies possible, especially since I was pursuing this degree after many years of working. Over the course of the years of my studies, I’ve been asked about what I seek to gain from this experience—a raise or a promotion. This was in many ways a purely selfish act, but not in the way most people think about it. The honest truth is I did all of this because I desperately wanted it. I wanted to push myself and learn in a way that I hadn’t been able to do in many years. However, at this point in my life, with all of its responsibilities of work and family, it would not be possible to succeed alone.

First and most of all, I need to thank my chair and advisor, Dr. Gregory Elmes. He never ceased to offer encouragement to me. Besides my family, he was my biggest cheering section and provided support especially when I doubted myself. I have come to enjoy our regular Saturday coffees with you and Jeanne to talk about my progress and to share the usual exploits of our families. I also am grateful for the guidance provided by the rest of my committee. I consider myself very lucky to have such a diverse group of committee members to challenge and push my work.

I would be remiss to not mention the assistance provided by key individuals from the Knoxville Police Department and the University of Tennessee Police Department—specifically Jonne Crick, Deputy Chief Gary Holliday, and Deputy Chief Emily Simerly. This study would not be possible without their support and assistance.

I must thank my dear friends at home, work, and church, as well as my parents and family. I have so many people to thank for their love, support, and assistance. I truly can say that I would not be able to have finished without you. In big and small ways, I have been blessed with help. Whether it was just a word of encouragement, a meal dropped off on my doorstep during my exams, a ride provided to my son to practice, a hug, or a celebration of a small success, these offerings have touched me and sometimes surprised me in the best possible ways. Please know that if you see even a glimmer of yourself in this note of thanks, it is because you should.

Finally, I need to specifically thank my husband and son, Terry and Patrick, for your unwavering presence during the last seven years. I will never be able to fully express my gratitude for the gift of this opportunity.

# 1 TABLE OF CONTENTS

---

1.1	Introduction .....	1
1.2	Format.....	1
1.3	Problem Statement.....	2
2	Missing Data in Uniform Crime Reporting and the Impact of Agency Type .....	5
2.1	Introduction .....	5
2.1.1	Sources of Variation in UCR Data.....	5
2.1.2	Law Enforcement Agencies.....	7
2.1.3	Agency Information in the UCR Program.....	7
2.1.4	Missing Data Imputation and Other Uses of Agency Information in the UCR Program .....	10
2.1.5	Nature of Missing Data in the UCR Program .....	14
2.2	Discussion and Results .....	14
2.2.1	Variation within the Reporting Agencies .....	14
2.2.2	Coefficient of Variation .....	14
2.2.3	Jackknifing.....	16
2.2.4	Results.....	17
2.2.5	Conclusion.....	21
2.3	Appendix .....	23
3	Crime and Place: Differences in Spatial relationship between Calls for Service and Recorded Incidents for Municipal and Campus Law Enforcement .....	24
3.1	Introduction .....	24
3.2	Crime and Place .....	25
3.3	Police in a Boundary Role .....	27
3.4	Methodology and Expected Results .....	28
3.4.1	Agency Priorities and Use of Data.....	28
3.4.2	Spatial Correlation and Co-location.....	31
3.5	Analysis .....	32
3.5.1	Cross-K Analysis.....	32
3.5.2	Global Colocation Quotient .....	34
3.5.3	Local Colocation Quotient.....	35
3.6	Discussion and Conclusion.....	38
4	The Recording of Crime Data and Spatial Scale: The Influence of Law Enforcement Recording Behavior and the Uncertain Geographic Context Problem .....	41

4.1	Spatial Uncertainty Associated with Law Enforcement Data .....	41
4.2	Background .....	44
4.2.1	Crime Data .....	44
4.2.2	Uncertainty, Crime Data, and Criminality of Place .....	45
4.3	Methodology and Data .....	48
4.3.1	Dasymetric Mapping and Other Areal Interpolation Methods.....	49
4.3.2	Law Enforcement and Census Geography .....	50
4.3.3	Data.....	51
4.3.4	Study Methodology.....	54
4.4	Results .....	58
4.5	Discussion and Conclusion .....	61
5	Conclusion.....	63
5.1	Big Geodata.....	63
5.2	The place of the individual.....	65
5.3	Future Research on Place, space, and the “placial turn” .....	67
5.4	Discussion.....	68
6	Citations .....	70



## LIST OF TABLES

Table 2-1. Number and percentage of agencies by type and population category, 2012 .....	8
Table 2-2. Heterogeneity of population group by agency type, 2012 .....	9
Table 2-3. Percentage of imputed data in national estimates for Part I crimes, 2012 .....	11
Table 2-4. Reporting pattern by population category for original Part I offenses and arson, 2012 .....	13
Table 2-5. Number of Agencies within each stratum by stratification technique .....	23
Table 3-1. Number of offense "tags" by agency .....	29
Table 3-2. Counts and percentages of CFS and incident reports by agency and offense type, 2014.....	30
Table 3-3. Pearson correlations of monthly counts of CFS and incident reports, .....	31
Table 3-4. Global colocation quotients, CFS to incident reports with property crime, April 2014 .....	35
Table 3-5. Offense categories <sup>a</sup> from the UCR Program .....	40
Table 4-1. Descriptive statistics for CFS and recorded incidents .....	51
Table 4-2. Preset densities used in dasymetric mapping .....	56
Table 4-3. Goodness-of-fit for dasymetric-mapping produced predictions .....	58

## LIST OF FIGURES

Figure 2-1. Percent of population represented by agencies reporting complete offense data to the UCR Program.....	12
Figure 2-2. Specialized law enforcement agencies by county as a percent of total law enforcement agencies, 2012 .....	13
Figure 2-3. Coefficients of variation within violent offenses, 2012 .....	15
Figure 2-4. Coefficients of variation within property offenses, 2012.....	16
Figure 2-5. Difference in mean coefficient of variation based in strata based on state and population group using 95 percent confidence interval .....	18
Figure 2-6. Difference in mean coefficient of variation in strata based on region and population group using 95 percent confidence interval.....	19
Figure 2-7. Difference in mean coefficient of variation in strata based on division and population group using 95 percent confidence interval.....	20
<i>Figure 3-1. CFS and incident reports for property crime in study area, April 2014.....</i>	<i>32</i>
Figure 3-2. Cross-K analysis from CFS to incident reports, UTPD .....	33
Figure 3-3. Cross-K analysis from CFS to incident reports, KPD.....	34
Figure 3-4. Local colocation quotients, CFS to incident reports with property crime, April 2014 .....	37
Figure 4-1. Scatterplot of recorded incidents by CFS aggregated by police beat.....	52
Figure 4-2. Scatterplot of Recorded Crime by Officer-Initiated Calls Aggregated by Police Beat .....	53
Figure 4-3. Scatterplot of Recorded Crime by CFS Aggregated by Census Tract .....	53
Figure 4-4. Scatterplot of Recorded Crime by Officer-Initiated Calls aggregated by Census Tract .....	54
Figure 4-5. Density maps used as Ancillary Data for Dasymetric Models.....	57
Figure 4-6. Prediction Error for Census Tracts (Clipped) .....	59
Figure 4-7. Prediction Error for Police Beats.....	60

## 1.1 INTRODUCTION

Crime data is the foundation of analysis and policy formation in contemporary law enforcement. Yet the discovery and recording of crime incidents is deeply contextualized in the culture of law enforcement and this culture is not monolithic in its approaches towards policing or crime analysis. It is often assumed that uniformity is provided by the common standards of the national statistics from the Uniform Crime Reporting (UCR) Program. Indeed, its name implies that. However, the FBI cautions users of UCR data against simplistic comparisons between places—such as ranking—due to the numerous variables that can affect the discovery and collection of UCR data and thereby introduce uncertainty and incompatibility between sources (Federal Bureau of Investigation, 2014).

In the light of the need for accurate and reliable statistics, especially in the current climate of public and professional concern about policing equity across geography, demographic and socio-economic groups, this research study examines the various influences on the collection of crime data across boundaries of jurisdictions and the boundaries between law enforcement and the communities they serve. The study has the eventual objective of suggesting means to improve the use and analysis of crime statistics through the incorporation of ancillary information, including the imputation of missing data. To this end, the research framework investigates the differing dimensions of policing adopted by law enforcement agencies and seeks to link the policing behavior produced by these different characteristics to the way official crime statistics are recorded. If evidence of police behavior and agency patterns are revealed by this study to influence the way crime is reported, then information about this relationship may be used to improve the fit of analytical and imputation models.

## 1.2 FORMAT

The main portion of this dissertation is structured as three papers situated between this introduction and an end section which draws the findings and their implications together. The first paper sets the stage for the need for additional exploration of policing and crime data and is titled “Missing Data in Uniform Crime Reporting and the Impact of Agency Type.” The analysis examines the variability of crime reporting by law enforcement agencies and whether the *population group* (a category created by combining population served and whether the agency serves a municipal area or not) is sufficient to capture the observed variability. In addition, the analytical approach uses a jackknifing procedure to test whether any detected differences are statistically significant. Jackknifing techniques provide an opportunity to test whether detectable differences can be attributed to more than random variation. Consequently, any explained variation is a candidate for inclusion in improved models.

The second paper is titled “Crime and Place: Differences in Spatial Relationship Between Calls for Service and Recorded Incidents for Municipal and Campus Law Enforcement.” It compares the spatial association between calls for service (CFS) and recorded incident reports by offense type (e.g., crimes against persons, crimes against property, crimes against society) to measure the potential influence of calls on recorded crime data. The recording of crime data is presumed to be related to a combination of the type of law enforcement agency and the style of policing used by the agency. For example, the analysis of data by University of Tennessee Police Department (UTPD) is could be influenced by the reporting requirements of the Clery Act rather than other priorities. The working hypothesis is that there is less association between CFS and incident data for UTPD because the need to record certain crimes is driven by the needs of the Clery Act, rather than where there might be reports of all crime.

The underlying assumption is that there should be a close relationship between CFS and reports that expresses what the needs are, and that CFS and reports should be in close alignment. The Co-location Quotient is used to test this assumption.

The third paper, “The Recording of Crime Data and Scale: The Influence of Law Enforcement Recording Behavior and the Uncertain Geographic Context Problem,” investigates the recording of crime data and scale by focusing on the influence of law enforcement recording behavior. An increasing amount of data is being made available by law enforcement, but frequently that data is not available at a comparable level of aggregation and scale. Factors such as the Modifiable Areal Unit Problem (MAUP) and the Uncertain Geographic Context Problem (UGCoP) make rectifying data across differing scales problematic. Central to this problem is the dynamics of recording crime data and whether variation in law enforcement activity is a key influence that should be accounted for in crime data models. Using data from a midsized, southern municipal police department, two dasymetric allocation techniques, one using street networks and the second, street networks weighted by calls for service, are used to test potential improvement through the introduction of law enforcement activity into allocation models for recorded crime data.

The dissertation is embedded within an ongoing discussion of space and place in the disciplines of geography and criminology and raises elements of the handling of “Big Data”, specifically big geo-data within the field of crime analysis. The findings of the three studies raise important questions for the geographic community on the heterogeneous nature of law enforcement. More explicitly, combining the conclusions of these three papers contributes to an evolving understanding of the representations of place by geographic information science (GISc.) and in criminology, and the construction of place through the roles and behaviors of individuals, and the increasing use of “Big Geodata”. Future research with data collected from official police activities should consider the degree of uncertainty introduced by the nature of the activities themselves — especially considering the growing use, influence and reliance on georeferenced data produced by individuals not particularly informed about the nuances of geography. Researchers should be cognizant of the diversity sometimes suppressed by apparent uniformity.

### 1.3 PROBLEM STATEMENT

While police practices are listed as one of the variables that can affect the collection of UCR data (Federal Bureau of Investigation, 2014), little has been done to investigate what the variation of those policing practices from jurisdiction to jurisdiction may have on crime data reported as statistics. The UCR Program collects statistical information from law enforcement agencies across the nation. Since 1929, it has provided one of only two national measures of crime and criminality for the United States. UCR crime statistics are deeply immersed in the context of police activity rather than being a standardized reflection of the criminality of a particular place. The fact that the data flows from the administrative and official work of the law enforcement agency means that the data cannot capture any information that law enforcement agency does not record. Unreported crime data includes crimes that police are unaware of because victims do not make a report, or crimes that are handled informally by virtue of police discretion (Boivin & Cordeau, 2011; Lott & Whitley, 2003; M. D. Maltz, 1977; Michael D Maltz & Targonski, 2003; Pepper, Petrie, & Sullivan, 2010).

Policing has also evolved in its approach since the beginning of the UCR Program in the 20<sup>th</sup> century. Moreover, the data handling methods in the UCR Program do not consider the wide diversity now present in policing styles in the 21<sup>st</sup> century and how that diversity affects the ability of police to discover crime and record crime. In the early part of the 20<sup>th</sup> century when UCR was developed, there was a general professional model of policing used by most agencies that relied primarily on random patrols and responding to CFS. Since the 1980's, policing has evolved to incorporate a variety of approaches—both proactive and reactive—to address crime as a social problem in communities. One of the particular areas of growth is in the use of crime analysis to direct policing activities rather than relying on the more random and reactive methods of the past (Maguire & King, 2004; Vila & Morris, 1999; Zhao & Hassell, 2005).

Different approaches to policing can lead to differences among agencies in terms of the emphasis of the crime analysis products. For example, agencies that subscribe to zero-tolerance policing practices are more likely to be associated with many of the COMPSTAT-style practices, such as mapping the locations of certain types of lower-level crimes and directing patrols to those areas (Walsh, 2001; Walsh & Vito, 2004; D. Weisburd, Mastrofski, McNally, Greenspan, & Willis, 2003). Agencies that make policing decisions based more on intelligence on criminal actors (i.e., intelligence-led policing) will focus on the offender identities associated with known criminal problems and the linkages between the offender and other criminal actors and places (Ratcliffe 2008).

In addition to impact of the different approaches to policing to crime statistics, data submitted to the UCR Program are also reported at the jurisdictional level with no other geographic or location information attached to either the offense counts or incident-level information, even where it's available (Michael D Maltz, 2009; Michael D Maltz & Targonski, 2003). The geographically aggregate nature of the UCR data is unlikely to change soon. While most crime analysis at the local level has become reliant on crime mapping using point data for each criminal incident, the national data, of necessity, remains aggregated by jurisdiction. Even if the national UCR Program adds the capability to capture more precise location information beyond the jurisdiction, the information released will still need to include some obfuscation or masking of location to avoid violating the privacy of the individuals associated with the incident. This is not unlike the approach of the Census Bureau and health agencies in their policies regarding the release of data under the auspices of HIPAA (Gatewood, 2001).

Rather than focusing solely on either qualitative or quantitative methods, this study employs elements of both. The rationale for the application of both qualitative and quantitative models is twofold. First, the qualitative assessment of analytical products produced by law enforcement agencies will allow for a deeper understanding of the ways that policing can ultimately affect the recording of data. The incorporation of qualitative data raises questions associated with space, place and big data that are not immediately apparent with quantitative data. Secondly, having a better understanding of the influence of the analytical priorities of the agency increases the utility of the crime data, as well as improving the interpretation of the data.

Specifically, I will address the following objectives:

- What characteristics can be used to identify peer agencies for crime reporting data given the current diversity of policing? What spatial and other dimensions seem to matter and at what scale?

- Can information on the location of calls for service (CFS) improve the prediction of law enforcement recording of crime incident data for crime types identified as a priority through an assessment of analytical products? Can differences among law enforcement agency types be measured in the spatial properties of recorded incidents?
- Crime statistics in UCR are a result of police activities connected to both calls-for-service (CFS) and officer-initiated acts. CFS represent “demand” for police services, but what can be said about discretionary behavior as measured by officer-initiated activities? Do officer-initiated activities exhibit the same spatial patterns as CFS?

The case study focuses on an area in Tennessee identified for the collection of data which centers on the metropolitan area of Knoxville. Using data collected by law enforcement agencies with differing responsibilities that serve the same geographic area differences are expected to emerge that can inform the understanding of the impact of law enforcement actions on crime statistics. While the results from these three studies will not necessarily be generalizable in a statistical-probabilistic fashion to the national level, they suggest a method that could be used and validated through replication by other investigators. An underlying theme is the search for veracity between the data and the imputed results that reflects preparation for dealing with far greater volumes of data of greater variety than is experienced presently. The ending section of this dissertation discusses at some length some of the implications of space, place and diversity for future crime data processing and analysis that have emerged during the research.

## 2 MISSING DATA IN UNIFORM CRIME REPORTING AND THE IMPACT OF AGENCY TYPE

---

### 2.1 INTRODUCTION

In 2016, the United States Department of Justice provided \$274.9 million in block funding to states for crime control programs (Cooper & Hyland, 2015). This money is disbursed to the states based on a formula using violent crime information from the Uniform Crime Reporting (UCR) Program. The UCR Program data are collected from a universe of over 18,000 law enforcement agencies representing a subset of crimes that come to the attention of law enforcement by either victim reports or through law enforcement actions and observations. It is the only source of crime data that can be associated with a specific agency or place. However, reporting data to the UCR Program is strictly voluntary from a federal perspective<sup>1</sup>, which can lead to occasional or pervasive missing reports either for an individual month or the whole year's data. Because participation levels can and do vary from year to year, the UCR Program needs to have a method to impute crime data for agencies that either provide incomplete data or no data at all to publish national and state annual estimates of the level and trend of crime. These estimates often form the basis for larger decisions related to criminal justice and other social policy as referenced by Cooper and Hyland (2015). As such the methods used by the UCR Program to impute missing data are critical to the validity of these estimates and any resulting decisions.

This paper will build upon the research opened by Lynch and Jarvis (J. P. Lynch & Jarvis, 2008) and Maltz and others (Michael D. Maltz & Targonski, 2002; Michael D Maltz, 2006; Michael D Maltz & Targonski, 2003, 2004) which explores the validity of and provides possible improvements to current imputation methods by investigating the impact of agency type on the variability of the data. Specifically, this paper will propose a method for testing the fit of classification methods used in these crime imputation models. After providing background on the currently used imputation methods and laying out the theoretical concerns of missing data, the author will address two major issues related to the possible introduction of bias into national estimates for criminal offense counts through missing data. The paper will initially detail an analysis of the current composition of agencies. Also, the paper will explore the existing variation in the UCR contributing agency pool and test for possible improvements to the categorization of UCR contributing agencies using jackknifing techniques to estimate a standard error.

#### 2.1.1 Sources of Variation in UCR Data

Consistent themes in research on UCR data are variation and error and their sources. As early as the mid-1960s, the "dark figure" of unreported crime was a point of concern regarding the validity of the UCR data and what conclusions could be derived from it. Because UCR data reflects reports taken by law enforcement agencies and formally recorded, any crime that does not come to the attention of a law enforcement officer would not be captured in the statistics from the UCR Program. Early research by Biderman and Reiss (1967) focused attention on variation introduced by unreported crime (Biderman & Reiss, 1967; Boivin & Cordeau, 2011; Messner, Beaulieu, Isles, & Mitchell, 2014; Skogan, 2009). Following the concern of unreported crime, later research has delved into the factors that affect whether a victim of a crime reports it to the police. Research supported that the perception of how serious the crime is by the victim is the primary factor in the

---

<sup>1</sup> There are examples of states that have mandated that law enforcement agencies report crime data to a state agency, but that mandate is not from the national UCR Program.

individual decision to report it to the police, though it may be difficult to establish broad patterns given the personal nature of the decision (Gove, Hughes, & Geerken, 1985; Hart & Rennison, 2003). Little research has been done on any regional patterns in victim reports to the police. Regional patterns are difficult to assess primarily because the National Crime Victimization Survey, the only national survey of victimization in the United States, cannot provide estimates for areas smaller than the four United States regions.

However, there are areas other than unreported crime where variation and error can be introduced to UCR data as a function of the actions of law enforcement rather than the victim. The premise of the UCR Program is to provide a consistent classification structure for law enforcement to use despite the different laws and statutes that exist from one location to the next. This classification system is itself subject to interpretation by law enforcement and should be considered part of the overall error structure of reported crime (Nolan, Haas, & Napier, 2011). For example, Rosenfeld (2007) found that violence among intimate partners and family tends to be “up-classified” from simple to aggravated assault because of the changing attitudes of police on the problem of domestic violence. Beyond classification, the discretion of law enforcement officers and agencies to report reports of crime can also affect data introduced into the UCR Program (Boivin & Cordeau, 2011).

After data has been recorded and classified, the handling of the data can also affect variation in UCR data. Given that one of the purposes of the UCR Program is to provide an accounting of rates and trends for the nation, regions, and states, UCR data are aggregated into larger groups of data. The very act of aggregation by spatial units affects the interpretation of those aggregates; a problem also referred to as the Modifiable Areal Unit Problem (MAUP) in spatial analysis (Rengert & Lockwood, 2009; Wong, 2009). Maltz and Targonski (2003) found that UCR data becomes less stable at smaller units of geography. The aggregated data tend to follow some mathematical principles (i.e., the laws of natural numbers), while the smaller areas and disaggregated data do not (Hickman & Rice, 2010).

Other agency-level factors that affect the measurement of crime and the variation in UCR data center on its voluntary nature. Law enforcement agencies are not required to submit UCR data to the FBI. This can lead to “holes” in the data for particular items and reports (Addington, 2004) or for periods of time (Michael D Maltz, 2006). To provide national, regional, and metropolitan estimates, the UCR Program imputes this missing data based upon a set of geographic and population-based auxiliary variables. This approach has come under increasing scrutiny, however, given mounting evidence that these auxiliary variables are not adequate to construct reliable imputation models (Lott & Whitley, 2003; J. P. Lynch & Jarvis, 2008; Michael D Maltz, 2006; Tabarrok, Heaton, & Helland, 2010).

In their (2008) article, Lynch and Jarvis analyzed the current auxiliary variable used by the UCR Program in its mean substitution approach to the imputation of missing agencies. The current method of stratification is based on a combination of agency size as measured by population served, metropolitan status, and geography (Cynthia Barnett-Ryan, 2007). When Lynch and Jarvis decomposed the variance of the crime reports within each of the strata, they found that there was significantly more within-group variance than between-group variance. This means that the predictive power of the population group is low and would not be a suitable auxiliary variable for an imputation method (2008). Given the nature of how law enforcement data is recorded and its close ties to the way law enforcement agencies execute their duties, more could be done to bring the relationship of police recording and reporting of crime into the modeling of crime data and imputation.



### 2.1.2 Law Enforcement Agencies

In the United States, the more common types of police forces with primary jurisdiction are municipal police and county sheriff's offices (US Department of Justice, 2012). Police departments usually have primary jurisdiction over incorporated areas, while sheriffs' offices typically serve in unincorporated areas of a county. The duties of police departments tend to be rather consistent from one location to the next — especially in larger metropolitan areas. However, the duties of the sheriff may range from limited or restricted- service to full-service law enforcement offices. These two types of law enforcement agencies have been in existence in one form or another since early United States history (Vila & Morris, 1999). These types of agencies— especially large metropolitan agencies—dominate the research literature (Falcone, Wells, & Weisheit, 2002). It is less likely to find discussions that investigate small-town agencies or specialized agencies and their practices (Falcone et al., 2002; Pelfrey, 2007).

Research suggests that small-town police have a much broader array of tasks than their metropolitan counterparts. This is often linked to the close ties that members of the small-town police force have with members of their community. Conversely, federal law enforcement and other state law enforcement bureaus are often less responsive to their immediate community and are typically more influenced by their organizational culture. Larger urban agencies have seen a push towards specialization (Falcone et al., 2002; Pelfrey, 2007). At the same time, there has been a lack of research on the nature of specialized agencies (such as campus police, park police, or conservation police) that assume law enforcement responsibilities in areas that are also served by primary law enforcement agencies (Falcone et al., 2002; O'Connor Shelley & Crow, 2009; Paoline III & Sloan III, 2003; Pelfrey, 2007). The concurrent jurisdiction of multiple agencies creates a complex situation for reporting crime for a specific location.

As new types of agencies have been created, the duties of both traditional and specialized agencies have evolved. Many of these agencies, such as campus police, operated initially as security forces without full law enforcement powers. Since the mid-twentieth century, specialized police agencies have evolved in their duties to include training at state law enforcement academies and full law enforcement powers. However, relatively little has been studied or written on the subject of the roles or history of these specialized agencies (Kelling & Moore, 1988; Marvell & Moody, 1996; Vila & Morris, 1999; Williams, Murphy, Zinn, Stewart, & Moore, 1990).

In the evolution of these specialized agencies from security forces and private entities into public, professional law enforcement, many have come to model themselves after traditional city and county police and sheriffs. These developments are primarily seen in their organizational structure that mimics the hierarchical management structure that is often described as para-militaristic (Kelling & Moore, 1988; Moore, Mark H. and Kelling, 1983; Vila & Morris, 1999; Weiss, 2011). Specialized agencies, however, face unique challenges depending on the community they serve. In addition, many agencies work within limitations set not only by geographic constraints but also by the type of offenses that they handle. These limitations have the potential to influence the type and amount of crime information recorded for the UCR Program resulting in potential bias and uncertainty.

### 2.1.3 Agency Information in the UCR Program<sup>2</sup>

As a starting point for the discussion of agency type and activity and their relationship to crime reporting, it would be useful to explain the current manner that the UCR Program describes agencies. The UCR Program approaches the description of law enforcement agencies in three primary ways. The first method would be

---

<sup>2</sup> This analysis uses 2012 Uniform Crime Reporting Program data. In 2013, the UCR Program implemented a major change in definition to forcible rape that make comparisons from one location to the next difficult.

based on agency function and type. A little over 50 percent of the agencies on file with the UCR Program as of 2012 are described as a city police department, while approximately 14 percent represent the unincorporated areas of the county served by a sheriff's department. These two types of agency typically represent primary law enforcement for the population of a given area. The additional categories of colleges/universities, state police, other agencies, other state agencies, and tribal agencies make up about 16 percent of agencies that represent specialized agencies. The remaining 17 percent are agencies that are "covered-by" another agency. These are agencies that agree to have another agency report their crime information to the UCR Program on their behalf. (See Table 2-1).

Table 2-1. Number and percentage of agencies by type and population category, 2012

Agency Type, 2012	Total		Zero-Population Agencies		Most-in-population Agencies		All Other Agencies	
	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
Covered-by	3,757	17.0%	1,968	35.1%	4	0.9%	1,785	11.2%
Cities	11,630	52.7%	3	0.1%	291	62.4%	11,336	71.0%
Counties	3,032	13.8%	123	2.2%	164	35.2%	2,745	17.2%
Colleges/ Universities	748	3.4%	748	13.3%	-	0.0%	-	0.0%
Other agencies	1,134	5.1%	1,023	18.2%	7	1.5%	104	0.7%
Other state agencies	495	2.2%	495	8.8%	-	0.0%	-	0.0%
State Police	1,057	4.8%	1,057	18.8%	-	0.0%	-	0.0%
Tribal	195	0.9%	195	3.5%	-	0.0%	-	0.0%
Total	22,048	100.0%	5,612	100.0%	466	100.0%	15,970	100.0%

An additional method of classifying agencies is in broad categories based upon population characteristics or a "population category." The "population category" is used primarily within the UCR Program for managing data collection efforts more than how the data is presented to the public. This concept has two primary groups: *zero-population* agencies and *most-in-population* agencies, while all other agencies beyond these two fall into a remainder group. *Zero-population* agencies are those agencies who share an overlapping area of geography and its inhabitants with another agency. By UCR policy, the inhabitants of an area such as a city or county are primarily assigned to a city police department or a county's sheriff's department. To avoid double-counting, the other agencies that might have concurrent jurisdiction are assigned no or "zero" population. In contrast, those agencies with 100,000 or more in assigned population are operationally assigned the status of *most-in-population* (MIP). Any remaining agencies that include medium to small primary law enforcement agencies fall into an *All Other Agencies* category. Of *zero-population* agencies, 62.6 percent are one of the specialized law enforcement groups of colleges/universities, other agencies, other state agencies, state police, and tribal agencies (Table 2-1). In contrast, only a small portion of *most-in-population* agencies or *all other agencies* fall into one of the specialized categories—namely, *other agencies*.

In addition to these two schema, agencies are also assigned to a population group. While the concepts of agency type and population category impact ways in which UCR data are displayed and published, it is a third category, population group, that serves as the basis of the current imputation method of UCR offense data described later in this paper. The population group is used to categorize an agency by the size of the

population being served, whether it serves a population in an incorporated (city) or unincorporated (county) area, and whether it is part of a Metropolitan Statistical Area. There is also a provision for identifying state police, as well. For example, agencies in the Population Group 1A serve cities with 1,000,000 or more inhabitants, while Population Group 8A serves counties (unincorporated areas) that are not part of a metropolitan statistical area (MSA) with 10,000 or more in population (See Table 2-2).

Table 2-2. Heterogeneity of population group by agency type, 2012

Population Group, 2012	Covered-by	Agency Type, 2012							Total	Agency count
		Cities	Counties	Colleges/Universities	Other agencies	Other state agencies	State Police	Tribal		
Possessions	100%	-	-	-	-	-	-	-	100%	4
Cities 1,000,000 or over	-	100%	-	-	-	-	-	-	100%	10
Cities from 500,000 through 999,999	-	100%	-	-	-	-	-	-	100%	25
Cities from 250,000 through 499,999	5%	95%	-	-	-	-	-	-	100%	43
Cities from 100,000 through 249,999	1%	99%	-	-	-	-	-	-	100%	217
Cities from 50,000 through 99,999	1%	99%	-	-	-	-	-	-	100%	483
Cities from 25,000 through 49,999	2%	98%	-	-	-	-	-	-	100%	908
Cities from 10,000 through 24,999	3%	97%	-	-	-	-	-	-	100%	1,979
Cities from 2,500 through 9,999	8%	92%	-	-	-	-	-	-	100%	4,380
<b>Cities under 2,500</b>	<b>25%</b>	<b>55%</b>	<b>10%</b>	<b>1%</b>	<b>4%</b>	<b>3%</b>	<b>3%</b>	<b>100%</b>	<b>7,295</b>	
Non-MSA Counties 100,000 or over	-	-	33%	-	67%	-	-	-	100%	3
Non-MSA Counties from 25,000 through 99,999	-	-	94%	-	6%	-	-	-	100%	298
Non-MSA Counties from 10,000 through 24,999	-	-	97%	-	3%	-	-	-	100%	677
<b>Non-MSA Counties under 10,000</b>	<b>23%</b>	<b>-</b>	<b>41%</b>	<b>-</b>	<b>20%</b>	<b>2%</b>	<b>14%</b>	<b>-</b>	<b>100%</b>	<b>2,418</b>
<b>Non-MSA State Police</b>	<b>10%</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>74%</b>	<b>1%</b>	<b>14%</b>	<b>-</b>	<b>100%</b>	<b>97</b>
MSA Counties 100,000 or over	-	-	97%	-	3%	-	-	-	100%	168
MSA Counties 25,000 through 99,999	-	-	93%	-	7%	-	-	-	100%	500
MSA Counties from 10,000 through 24,999	-	-	96%	-	4%	-	-	-	100%	290
<b>MSA Counties under 10,000</b>	<b>43%</b>	<b>-</b>	<b>9%</b>	<b>-</b>	<b>19%</b>	<b>8%</b>	<b>21%</b>	<b>-</b>	<b>100%</b>	<b>2,147</b>
<b>MSA State Police</b>	<b>24%</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>43%</b>	<b>3%</b>	<b>30%</b>	<b>-</b>	<b>100%</b>	<b>106</b>
<b>Total</b>	<b>17%</b>	<b>53%</b>	<b>14%</b>	<b>3%</b>	<b>5%</b>	<b>2%</b>	<b>5%</b>	<b>1%</b>	<b>100%</b>	<b>22,048</b>

NOTE: Selected rows indicate population groups with significant heterogeneity of agency type

How does the UCR Program handle assigning agencies with no population (zero-population agencies) in a schema primarily built around population? While some population groups imply quantified population, *zero-population agencies* are officially assigned to one of five categories: cities under 2,500; Non-MSA Counties under 10,000, Non-MSA State Police; MSA Counties under 10,000; and MSA State Police. The result is that these five categories capture much more variation in agency type when compared to others. Table 2-2 is provided to show the intersection of two of these classification schema: agency type and population group. The inclusion of the *zero-population agencies* in these population groups has resulted in a more heterogeneous make-up compared to the other population groups. As will be discussed further in this paper,

this pattern becomes a critical issue in the choice to use population group in the imputation of missing data in the UCR Program.

#### 2.1.4 Missing Data Imputation and Other Uses of Agency Information in the UCR Program

As has been stressed, agency characteristics feature prominently in all aspects of management of UCR data. These characteristics are used in the display of the data, such as the tables that are included in the annual publication of *Crime in the United States* (US Department of Justice, 2012), as well as the imputation techniques used by the UCR Program to account for missing data. For publication purposes, the UCR Program features national and subnational estimates of the original Part I offenses. The original Part I offenses are murder and nonnegligent manslaughter, rape, robbery, aggravated assault, burglary, larceny-theft, and motor vehicle theft. The current methods for imputing missing UCR data by the FBI follow two paths. The first is applied to agencies that are viewed as incomplete or partial reporters. If an agency has provided at least three months of crime data, then the reported data are pro-rated to a full twelve months of data without regard to seasonality. For example, if an agency reports five months of data, an average number of crimes per month is calculated and then multiplied by twelve to arrive at an estimate for the year. For the second method, if an agency provides less than three months of crime data or no data at all, then the crime rate per capita for agencies of similar size within either the same state, geographic division, or region is multiplied by the agency's population in order to arrive at an estimate for the year. In essence, this is a population-adjusted mean substitution approach to imputation (Barnett-Ryan, 2007).

For the offense of arson, however, the data are treated differently. The collection of arson incidents was added as a permanent part of the UCR Program in 1980. However, for many years, the collection of arson data was relatively spotty in comparison to the other UCR Index offenses. Nonetheless, an analysis of 2012 data shows that arson data have been reported by almost the same number of agencies as the other UCR offenses (US Department of Justice, 2012). However, the UCR Program has never produced national estimates of reported arsons. This reluctance is primarily due to the concern that reported data might be biased and a general concern that the current imputation methods are not appropriate for producing national estimates of arson rates ("Arson," 2012).

Since the late 1950s, the field of statistics, however, has evolved in its understanding and treatment of missing data (Rubin, 1976). Rather than ignoring or excluding incomplete information, modern methods of imputation emphasize managing the bias that can be introduced into the data when portions are missing. Part of that process is ensuring that there is an understanding of the reasons why the data are incomplete. However, the UCR Program still treats missing data as though it were the same as the reported data in both volume and character. The UCR Program has never systematically studied the nature of nonresponse of law enforcement agencies and the potential bias that this may introduce into the estimates using the current procedures.

Table 2-3. Percentage of imputed data in national estimates for Part I crimes, 2012

	Reported	Imputed
Murder and Non-negligent Manslaughter	98.8	1.2
Rape	97.0	3.0
Robbery	98.9	1.1
Aggravated Assault	98.0	2.0
Burglary	97.2	2.8
Larceny-theft	96.9	3.1
Motor Vehicle Theft	98.0	2.0

As has been established, there has been very little study of the types of missingness exhibited by UCR data or its statistical importance. Based on 2012 data, the percentage of the estimated values for Part I offenses that are based on actual reports range from 97.2 percent for burglary to 98.9 percent for robbery and motor vehicle theft (See Table 2-3). Instead, the UCR Program has traditionally described agency missingness in terms of the quantity of missing data. The first group is composed of agencies that are essentially or completely absent from a year’s data collection—so-called unit missing. If an agency provides less than three months of crime data, it is completely absent for that year and grouped with agencies that provided no information at all. The second group includes the agencies that provided anywhere from three to eleven months of crime data to the UCR Program. These agencies reflect a pattern of item missing rather than unit missing. The reason behind the construction of these two categories of missingness in UCR is that each group receives a different imputation technique. However, the description of the amount of missingness does not necessarily provide any explanation as to the underlying mechanism of the patterns of missingness.

Given the scant amount of missing data from a national aggregate, the question remains why we would want to be concerned about potential bias in imputed data. The problem lies in that the UCR Program also publishes estimates for smaller areas such as regions, geographic divisions, states, and some MSAs. As can be seen in Figure 2-1, the amount of missing data by county is not evenly distributed around the country. There are counties have very few agencies of low population providing crime data, as well as counties that are completely missing. Also, it is apparent that some of the variation in reporting is more common in some states rather than others—especially for states that lack state-level aggregators (i.e., state UCR Programs) such as Mississippi. In addition, robust reporting and the relatively low impact of missing data has been established through almost 90 years of reporting history. However, the upcoming transition of the UCR Program to incident-based reporting will create the need to understand past reporting and variation of agency reports of crime.

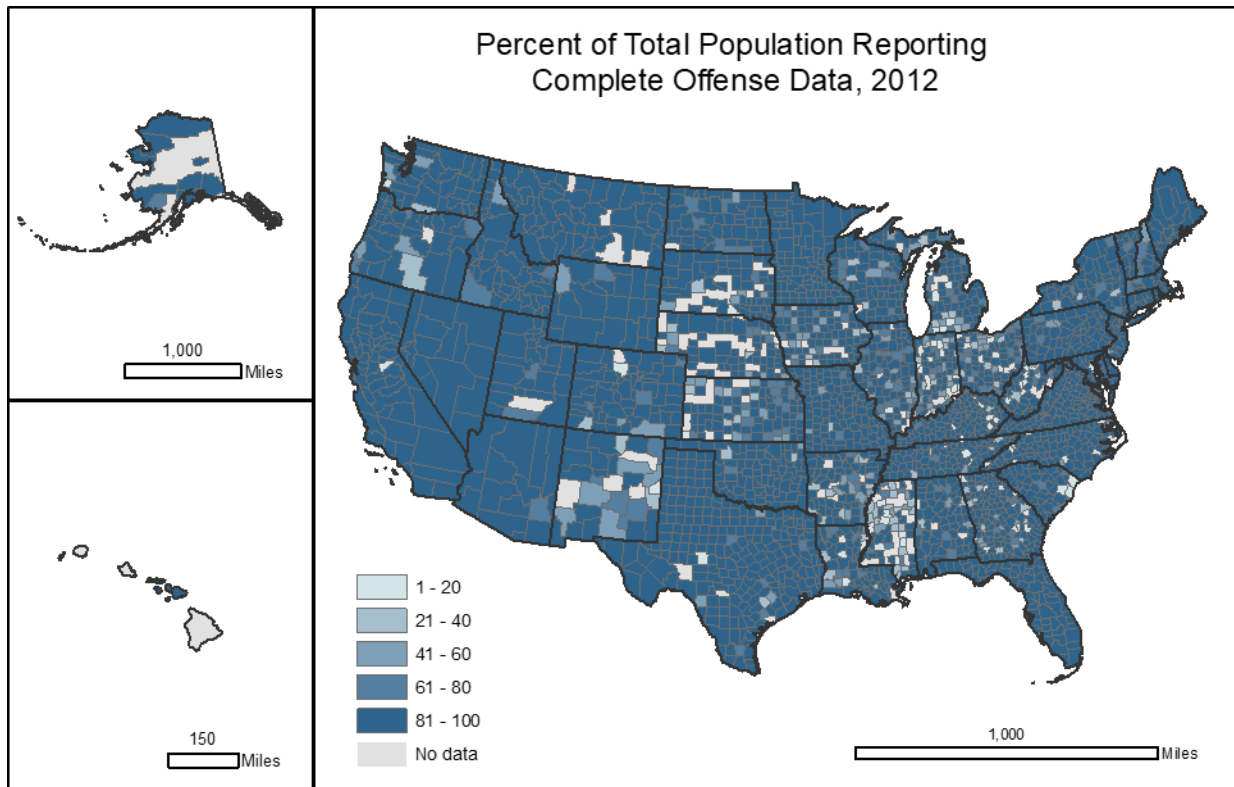


Figure 2-1. Percent of population represented by agencies reporting complete offense data to the UCR Program

A brief analysis of the data reveals that the *MIP agencies* are more often than not providing complete data. Because of the role larger agencies play in multiple publications and the overall influence of the crime counts on national measures, the MIP agencies are given very specific attention to ensure that their data have been reported. *Zero-population agencies* and agencies with small populations are responsible for much of the unit and item missing data (see Table 2-4). In all cases, it is more likely for agencies not to provide any data (unit missing) rather than providing incomplete data (item missing). Almost half of all non-reporting agencies in 2012 were *zero-population* agencies. The reasons behind this tendency towards non-reporting by *zero-population* agencies may be as diverse as the agencies themselves. Little research has been conducted to explain the lack of participation by specialized agencies, but it is likely connected to management practices valuing the contribution of crime counts by larger agencies. However, as the UCR Program transitions to incident-based reporting of a broader array of offenses, the contribution of crime measures from specialized agencies may have a greater impact.

A critical problem surrounding this pattern is found in the current method for imputing missing data for *zero-population* agencies. Because the missing data imputation method uses population to arrive at an estimated count of reported crimes, the current imputation process always estimates that no crimes were reported to non-reporting *zero-population* agencies; even though it is known that many *zero-population* agencies do have crimes reported to them. While this process introduces bias into the national estimate of reported crimes, there is almost no understanding of the relative impact of this process. The proportion of *zero-population* agencies has changed from 0.8 percent of total agencies in 1960 to 16.5 percent in 2012. The original decision not to include *zero-population* agencies in the imputation method would have had minimal impact. However, the proportion of *zero-population* agencies has grown to the point where the UCR Program can no

longer ignore them. In addition, the distribution of zero-population agencies is also not evenly divided across the nation. As is illustrated in Figure 2-2, these specialized agencies appear to be more common for the Northeastern, South Atlantic, and Western states.

Table 2-4. Reporting pattern by population category for original Part I offenses and arson, 2012

		Reporting Patterns for original Part I Offenses, 2012							
		Missing		Incomplete		Complete		Total	
		Count	Percent	Count	Percent	Count	Percent	Count	Percent
Population Category, 2012	Zero-Population Agencies	2,603	46.4	259	4.6	2,750	49.0	5,612	100.0
	"Most in Population" Agencies	4	0.9	1	0.2	461	98.9	466	100.0
	All Other Agencies	3,037	19.0	807	5.1	12,126	75.9	15,970	100.0
Total		5,644	25.6	1,067	4.8	15,337	69.6	22,048	100.0

		Reporting Patterns for Arson, 2012							
		Missing		Incomplete		Complete		Total	
		Count	Percent	Count	Percent	Count	Percent	Count	Percent
Population Category, 2012	Zero-Population Agencies	2,541	45.3	397	7.1	2,674	47.6	5,612	100.0
	"Most in Population" Agencies	9	1.9	5	1.1	452	97.0	466	100.0
	All Other Agencies	3,291	20.6	1,025	6.4	11,654	73.0	15,970	100.0
Total		5,841	26.5	1,427	6.5	14,780	67.0	22,048	100.0

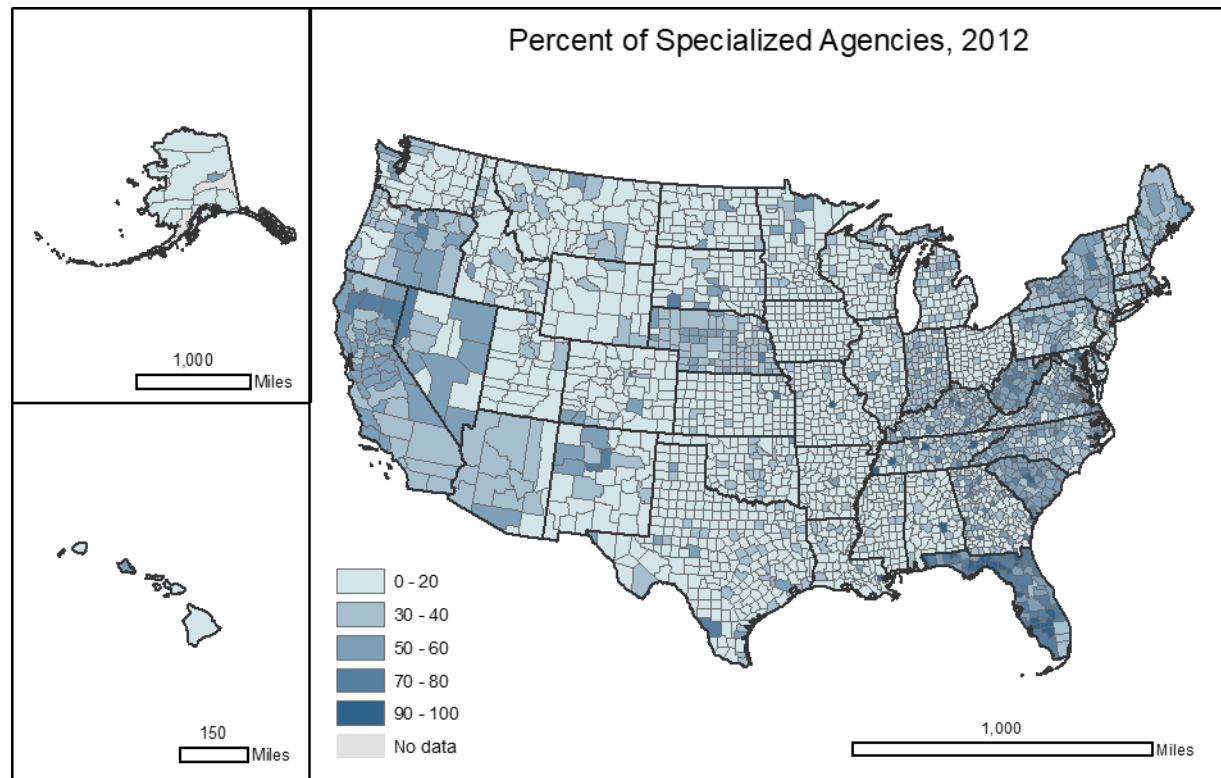


Figure 2-2. Specialized law enforcement agencies by county as a percent of total law enforcement agencies, 2012

### 2.1.5 Nature of Missing Data in the UCR Program

A better understanding of the nature of the missing data and those agencies that do not report will reveal the potential impacts of missing data on the validity of national crime estimates produced by the UCR. Rubin (1976) introduced a classification method of missing data to describe both the mechanism of the missingness and its potential impact. The first major category of missing data is data that are “Missing Completely at Random” (MCAR). Data that are MCAR exhibit no structured pattern to the quality of missingness. In essence, MCAR can be described much in the same way as a simple random sample. Because of its random nature, the impact of MCAR is minimal, unbiased and can be effectively ignored. While MCAR is the ideal situation when it comes to missing data, it is not often observed except in areas when it is purposefully introduced into the research design itself.

A second possibility is for data to be “Missing at Random” or MAR. Data that are MAR are not subject to the random process described above as MCAR. Instead, MAR exhibits signs of systematic bias; however, the pattern can be associated with known information about each case. This is roughly equivalent to a stratified random sample. The key is to discover the underlying reason for, or at least a variable that seems to be highly correlated with, the missingness. Once the covariate with missingness is accounted for, the potential bias is controllable (Rubin, 1976).

Finally, there are situations when the data are systematically missing, but by means that are unknown. These data are described as “Not Missing at Random” or NMAR. Because it is not possible to account for the bias that is introduced by NMAR data, these data have the greatest potential to affect the results of analysis adversely. The key to dealing with missing data is to try to transform any NMAR process to a MAR process through a careful exploration of the data set. The reality is that all datasets start as NMAR until the patterns of missingness are analyzed to ascertain the mechanism for the missing data. Once that mechanism can be described, it can be introduced into either an analytical model or an imputation model to avoid (or at least reduce) biasing the results. Metadata can be crucial in identifying this process.

## 2.2 DISCUSSION AND RESULTS

### 2.2.1 Variation within the Reporting Agencies

The current imputation method in the UCR Program presumes that missing data are a MAR process that can be accounted for by population, population group, and geography. However, there is evidence that these three elements do not effectively capture all the underlying correlates of missingness for UCR data (Lynch & Jarvis, 2008; Maltz, 2006). Based upon the initial descriptive analysis of missing data, the type of agency—especially *zero-population* agencies—appears to be related to reporting levels of UCR data. Given the variation in the types of agencies within population groups capturing zero-population agencies, the variation and heterogeneity among *zero-population* agencies could interfere with the application of the mean-substitution imputation technique (See Table 2-2).

### 2.2.2 Coefficient of Variation

To explore the amount of variation that may exist within the given categories of agency type and population group, the coefficient of variation (CV) was calculated using those agencies that submitted a full year of data for 2012. The coefficient of variation (also known as the relative standard deviation) is arrived at by taking the ratio of the standard deviation to the mean. The primary advantage of using the CV is that it is unit-less or standardized across different units of measurement. As a standardized measure of dispersion, the CV can be used to compare the amount of variation among different variables even in situations where the measurements are different (Institute for Digital Research and Education, n.d.). While, the measurements of



crime categories are each counts, the way in which these crimes are counted may introduce a difference in the underlying measurement. Within the UCR Program, law enforcement agencies are requested to count only one offense per victim of any crime against person (e.g., murder and non-negligent manslaughter, rape, and assault). Agencies count one offense per criminal transaction at approximately the same time or place (also referred to as an incident or distinct operation) for the crimes against property of robbery, burglary, larceny-theft, and arson. Motor vehicle theft is counted as one offense per vehicle stolen or attempted to be stolen.

Figure 2-3 and Figure 2-4 show the CV for each of the seven offenses appearing in the original Part I offenses collection, as well as Arson, for each of the population groups. The most notable finding is that the population groups that represent the smallest agencies and *zero-population* agencies have the largest variation across all offense categories. This finding is not necessarily surprising given the heterogeneity of those population groups. (See Table 2-3.)

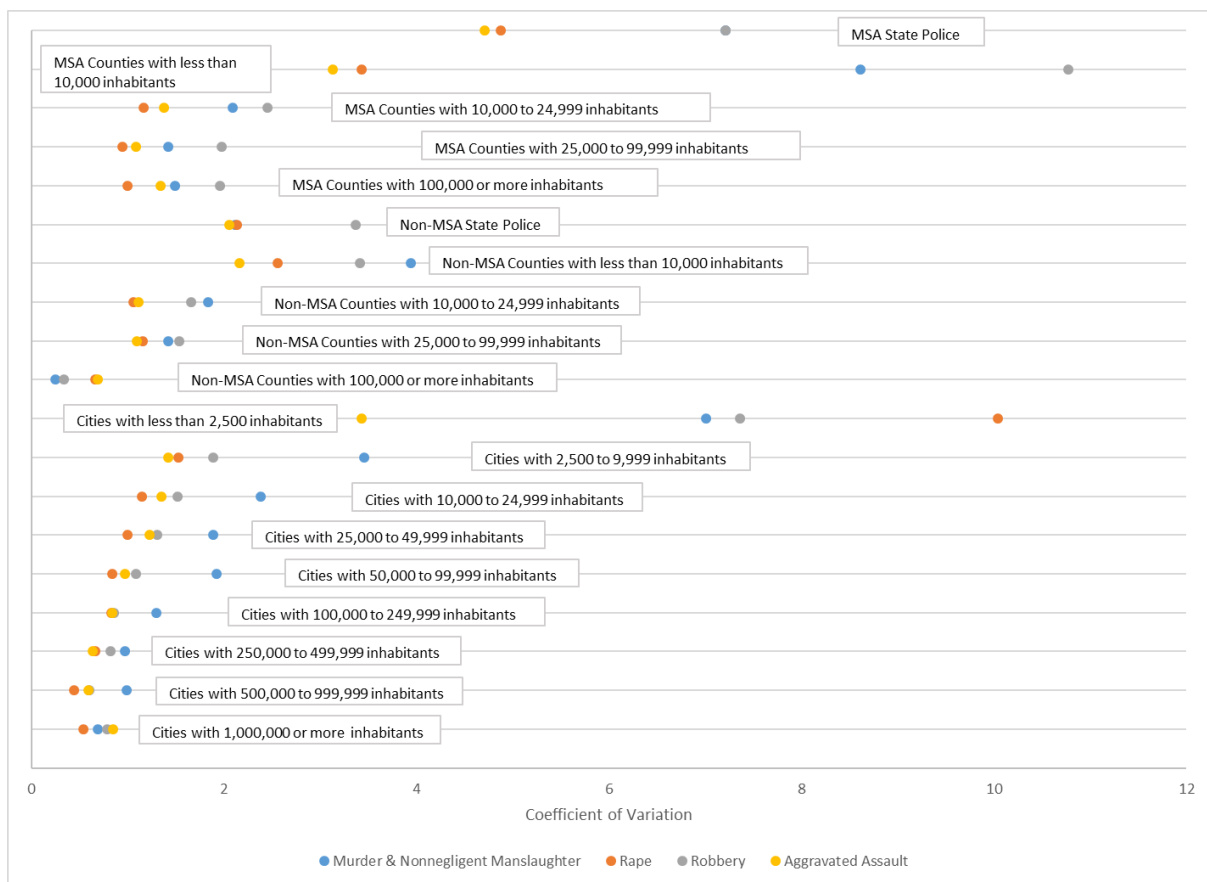


Figure 2-3. Coefficients of variation within violent offenses, 2012

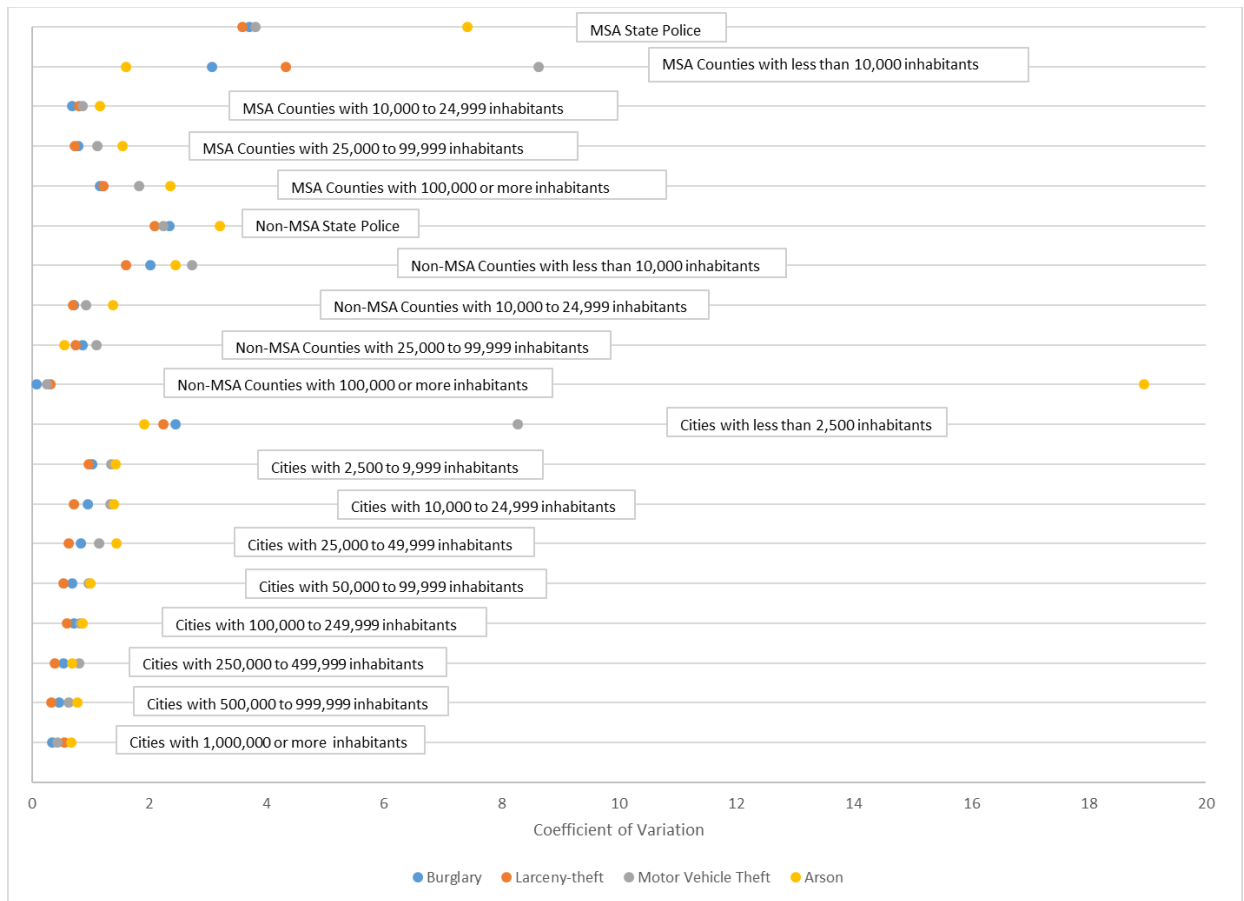


Figure 2-4. Coefficients of variation within property offenses, 2012

These findings raise the question of how to best subdivide the law enforcement agencies that report data to the UCR Program, and whether accounting for the *zero-population* status may account for additional factors for the process of missing offense reports. A simple comparison of measurable variation between a stratification schema with the *zero-population* agencies and without may provide some indication of the possible impact of introducing population category into the model. While the “unit-less” quality of the CV has its strengths, there is no obvious statistical test to assess whether any measurable improvement in the CV is more appropriately attributed to random variation in the variable. Parametric statistics that would address measures of central tendency rely upon theory and known distributions to inform a decision of whether differences are “statistically significant.” Being nonparametric, the CV does not have a theoretical distribution to provide this insight. Also, the UCR dataset does not have the benefit of a structured sampling strategy that is usually part of the underlying assumptions tied to parametric statistics. For these reasons, this paper will explore the benefits of using a resampling method, jackknifing, to decide on the significance of the differences in the CV. This paper will also extend the use of the CV by calculating the average CV across a vector of CV values.

### 2.2.3 Jackknifing

Bootstrapping, jackknifing and other resampling methods “allow us to quantify uncertainty by calculating standard errors and confidence intervals and performing significance tests” (Hesterberg, et al. 2003, p.18-4). This paper will use a jackknife variance estimator to introduce a method to calculate a standard error. First introduced by Quenouille (1949, 1956), the jackknife was further exploited by Tukey (1958) as a method to remove bias from the estimator of variance and interval estimation (Boos & Osborne, 2013). The jackknifed

standard error is a standard error estimated "computed from a sample of size N of vectors; each vector computed from a computer-generated random sample" (Boos & Osborne, 2013, p. 2). Boos and Osborne provide an easy to replicate set of R code to encourage the use of the "leave 1 out" jackknife to calculate the standard error (2013). The jackknifed standard error allows for a determination of whether any measurable reduction in variation could be statistically significant.

Central to this paper is the idea of a stratum of "like" agencies used to calculate the per capita rate used in the imputation missing agency data and whether the model captures the MAR process. Current UCR practices for imputation focus initially on the agency population group and state. When the geographic extent of the state does not have another similar agency for calculating a per capita rate for mean substitution, the per capita rate is based upon agencies within a larger geographic area. Initially, one of nine geographic divisions is used or then finally expanded to the larger area of one of the four geographic regions. All three levels of geography will be explored in this paper to assess the impact of geography on the overall variation of the data.

Within each of the three geographic levels, the average will be calculated for each vector of CVs of crime offense values based upon two different stratification methods. The first stratification method replicates the current process in UCR where only the population group and geography are represented. The second stratification method will remove *zero-population* agencies for comparison. Finally, the jackknifed standard error will be calculated for each of these vectors to establish a 95 percent confidence interval around each average CV value. (See Table 2-5 for counts in each vector based on stratification method.)

#### 2.2.4 Results

The first stratification method assessed includes a combination of state and population groups to arrive at a set of strata that has minimal internal variation within each group. For 48 of the 50 states, UCR data is aggregated by state agencies responsible for compiling crime data in a state repository across a common standard. The state UCR Programs then forwards state crime information to the FBI's national UCR Program applying the national standards. For this reason, the state is the smallest geographic area used in this research and has the benefit of incorporating any significant spatial autocorrelation that might exist among agencies within this smaller geographic area based upon the aggregated reporting units. The resulting comparison between the vector with *zero-population* agencies and the vector without those agencies shows that there are overall improvements in the calculated average CV and a reduction of variation around those CVs once specialized (i.e., *zero-population*) agencies are removed. However, it is only with the property crimes of burglary, larceny-theft, and motor vehicle theft that have statistically significant decreases in average CV in the vector without specialized agencies (see Figure 2-5). While this result is promising, there is an equal concern that there are many strata that only have one or two values. For the original seven Part I crimes, 22.8 percent of the strata have one or two values (for example, there is only one agency in the state of Alaska that has between 250,000 and 499,999 inhabitants), and 23.1 percent of the strata have only one or two values for offense reports of arson. When a stratum only contained one value, the CV would be missing(See Table 2-5 in the appendix for a summary).

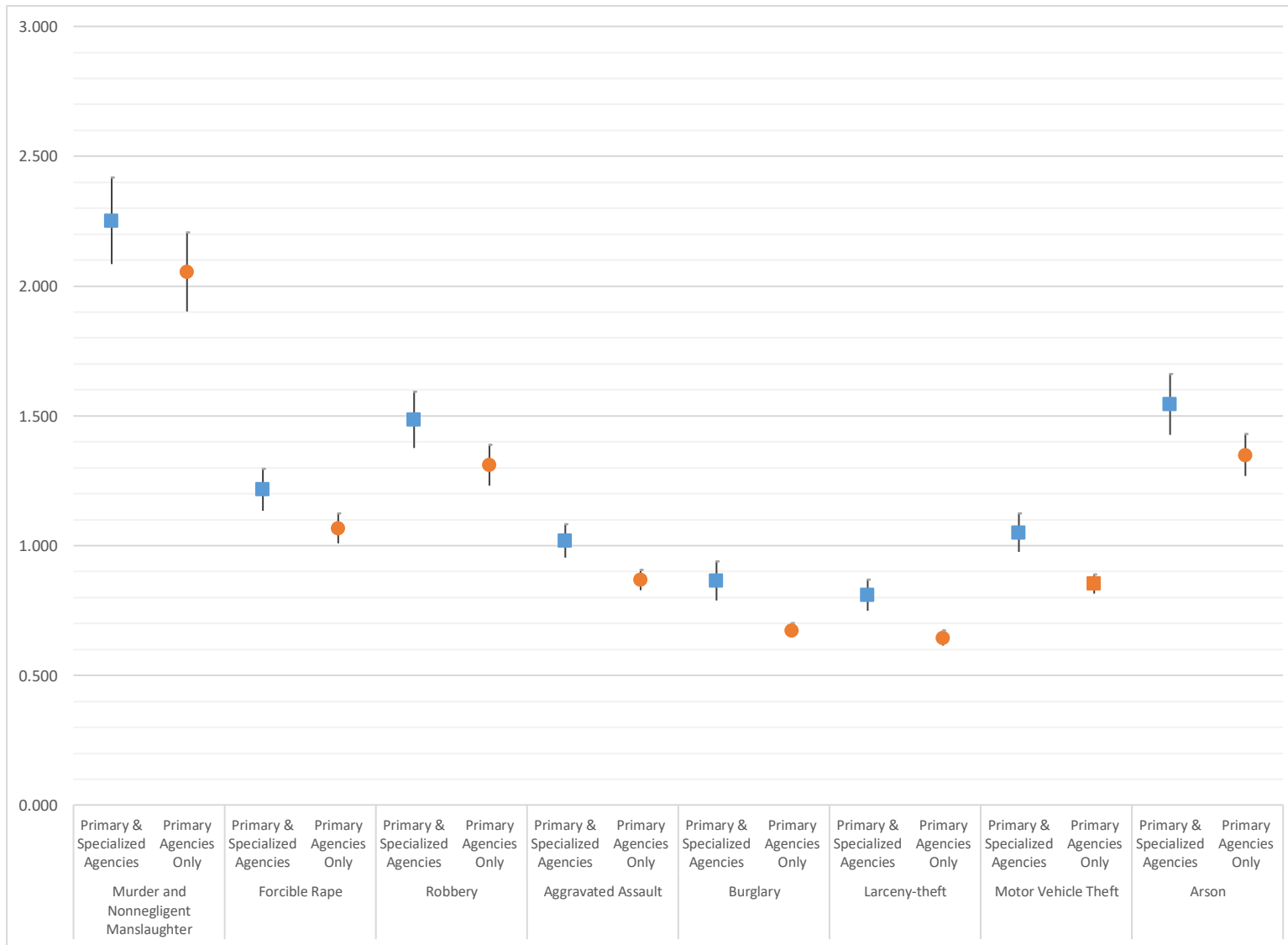


Figure 2-5. Difference in mean coefficient of variation based in strata based on state and population group using 95 percent confidence interval

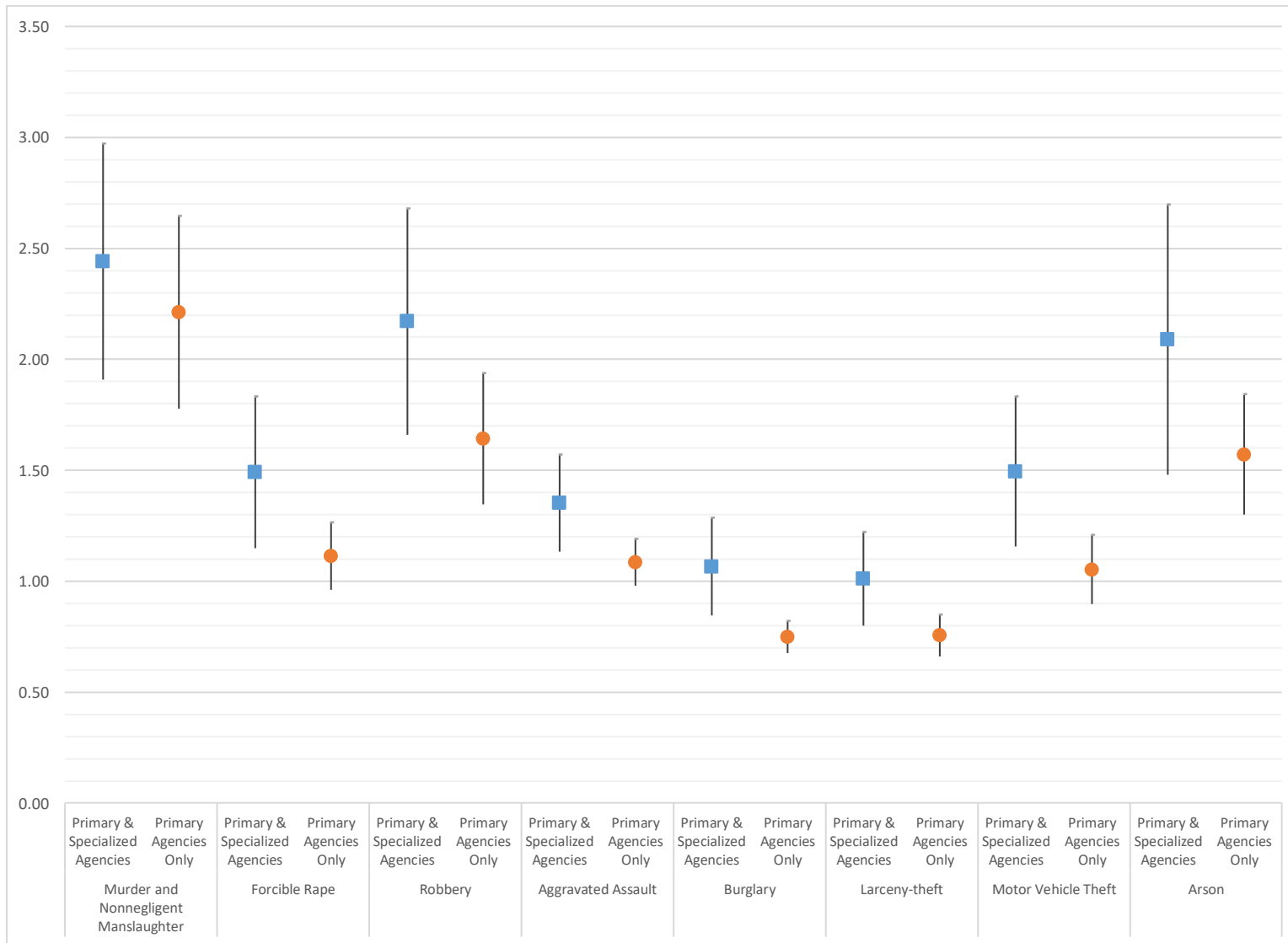


Figure 2-6. Difference in mean coefficient of variation in strata based on region and population group using 95 percent confidence interval

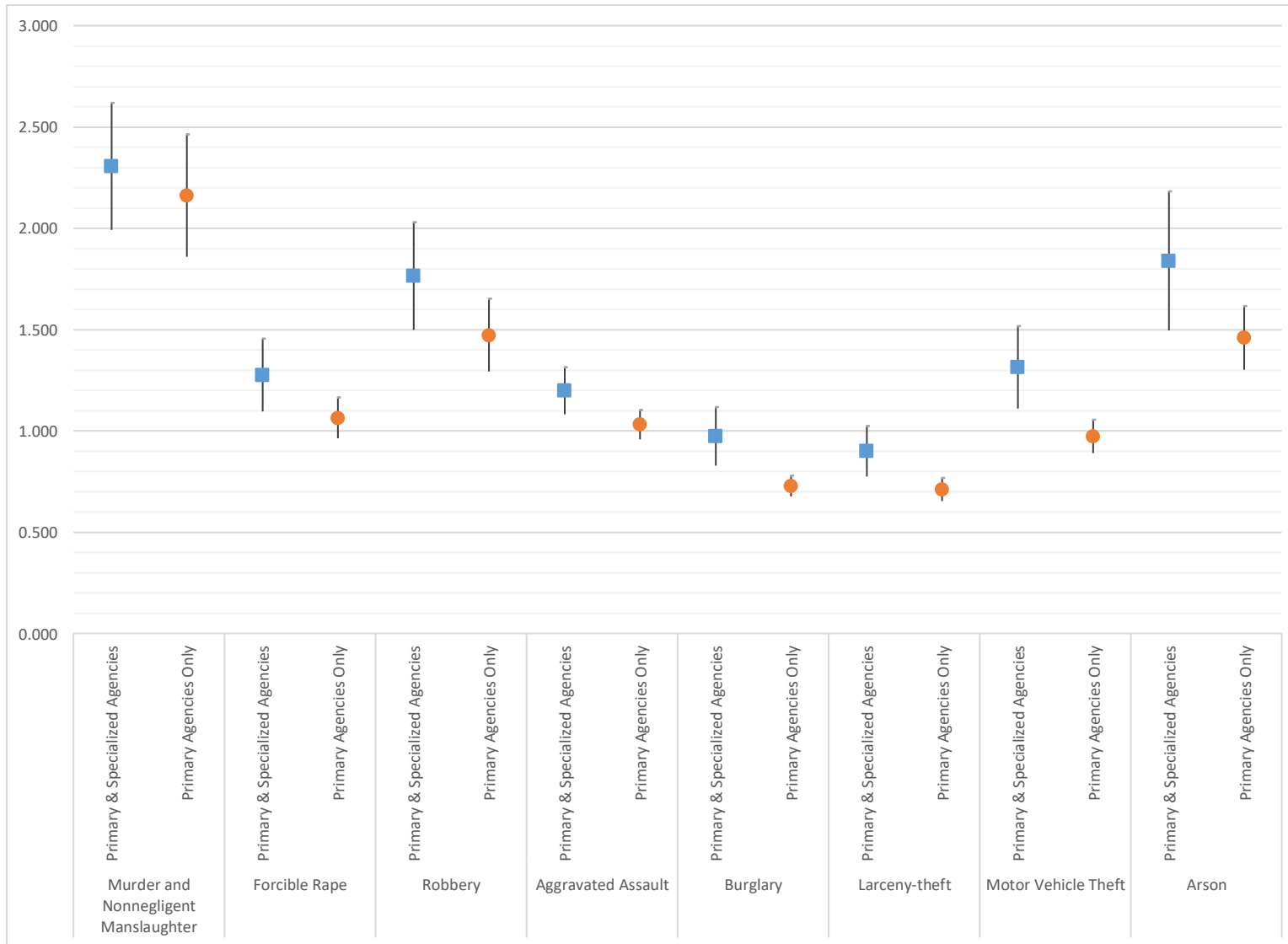


Figure 2-7. Difference in mean coefficient of variation in strata based on division and population group using 95 percent confidence interval

To address the issue of the small N in each of the strata based upon state, an alternative approach might be to use the four geographic regions used by the Federal statistical agencies (e.g., Northeast, Midwest, South, and West). Expanding out to the region increases the number of agencies in each of the strata. Under this regional approach, only 7.0 percent of the strata contain one or two values for the offense reports of the original Part I, just 4.3 percent of the strata contain one or two values for the offense reports of arson. The average CV is reduced with the elimination of the specialized agencies. However, the increased variation dilutes the measured improvement in overall fit within the strata by reducing the possible spatial effects. There is no statistically significant decrease to the CV based on the elimination of specialized agencies for any of the offense categories (see Figure 2-6).

As a possible compromise to increase the number of values in each stratum while also maintaining some captured spatial effect, the last comparison is based upon the geographic division. Much like region, the division is a consistent grouping of states used across all federal statistical areas. Each of the four regions is subdivided into two to three divisions for a total of nine divisions across the 50 states. The comparison of average CV using this stratification method replicates the results as the stratification method based upon state. Overall, there is a measurable improvement and reduction of variation across all offenses. However, it is only for the property crimes of burglary, larceny-theft, and motor vehicle theft that have statistically-significant improvements to the amount of variation within each stratum (see Figure 2-7). The strata created by this intermediate level of geography appears to strike a balance between a small enough geographic unit to capture similar experiences of crime while allowing for sufficient strata size.

### 2.2.5 Conclusion

The goal of this paper was two-fold. The first objective was to explore the impact of *zero-population* agencies on the variation within the population-based groups of law enforcement agencies. To accomplish the first objective, the second of this research was to introduce a possible new methodology for the comparison of UCR data using jackknifing techniques to produce jackknifed standard errors. The results of the analysis support the continued exploration of *zero-population* agencies and their crime-reporting characteristics—especially as it relates to property crimes. In addition, the jackknifing technique allowed for more refined conclusions to be drawn from those results by allowing both substantive and statistical significance to be assessed.

As has been discussed previously, there is strong evidence that the UCR-defined population groups have too much internal variation for measures of central tendency to reliably represent the whole. This research supports those assertions for population groups that capture most *zero-population* agencies. The UCR Program assigns those *zero-population* agencies to groups that contain the smallest populations. In general, the population groups with the smallest populations have most diversity in agency type. The members of these groups could include small municipal police departments, as well as campus police from colleges and universities, state agencies (such as a Department of Natural Resources), or transport police. Conversely, the results also point to the issue of variation being less problematic for population groups that are more homogeneous by agency type, such as those representing large metropolitan areas. The issue is less that population groups are not enough for representing the average crime experience of a group of agencies, rather that population groups may be improved upon if the confounding effects of *zero-population* agencies are accounted.

A deeper look at the results of the analysis also reveals that zero-population agencies also have the most measurable influence on property crime rates for burglary, larceny-theft, and motor vehicle theft. Small agencies, campus police, and other specialized law enforcement agencies may respond disproportionately to more reports of property crime when compared to larger agencies. There is some early evidence that smaller and specialized police agencies are more likely to refer serious violent offenses to larger agencies, such as the Sheriff's office, which may have more

personnel to deal with those serious offenses. While these anecdotal reasons may ultimately explain these results, there is not a body of research that specifically addresses the differences among these types of law enforcement agencies. Until more is known about zero-population agencies and their specialized nature, it will be difficult to account for their impact on crime data.

Finally, the jackknifing technique to calculate standard errors of the mean CV appears to address the stochastic nature of UCR data. Most research within the field of criminology would ignore the inherent uncertainty in the UCR data that can be attributed to random variation. Because it is not based on a random sample, it was thought it was either unnecessary to account for the random variation in the data or that there was no clear path to develop those measures given the unknown distribution of the data. Jackknifing provides a method to produce the variances and standard errors necessary to determine if measurable differences are likely to be attributed to random variation. The specific case of this research prevented assigning too much meaning to a measurable improvement in average CV for violent offenses. In addition to the ability to test differences between stratification methods, the measure of the jackknifed standard error provides insight into the amount of variation in a measure such as CV. The results of this analysis showed that the removal of *zero-population* agencies from stratification methods reduce the variation in average CV across all offenses. This finding means that measures of central tendency calculated within each stratum are more precise.

For future research, the results of this study support the need to focus on the variation within UCR data for both statistical and substantive reasons. Statistically, the risk of Type I error has been high for analyses using UCR data because researchers have left the idea of statistical significance unaddressed. Jackknifing and other resampling techniques show clear promise that should be replicated to understand how and when to apply these methods. Also, future research will need to concentrate on the changing dimensions of law enforcement and the impact on the data collected. Just as this article seeks to build upon the findings of Lynch and Jarvis (2008), other research is called for to look further into the differences among the various types of agencies and their policing activities. It is no longer justified to view all law enforcement through one lens given that there is an increasing diversification of law enforcement agencies and increasing specialization within each agency.



## 2.3 APPENDIX

Table 2-5. Number of Agencies within each stratum by stratification technique

	Geography	Specialized Agencies included?	N		Mean	Median	Mode	Minimum	Maximum
			Valid	Missing					
Original seven Part I	State	Yes	624	1	24.6	9	1	1	489
		No	596	0	21.1	7	1	1	489
	Region	Yes	71	1	216.0	88	1 <sup>a</sup>	1	1,718
		No	66	0	190.7	74	1 <sup>a</sup>	1	1,169
	Division	Yes	147	1	104.3	39	2 <sup>a</sup>	1	812
		No	138	0	91.2	35	4	1	812
Arson	State	Yes	611	14	24.2	9	1	1	489
		No	583	13	20.8	7	1	1	489
	Region	Yes	69	3	214.2	88	4	1	1,631
		No	63	3	192.2	84	4	1	1,085
	Division	Yes	144	4	102.6	39	4	1	741
		No	135	3	89.7	35	4	1	741

<sup>a</sup> Multiple modes exist. The lowest value is displayed

### 3 CRIME AND PLACE: DIFFERENCES IN SPATIAL RELATIONSHIP BETWEEN CALLS FOR SERVICE AND RECORDED INCIDENTS FOR MUNICIPAL AND CAMPUS LAW ENFORCEMENT

---

#### 3.1 INTRODUCTION

In this current era of open data, law enforcement agencies are encouraged to release more types and amounts of data collected typically through day-to-day operational activities to increase transparency and accountability. The alternative data sources provided through open data initiatives are tempting sources for the exploration of critical research topics. While purposeful data collections employing rigorous data collection techniques and methodologies have always been seen as the gold standard for research, researchers across many disciplines begin to consider nontraditional sources of information in efforts to explore and test concepts when fewer resources, such as personnel or monetary support, are available to support purposeful data collection (Currie, Paris, Paschetto, & Pierre, 2016; Fryer Jr., 2016; Janssen, Charalabidis, & Zuiderwijk, 2012; Molloy, 2011; Reichman, O J ; Jones, M.B.; Schildhauer, 2011). The challenge laid out for researchers—as also articulated by James Lynch in his 2017 American Society of Criminology Presidential Address (Lynch 2017)—is to understand the possibilities and limitations of these open data.

The most commonly available open data in criminal justice comes from administrative data that are a product of the work of law enforcement officers and executives. A variety of administrative data are available from law enforcement including accidents/crashes, citations, arrests, workforce, incidents and calls for service. Many law enforcement agencies have begun to participate in open data initiatives to allow for easier access by the public to their raw administrative data. The Police Data Initiative is one such example that has 130 law enforcement agencies participating from around the country contributing access to a total of 351 datasets (Police Data Initiative n.d. accessed on 5/5/2018).

The sheer volume of open data from law enforcement sources has led to increased interest in the possibilities of secondary analysis of existing digital information to explore criminal Justice concepts. While most crime statistics are also a product of administrative data collected through policing activities, they are usually not available until a formal process of compilation, editing and review is complete. Statistical data, such as Uniform Crime Reporting (UCR) data, are usually derived from official incident reports recorded by law enforcement agents and agencies. While also a product of policing work, these incident reports require some initial vetting on the part of law enforcement administrators to assign information into uniform criminal justice and statutory categories—much like the process of coding in a content analysis or other types of qualitative analysis—referred to as “classification” (US Department of Justice, 2013a, 2013b). Alternatively, calls for service (CFS) are generated by officer and public-initiated requests. These data are typically recorded in a system used to manage dispatch of law enforcement officers to a particular location. CFS records include both the location and time of the call along with a basic description of the event. CFS data can be quickly and easily generated by law enforcement agencies and is a rich source for the analysis of temporal and spatial crime patterns when the recorded incident UCR information takes longer to produce or sometimes lacks in geographic specificity.

However, if the goal for using CFS is to provide a proxy data source for the traditionally-used recorded incident data, little research has been done to test the extent to which CFS could stand in for recorded incidents—especially considering that these data sets are collected with differing goals. Recent examples of research utilizing CFS data include measuring fear of crime, collective efficacy, or other concepts closely tied to measuring police presence or the request for police presence (Alpert, Macdonald, & Dunham, 2005; Boivin & Cordeau, 2011; Conley & Stein, 2014; Junior & Muniz, 2005; Klinger & Bridges, 1997; Levitt, 1998; Nesbary, 1998; Renauer, 2012; Telep, Mitchell, & Weisburd, 2014). However, there remains an underlying question of whether CFS data have the potential to reveal more information related to recorded crime or could serve as a reasonable proxy given the differences between the two datasets.

CFS data are reflective of the self-reported nature of the information provided and a record does not necessarily reflect the situation as observed or investigated by law enforcement on arrival. These data are primarily collected for operating a computer-aided dispatch system. The key pieces of information are location, time, and a brief explanation or request to be used to manage the process of sending law enforcement to respond to a request for service by a member of the public (UCAD Project Committee, 2012). Alternatively, when and how incident information is recorded by law enforcement is influenced by the priorities and data demands of agencies that vary by agency type given that these data are used to manage the investigation and processing of recognized crimes. Past research has indicated that data recorded as CFS in computer-aided dispatch systems have significant measurement error often resulting in the misclassification of offense based upon initial accounts provided by the public rather than law enforcement (Klinger & Bridges, 1997). More recent scholarship indicates that CFS may vary depending upon the relationship between the community and police—especially in minority communities that have seen tensions and protests increase in the wake of officer-involved shootings and other negative interactions with law enforcement (Desmond, Papachristos, & Kirk, 2016; Gibson, Walker, Jennings, & Miller, 2010; Mazerolle, Bennett, Davis, Sargeant, & Manning, 2013). There are additional untested assumptions that the recording of data—both CFS and incident reports—operates in a similar fashion across all types of law enforcement agencies and at global and local units of scale.

This paper investigates whether CFS can serve as a substitute for purposefully-collected crime statistics, and how that relationship could be tested. The paper first seeks to establish the relationship—specifically, the spatial relationship—if any, between CFS and recorded crime data given the policing of a similar geographic area by distinct law enforcement agencies using both global and local statistics. Secondly, the paper will explore the spatial relationship between CFS and recorded crime data when comparing a municipal (traditional) law enforcement agency with a campus (specialized) law enforcement agency in the same area. This approach will explore the expectation that there is a close relationship between CFS and incident reports while also testing whether differing types of law enforcement agencies result in different recording patterns.

### 3.2 CRIME AND PLACE

In his 2014 Presidential Address to the American Society of Criminology, David Weisburd argued for criminology to turn its focus away from traditional explanatory models centered on the individual. Rather, he encouraged researchers consider the implications of “place” on crime and to turn towards spatial models (D. Weisburd, 2015a). In particular, Weisburd noticed a consistent finding in the literature that a disproportionately large number of crime reports tended to be associated with a limited

number of specific microgeographic units, consisting of street segments or addresses (Andresen, Curman, & Linning, 2017; LaRue & Andresen, 2015; Schnell, Braga, & Piza, 2017; Steenbeek & Weisburd, 2016; Wheeler, Worden, & McLean, 2016). These “hotspots” varied by type of offense but appeared to be surprisingly stable over time despite law enforcement activity directed to them. A key point from Weisburd’s address and others’ findings is the need to analyze geographic data at smaller, microgeographic units of scale for these persistent patterns to emerge. These findings raise critical questions about why crime might persist in particular spaces and scales. Why would some spaces consistently engender or be more vulnerable to crime compared to others and how would police recording behavior potentially affect this conclusion?

The ideas of space, place, and crime are closely associated with the law enforcement agency that collects and records the data used to measure crime. The sense of place may be perceived differently depending upon the type of agency involved. Data recorded by traditional law enforcement agencies are directly tied to a concept of place defined by the area in which a law enforcement agency and its officers have authority to operate. Traditionally, the law enforcement agency determines the extent of activity space of officers based upon administrative boundaries, such as city or county limits, that are associated with jurisdictional definitions. This definition may not be appropriate for specialized law enforcement agencies, such as campus police. Using the example of campus police, the Clery Act and the data required defines place through the activity space of students, rather than strictly jurisdictional boundaries (King, Mahaffie, & McLarnon, 2016).

With the passage of the Clery Act in 1990, campus law enforcement was mandated to meet certain criteria for recording and reporting of crime data, which are not necessarily required for other law enforcement agencies (“Clery Center,” n.d.; King et al., 2016; McCarthy, 2016). At the core, campus law enforcement must “collect, classify, and count” the offenses of murder, sexual assault aggravated assault, and the property crimes of robbery, burglary, motor vehicle theft, and arson (King, Mahaffie, and McLarnon 2016, p. 1-6). Campus agencies also gather information on hate crimes associated with the primary list of offenses mentioned above as well as the additional offenses of simple assault, intimidation, larceny-theft, and vandalism. Additionally, arrests for liquor law violations, drug law violations, and weapons law violations are included. This information is published in an annual security report and sent to the Department of Education. The daily crime log is open to review by the public, and agencies are expected to implement ways to issue crime alerts to the campus community and develop educational programs on key campus crime problems, such as dating violence or sexual assault.

In addition to requiring certain content in these reports, guidance documents provided to campus law enforcement to implement the Clery Act is careful to define the areas and boundaries for recording such data. There are distinct areas for which campus law enforcement is required to maintain records. Any crime reported on-campus or in any non-campus building owned or controlled by the institution is recorded. While the Department of Education’s concept of university-controlled areas or buildings is certainly tied to the definition of campus property, use agreements, and loans, it is equally defined by the movement of students through public property such as sidewalks, parking facilities, or bike paths. In the case of non-campus buildings and leased space, any property that staff and students have access to is subject to reporting requirements if the buildings are used in support of the university’s educational purposes. In addition, crimes occurring in public spaces or in public property adjacent to the campus frequented by students is required to be reported as part of Clery Act requirements (King et al., 2016).

### 3.3 POLICE IN A BOUNDARY ROLE

Beyond the more concrete geographic representations of boundaries such as lines separating jurisdictions, police officers, especially those working directly with the public such as patrol officers, can also be forming a boundary between the public and the organizational goals established by command staff. The boundary between these two groups is functional, and within the police force there are particular individuals who serve a role at the margins between those two groups—a boundary role (Aldrich & Herker, 1977; Klockars, 1985; Manning, 2010; Tyler, 2004). The two main functions of the boundary role are to provide information processing and external representation. These two functions represent the communication of information in two directions. The information processing function is a process by which the boundary role incumbents filter the information that is available in the environment and selects only what is seen as relevant to be passed along to the rest of the organization. The hierarchical nature of law enforcement organization sets an expectation that the command staff will establish the agency's mission (D. Weisburd et al., 2003). In turn, officers are in turn expected to represent organizational goals set by the mission to the public through their information communications outward to the public, such as press releases, social media posts, or themes surrounding organized community meetings (Aldrich & Herker, 1977; Klockars, 1985; Manning, 2010; Tyler, 2004).

The boundary role has far-reaching implications, starting with how officers observe their environment and their activity space. A patrol officers' response to her or his environment has direct consequences on how and which crimes are discovered. Once a criminal incident has been identified, the officers are responsible for recording the event's information according to organizational priorities. Additionally, which crimes are formally recorded and what information is provided on those crimes may vary from event to event, officer to officer and location to location (Aldrich & Herker, 1977). The function of external representation is the boundary role incumbent's ability to negotiate the differences between the environment and the organization's policies. An additional point of focus is the adaptability in the interaction between community members and patrol officers. The patrol officer is responsible for adapting his or her behavior to both the policies of the agency, as communicated through the command staff, and the surrounding environment. The nature of this interaction affects whether victims are willing to report crimes to the police and thus how much crime is discovered. For example, the NYPD's decision to use a zero-tolerance policing strategy and how that policy is implemented by the patrol officers will result in a different reaction from the community than in other cities that apply community policing tactics. The relationship that police maintain with the communities they serve affects what crime is handled by police, how it is handled, and whether information about that crime is recorded (Aldrich & Herker, 1977; Manning, 2010).

The increasing specialization of jurisdiction in both crime enforcement and geographic area results in differing policing strategies and in different data recorded to meet the current challenges, even in areas where two different agencies may occupy the same space. Modern law enforcement is a complex process that often involves the careful coordination between agencies serving overlapping jurisdictions. In 1960, the proportion of law enforcement agencies that were identified as having concurrent jurisdiction with another agency in the UCR Program was less than 1 percent. However, the proportion of those same agencies had grown to about 20 percent in 2012 ("Monthly Report of Offenses Known by Law Enforcement (Return A) Master File," 1960; "Monthly Report of Offenses Known by Law Enforcement (Return A) Master File," 2012). Multiple agencies may serve the same geographic area,

such as areas in cities that contain a university or college. In areas where more than one law enforcement agency would have overlapping jurisdiction, there may be duplication in CFS, but the recording of crime incidents would be negotiated among agencies to determine which agency would take primary responsibility for handling the investigation of that incident. Specialized law enforcement agencies (such as campus police) might record data differently from traditional municipal law enforcement even though these agencies may be reporting in the same area.

As campus police agencies professionalized in the decades from the 1960s to the early 1980s, they tended to model themselves on municipal law enforcement agencies in terms of organization and operations. Campus police, however, serve a community that is quite different in its background and makeup when compared to most municipal agencies. In addition, the types of crimes reported to campus police tend to consist of more property crimes than violent crimes in comparison to municipal agencies (Bromley, 2007). The Clery Act mandates information sharing practices with other local jurisdictions (King et al., 2016; Storch, n.d.). Furthermore, the Clery Act requirements direct the dynamic of recording and reporting of crime data for campus police that is not experienced by municipal law enforcement agencies and may impact the ability for CFS to accurately model the information typically measured by incident reports (Gregory & Janosik, 2003; Janosik & Gregory, 2003).

### 3.4 METHODOLOGY AND EXPECTED RESULTS

The analysis focuses on a case study of the mid-southern city of Knoxville, Tennessee—two law enforcement agencies (LEA)—primarily by conducting a quantitative analysis of data collected by these agencies. However, this quantitative analysis is informed by a qualitative assessment of the agencies use of the same law enforcement data to provide context to the results. Knoxville Police Department (KPD) is the primary municipal police department serving a population of approximately 186,000 inhabitants. The University of Tennessee-Knoxville is the campus police department for a large land grant institution serving a student body of approximately 27,800 (28,000). The primary campus of the university is surrounded on three sides by commercial and residential areas of the city and a river on the fourth. Both agencies provided information on CFS and incident reports for 2014, as well as examples of their analysis of that information for one month of that year (April). Having both the quantitative information recorded by the agency as well as an agency's use of that data provides insight into the focus of an agency's police mission for that period (see Figure 3-1).

#### 3.4.1 Agency Priorities and Use of Data

To explore the possible impact of the use of data on the recording behavior of law enforcement agencies, samples of analyses produced for each agency were provided for the month of April 2014. It was thought important to choose a month that represented a period when both law enforcement agencies would be actively engaged with a full complement of the communities each serves. Because it is anticipated that the campus LEA would see a dramatic reduction in its activities in the summer months, April was used as a time when most enrolled students would still be on campus.

The published standard operating procedure for the KPD LEA Crime Analysis Unit identifies their goals to be focused on improved operational understanding and decision-making. The analyses are broadly defined in three areas of administrative analyses that support budget and resource allocation; strategic analyses that support proactive crime reduction strategies; and finally, tactical analyses that support investigations. Communicating crime data to non-LEA personnel is approved on a case-by-case basis,

and in general, the public is typically directed to crime data published in the agency’s annual report (Knoxville Police Department, 2005, 2014).

For the campus LEA, much of the crime analysis focus is directed by the Clery Act’s requirements associated with the release of key statistical information regarding crimes on and around campus. The Clery Act mandates LEAs maintain a daily crime log with basic information about the offense, date and time, general location, and disposition. The daily crime log must be accessible by the public. In addition, campus LEAs must also produce an annual security report providing information on offenses, broad geographic locations, and other characteristics of the crimes reported. These are minimum requirements enforced by the U.S. Department of Education (King et al., 2016; State University of New York, n.d.).

The types of products examined for the paper included crime logs, reports, presentations for COMPSTAT<sup>3</sup> meetings, maps, and other such documents used for internal communications with officers and external communications with their communities. The city (KPD) and campus (UTPD) LEAs provided 11 and 19 products to review, respectively. Overall, the types of analytical products obtained from the campus LEA were much more diverse than those from the city LEA. The city LEA was more inclined to produce reports for internal consumption, while the campus LEA had a mixture of both internal and external reports. The campus external reports most often were crime logs that are mandated by the Clery Act. Additionally, the campus LEA had not implemented any geocoding of their data at the time, so their analytical reports were more descriptive of general locations of concern for directing patrols and did not provide any maps associated with recorded data.

*Table 3-1. Number of offense "tags" by agency*

	City LEA (KPD)	Campus LEA (UTPD)
Person Crimes	9	61
Property Crimes	64	126
Society Crimes	0	0
Hate Crimes	0	8
Traffic	0	50
Non-criminal	0	12
Total Codes	73	257
Total Sources	11	19

The information provided in each product was tagged using qualitative software to reflect either a visualization of data (chart, map, or photo) or text by type of crime (see Appendix for description of coding scheme). Using this methodology, it is possible to have more than one tag or type of information within a product or a single page of a report. Based upon an enumeration of tags provided or displayed for consumption by an agency’s officers, there was a heightened concern directed at property crimes

<sup>3</sup> COMPSTAT (COMPare STATistics) is a strategic management program first introduced by Commissioner William Bratton in the New York Police Department in the 1990s to introduce accountability of mid-level police management through analysis of precinct-level statistical analysis (B. D. Weisburd, Mastrofski, Greenspan, & Willis, 2004)

(such as thefts, vandalism, burglary, arson, motor vehicle theft, or robbery) during the month of April (See Table 3-1).

Upon examination of data for the entire year of 2014, similar patterns showing the preponderance of crimes recorded as property offenses can be found in the incident reports for both agencies. A property offense is recorded on 47.8 percent of incident reports for the municipal agency and on 40.9 percent of incident reports for the campus agency (See Table 3-2). Upon analysis of the CFS, property crimes are more often recorded on a CFS in comparison to crimes against persons and crimes against society for both agencies as well (See Table 3-2).

Table 3-2. Counts and percentages of CFS and incident reports by agency and offense type, 2014

Offense Type	KPD				UTPD			
	Calls for Service		Incident Reports		Calls for Service		Incident Reports	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
Person	17,575	5.9%	3,450	13.3%	56	2.2%	69	6.4%
Property	28,553	9.6%	12,367	47.8%	445	17.1%	439	40.9%
Society	26,445	8.9%	3,353	13.0%	242	9.3%	474	44.1%
Traffic	37,402	12.6%	-	0.0%	334	12.8%	-	0.0%
Welfare	18,793	6.3%	38	0.1%	19	0.7%	109	10.1%
Other	167,665	56.6%	8,992	34.8%	1,506	57.8%	262	24.4%
Total <sup>a</sup>	296,433		25,874		2,604		1,074	

<sup>a</sup> CFS and Incident Reports can have more than one offense type recorded on an individual record. Counts and percentages by category will not equal the total.

However, each agency exhibits differing patterns of correlation when these same categories are analyzed by month to measure the level of association between the counts for CFS and counts of incident reports. The municipal agency had less measured association for the three primary crime types than the same categories for the campus agency. The CFS and incident reports had strong, statistically-significant correlations ranging from .820 to .904 for the campus agency, while the municipal agency had only one statistically-significant correlation which was for property crime (.667). (See Table 3-3.) Further exploration of CFS and incident report data from KPD revealed that the months of October, November, and December had patterns of higher incident reports than would be anticipated from the levels of CFS as compared to the other months. This nonlinear pattern is reflected in the lower correlations for KPD. Based upon these findings, the quantitative information for property crime recorded in April will be analyzed.



Table 3-3. Pearson correlations of monthly counts of CFS and incident reports, city and campus agency, 2014

	City LEA (KPD)	Campus LEA (UTPD)
Person	.393	.885*
Property	.667*	.904**
Society	.043	.820*

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

\* Correlation is significant at the 0.05 level (2-tailed)

### 3.4.2 Spatial Correlation and Co-location

Each agency provided a dataset from their computer-aided dispatch (CAD) system recording CFS for 2014 and incident reports recorded in a records management system (RMS) for the same year. Along the continuum of data inclusiveness with police data with arrest reports on one end being the most restrictive, while CFS have the fewest restrictions, incident reports fall in the middle of the spectrum where the decision to record the data is more reflective of the agency’s priorities rather than the demand for policing services as is the case for CFS.

The city data had been geocoded by the agency, however, the university data required geocoding before the analysis. Geocoding was accomplished by associating the listed (or closest) address to the name of the building, facility, or parking lot. After each record was geocoded, the reason for the call or recorded crime was recoded into broad categories to match Table 3-2. These included crimes against persons, crimes against property, crimes against society, traffic violations and stops, welfare checks, and a miscellaneous “other” category. Combining the categories also removes any potential issues where the individual calling for service does not accurately specify the type of property offense. For example, it is common for individuals to use burglary and robbery interchangeably even though they are distinct crimes. Table 3-5 in the appendix details the offense types that are included in crimes against property, which is the subject of this analysis.

After exploring the underlying categorical and spatial patterns of each data set, this research will discuss the application of the co-location quotient to assess the spatial association between CFS and incident reports. This will be accomplished in two stages. The first will be an assessment of the global co-location quotient (GCLQ) to determine if any spatial association exists between the two datasets. In addition, the local co-location quotient (LCLQ) will be calculated to assess if there are any smaller areas of association that can be measured on a point-by-point basis that are not detectable in a global measure of association. Analyses at micro-geographic scales will allow for these hotspots to emerge in ways that analyses at larger scales will not.

Despite the significant correlations between the counts of incident reports and CFS, it is anticipated that there will be less spatial association between CFS and incident data for the university data as measured by the GCLQ because the need to record certain crimes are driven by the needs of the Clery Act or other organizational pressures rather than simply recording the locations of crimes and other requests for

assistance. Conversely, there should be more spatial association between CFS and incident data for the city data as measured by the GCLQ. Finally, the LCLQ could reveal differences in spatial patterns between the two agencies—especially in the boundary area where there is concurrent jurisdiction around campus.

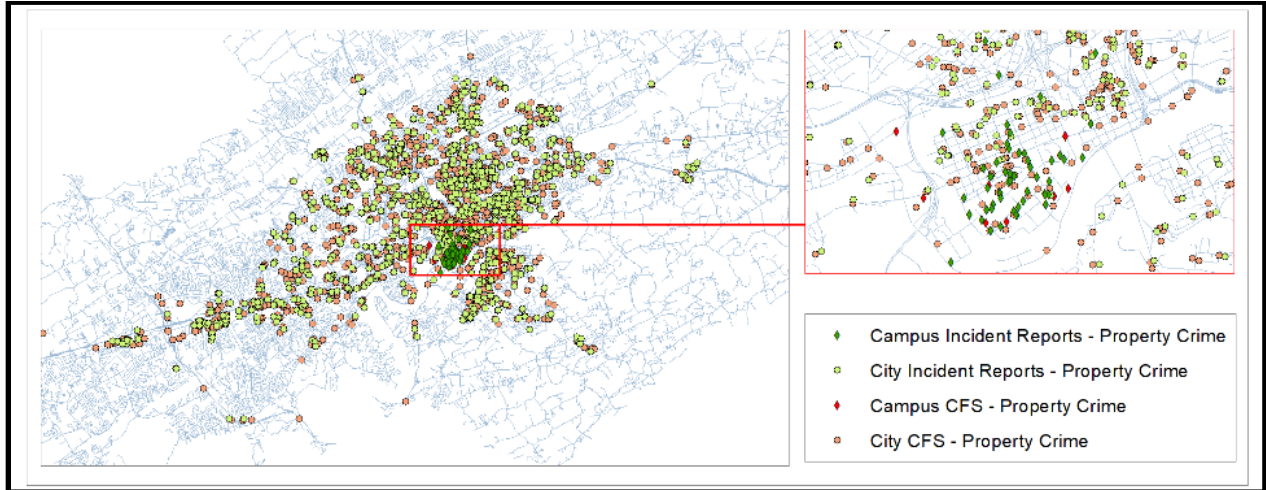


Figure 3-1. CFS and incident reports for property crime in study area, April 2014

### 3.5 ANALYSIS

#### 3.5.1 Cross-K Analysis

While spatial autocorrelation measures the spatial correlation of an observation with other objects like itself, spatial correlation (or co-location) measures the spatial relationship with differing objects. The Cross-K or bivariate K function—an extension of the Ripley’s K function testing—is a commonly used measure of clustering patterns of a point dataset around another represented as type  $i$  and  $j$ . Cross-K analysis measures the overall density ( $\lambda$ ) of one set of points (for example type  $j$ ) within a specified distance of another set of points (type  $i$ ) compared to that would be observed by chance. It is an asymmetric measure based upon distance of points  $j$  from points  $i$ , as such, it considers all possible pairs of points, not just the nearest ones. The distance between the  $k$ th location of type  $i$  and the  $l$ th location of type  $j$  is represented as  $d_{i_k, j_l}$  for a given area ( $A$ ) (Dixon, 2002).

When edge corrections are not applied, the Cross-K function is formulated similar to the univariate Ripley’s K as:

$$\hat{K}_{ij}(t) = (\hat{\lambda}_i \hat{\lambda}_j A)^{-1} \sum_k \sum_l I(d_{i_k, j_l} < t)$$

When edge corrections are applied, the Cross-K function is independent of the shape of the area. The results were calculated using the *Kcross* function in Spatstat package in R (Baddeley, Rubak, & Turner, 2015). The null hypothesis assuming complete spatial randomness (CSR) is provided by the  $K^{pois}$  in the results in Figure 3-2 and Figure 3-3. The other three results indicate the Cross-K results based upon three different edge corrections:  $K^{iso}$ : isotropic correction (or Ripley’s);  $K^{trans}$ : translation correction (or

toroidal);  $K^{bord}$ : border correction (or guard area). The estimates indicate clustering between CFS and incident reports for both the city and the campus law enforcement. In addition, the three differing edge correction techniques show very similar profiles, which indicate that there were no difficulties estimating the Cross-K function. While the results from the Cross-K analysis indicate significant clustering between the CFS and incident reports, these estimates could be biased if the distribution of the underlying population also exhibit a significant spatial clustering pattern or potentially overwhelm any detectable pattern between the two datasets solely represent the clustering of the underlying population with crime data (Leslie & Kronenfeld, 2011; Wang, Hu, Wang, & Li, 2017).

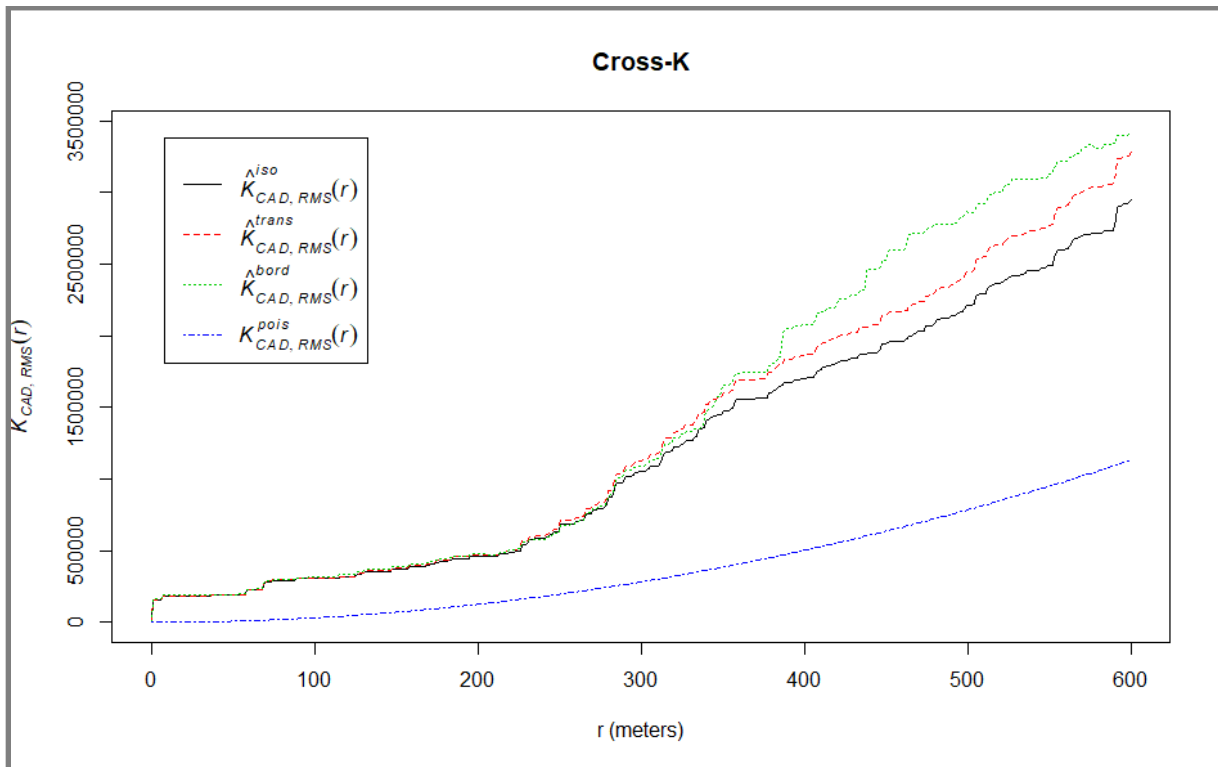


Figure 3-2. Cross-K analysis from CFS to incident reports, UTPD

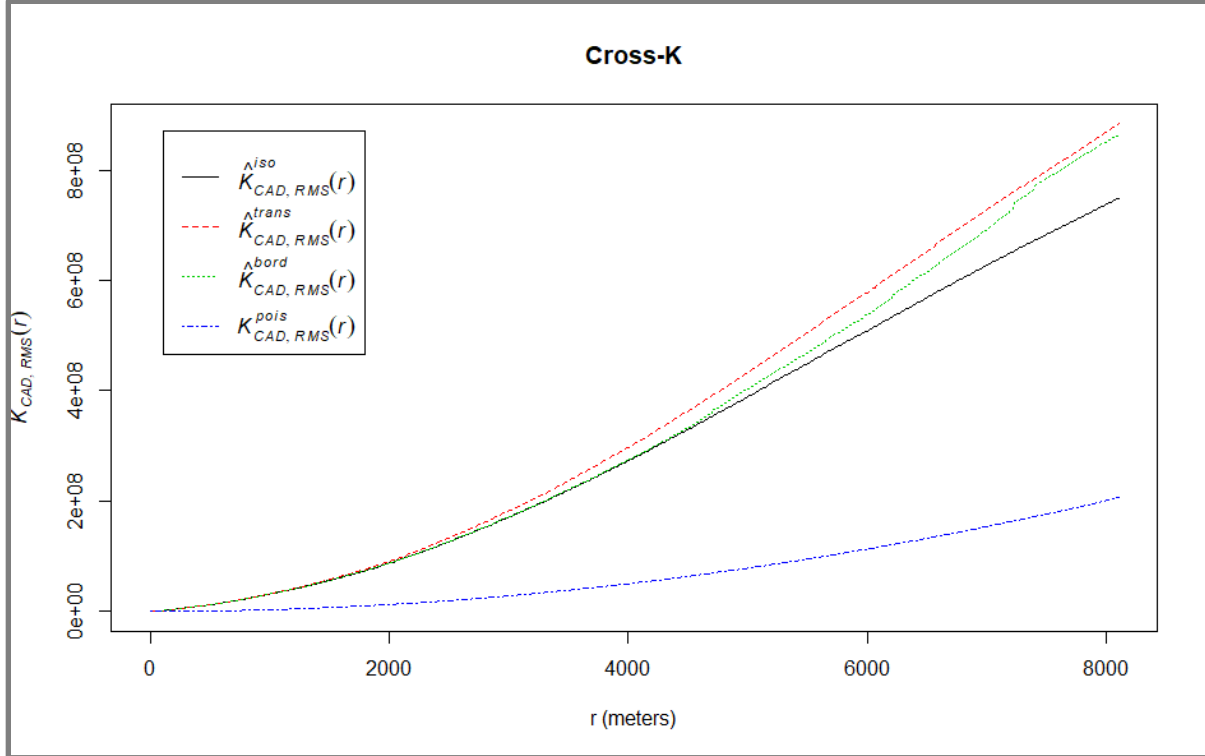


Figure 3-3. Cross-K analysis from CFS to incident reports, KPD

### 3.5.2 Global Colocation Quotient

Given the potential for bias in the Cross-K analysis, Wang, et al. (2017) suggest that the GCLQ provides a better measure for the degree of spatial dependence between two point data sets. Originally developed by Leslie and Kronenfeld (2011), the GCLQ is seen as advantageous over the join count statistic and the Cross-K function because it is less sensitive to the bias of clustering of the underlying population as is common with procedures based on distance. The GCLQ is an asymmetric measure formulated as

$$CLQ_{A \rightarrow B} = \frac{N_{A \rightarrow B} / N_A}{N_B / (N - 1)}$$

where  $N_A$  and  $N_B$  represent the counts for each of two groups, respectively.  $N_{A \rightarrow B}$  is the number of type A points that have a Type B point as its nearest neighbor. The function calculates the ratio between the observed proportion of points and the expected proportion by chance. To address the presence of multiple nearest neighbors within the search radius, each point is treated equally within the calculation of the GCLQ in the following manner. The number of points of type A are noted as  $i$  and  $nn_i$  is the number of nearest neighbors to point  $i$ ;  $j$  are point  $i$ 's nearest neighbors  $nn_i$ . A binary variable that indicates whether a point is of type B, with yes as 1, is represented as  $f_{ij}$  (Wang et al., 2017).

$$N_{A \rightarrow B} = \sum_{i=1}^{N_A} \sum_{j=1}^{nn_i} \frac{f_{ij}}{nn_i}$$

The null hypothesis for the GCLQ is stated as, “given the clustering of the joint population, there is no spatial association between pairs of categorical subsets” that would result in a value of 1.000 (Leslie & Kronenfeld, 2011, p. 312). Any value over 1.000 would indicate points of type A were spatially collocated with points of type B, while values less than 1.000 would show a level of spatial dispersion.

Wang et al. extend the research of Leslie and Kronenfeld by developing a freely-available tool to calculate the statistical significance of the GCLQ using simulation techniques. The analysis of the crime datasets in this paper again focus on property crime for April 2014 assessing the spatial dependence of incident reports on the CFS for each of the agencies. An adaptive bandwidth was used with a maximum search radius of 300m using a Gaussian distribution for both the 5 nearest neighbors and the single nearest neighbor. The adaptive bandwidth uses a rank rather than a fixed measured distance to ensure the same number of points are used in the calculation. The maximum search radius was chosen to capture an approximation of city block size. Recently published work by Weisburd and his colleagues (Groff, Weisburd, & Yang, 2010; Steenbeek & Weisburd, 2016; Telep, Weisburd, Gill, Vitter, & Teichman, 2014; D. Weisburd & Amram, 2014; D. Weisburd, Groff, & Yang, 2013) suggest that micro-geographies of the city block reflect the most salient geographic scale to explore crime and place. The p-values were calculated using a replication of 100 times.

Table 3-4. Global colocation quotients, CFS to incident reports with property crime, April 2014

	Global Colocation Quotient		
	5NN	1NN	N
KPD	1.09756	1.18539	
(p-value)	(0.0)	(0.0)	3,469
UTPD	1.01295	1.00654	
(p-value)	(0.74)	(0.96)	88

The results indicate that there is a statistically-significant clustering of incident reports on property crime around areas that experience CFS related to property crime for the city agency. While the increase in likelihood is somewhat modest, the strongest effect appears to occur with the single nearest neighbor. This result is not supported for campus agency. For the campus law enforcement agency, the results do not deviate substantively from 1.000 showing no increased or decreased likelihood for an incident report for property crime to be recorded near a CFS for a property crime. In addition, the results are not statistically-significantly higher than 1.000, thus failing to reject the null hypothesis (See Table 3-4).

### 3.5.3 Local Colocation Quotient

Global measures, such as the Cross-K and the GCLQ, assume that clustering would be stable across space. Local measures allow for a more detailed analysis to understand either the process behind clustering or dispersion. In addition, local measures can identify occurrences of clustering or dispersion that might be overwhelmed by a majority that would not be identified as either clustered or dispersed. The LCLQ was calculated to assess the spatial influence of CFS on the RMS data for property crime in

April 2014 for each law enforcement agency in this study. The LCLQ is similar to the GCLQ as it is an asymmetric of spatial dependence between point data sets. The calculation of the LCLQ is similar to that of the GCLQ. However, the values of  $N_{A_i \rightarrow B}$  is the weighted average number of type B points that are the nearest neighbors to point  $A_i$ . The weights ( $w_{ij}$ ) reflect a measure of importance as a function of distance, and  $f_{ij}$  is a binary variable used to indicate whether the point is a marked type B point (1 = yes and 0 = no). In essence, the further away a neighbor is from a point, the less important it is. The third formula below represents a Gaussian kernel density function used to assign a greater weight to points closer to the  $i$ th A point with  $d_{ij}$  being the distance between the  $i$ th A point and point  $j$  and  $d_{ib}$  being the bandwidth distance around the  $i$ th A point. Similarly to the GCLQ, the LCLQ also uses an adaptive bandwidth to ensure the same number of points are used in the calculation. (Wang et al., 2017).

$$LCLQ_{A_i \rightarrow B} = \frac{N_{A_i \rightarrow B}}{N_B / (N - 1)}$$

$$N_{A_i \rightarrow B} = \sum_{j=1(j \neq i)}^N \left( \frac{w_{ij} f_{ij}}{\sum_{j=1(j \neq i)}^N w_{ij}} \right)$$

$$w_{ij} = \exp \left( -0.5 * \frac{d_{ij}^2}{d_{ib}^2} \right)$$

For each point, the LCLQ will indicate spatial colocation for individual points A around a B type point when the coefficient is over 1.000, and spatial dispersion with the LCLQ is less than 1.000.

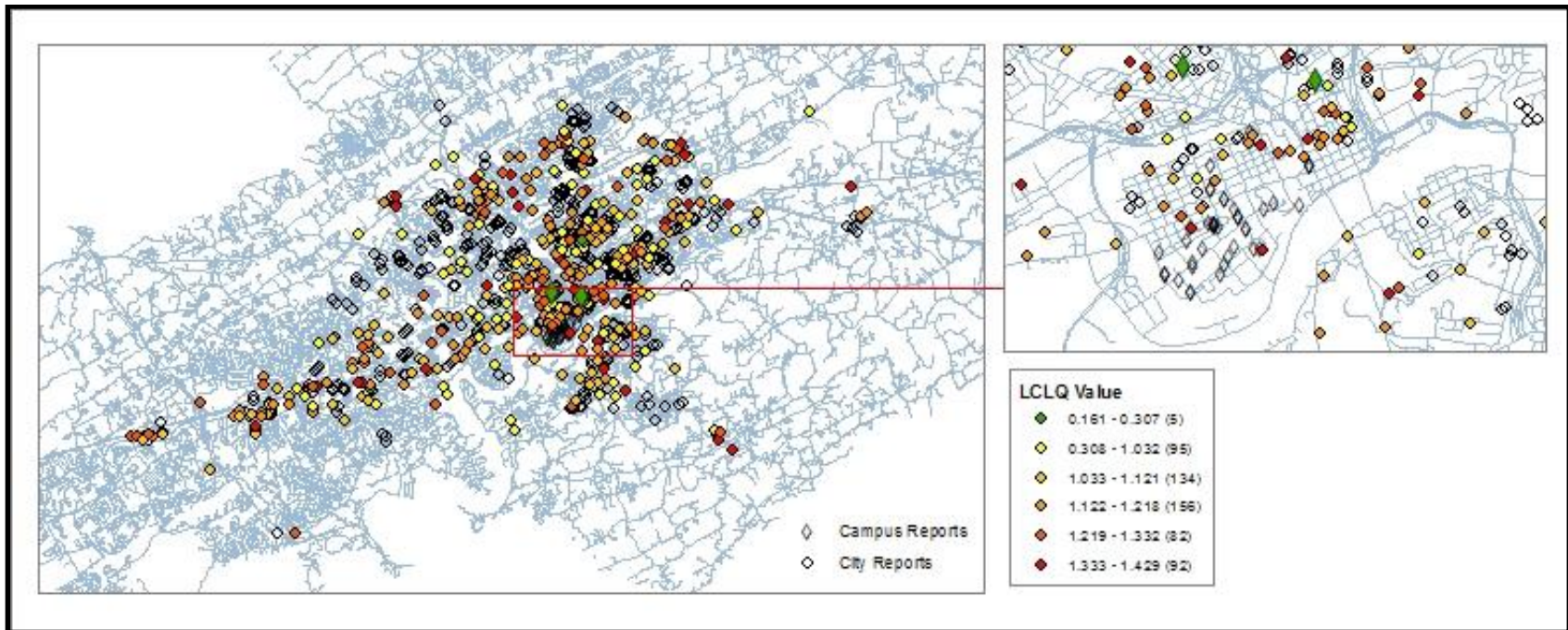


Figure 3-4. Local colocation quotients, CFS to incident reports with property crime, April 2014



Over half of the incident report locations were spatially colocated with its nearest CFS neighbor (54.1 percent). Upon examination of the 564 statistically-significant clusters, the minimum value of the calculated LCLQ was 0.161 and the maximum value was 1.429. Only approximately 10 percent of the locations indicated spatial dispersion with an LCLQ value less than 1.000. Local results show that statistically-significant clusters do not appear to be associated with one area over another for the city agency. Local results of the campus show almost no property crime incident reports that were spatially associated with a similar property crime related CFS. These campus incident reports lie north of the main campus and represent noncampus buildings under the jurisdiction of the campus law enforcement agency. Results, again, support more spatial colocation between CFS and incident reports for the municipal law enforcement agency than the campus law enforcement agency. While the CFS for the municipal law enforcement agency has a modest amount of success in predicting where incident reports are made, the same is not true for the more specialized type of agency represented by this campus law enforcement agency (see Figure 3-4).

### 3.6 DISCUSSION AND CONCLUSION

The “digital turn” in data and data collection challenges the research community to consider the assumptions made with the use of “open data” (Ash, Kitchin, & Leszczynski, 2018). An increasing amount of data available to the research community as open data has become the norm across all sectors of analysis and research. This is especially true in law enforcement given the number of operational databases available. It is tempting to use the administrative data made available by law enforcement to further important research goals without regard to the source and manner in which the data was originally collected. However, data are not neutral. The impact of data collection choices is greater for operational or administrative data such as criminal justice data. The collection of any operational data is mediated by politics and the choices made by individuals serving in a boundary role in the collection of information. In order to know limitations associated with administrative data it is important to take into account whose needs are being met with the data and how that impacts who has access to data (Boyd & Crawford, 2012; Elwood & Leszczynski, 2013; Kitchin, 2014; Kwan, 2016; Mayernik, 2017; Miller & Goodchild, 2014).

Given the lack of measurable spatial association between CFS and incident reports, how does this challenge our assumptions of law enforcement data and the ability to use various sets of data interchangeably? The primary focus for CFS data is the management and dispatching of law enforcement and emergency resources. The CFS data tends to be location-based based on this goal. On the other hand, incident data tends to be person- or object-based to manage the investigative process of law enforcement. Each have differing measurement issues or errors centered on the primary needs met by each. CFS data often has error in the description of the event given it’s often based on the public’s account but can be useful as a measure of demand or confidence in law enforcement. Incident data have better descriptions of the event reflecting an initial assessment of the even by law enforcement. However, law enforcement discretion in data recording may affect data completeness and this effect may vary by location as well as type of law enforcement agency. Findings from this study show that that CFS are not strong spatial proxies for incident data. Traditional city LEAs may have a stronger relationship than campus LEAs. However, the relationship is weak and still show evidence of discretionary data recording.



Each LEA operates from a different set of assumptions about what generates crime in places that it operates and differing needs for the use of data generated by these assumptions and operations. Campus law enforcement has a mandate by the Clery Act to produce information to be consumed by their student and parent communities. Data collection supports the process of both compliance with the law and providing overt information on the campus agency's ability to address both criminal and safety issues impacting the student population. However, municipal law enforcement has greater latitude to determine what information is provided directly back to the communities they serve and what is produced for internal consumption. The simple difference could lead to vast differences in the amount of information available for analysis. While both agencies showed concern for property crimes, campus agencies recorded a wider array of crimes and recorded more crimes against society (such as drug and alcohol-related offenses) and welfare-related incidents such as mandated by the Clery Act in comparison with the municipal agency.

Using this one example, data collected on CFS do not always manifest itself in the same way as recorded incident data depending upon the type of agency or perhaps even differing priorities. The findings show that CFS has a modest ability to measure reports for property crime for a traditional municipal law enforcement agency. However, this clustering of incident reports around CFS all but disappears for the campus agency in the same city. In addition, the overall lack of strength of spatial association of incident reports to CFS could call into question whether CFS would be a reasonable substitute for incident reports when location is considered. Given earlier significant results of the Cross-K function and the more modest results of the co-location quotients, there is evidence that clustering was more indicative of a spurious relationship, rather than a direct influence of CFS on incident reports.

Past research studies that were more concerned with levels of crime for aggregate areas may not have been impacted by this lack of association as illustrated by the initial correlations between CFS and incident reports. However, current research is highly focused on micro-geographies. The confluence of the desire to use "big" or "open" data with added attention paid to more exact location information should call researchers to more closely assess the assumptions being made about their data. Whether or not data generated by forces outside of the LEA, such as a CFS initiated by a member of the public would closely model the data generated by actions of the LEA would be highly dependent on whether the priorities of the LEA are focused on the response to public requests or not. Given that CFS often have more easily attained location information compared to incident reports, it is critical for researchers to examine this relationship and assess whether CFS would be the appropriate measure.

As more research begins to mine the possibilities of "open" and "big" data to explore and test ideas about crime and criminality, data should be used with a critical eye to understand the possibility of generalization. While this research cannot be necessarily generalized or transferred to other areas, future work should seek to test these same assumptions with other agencies or locations to see if they hold. This research demonstrates the need to focus on smaller areas and varying levels of scale for potential research on the interaction between law enforcement activity and the data they record. In addition, research should expand beyond cities. How do the diverse types of place differ in, as Weisburd and Amram called it, the "regular recurring rhythms of activity" (2014, p.105)? As law enforcement becomes more specialized and overlap in geographic jurisdiction, it will become increasingly important to account for the impact of agency priorities on the data it collects.

APPENDIX

Table 3-5. *Offense categories<sup>a</sup> from the UCR Program*

Crimes against Persons	Crimes against Property	Crimes against Society
<ul style="list-style-type: none"> <li>• <b>Assault offenses</b></li> <li>• <b>Homicide Offenses</b></li> <li>• <b>Human Trafficking</b></li> <li>• <b>Kidnapping &amp; Abduction</b></li> <li>• <b>Sex Offenses</b></li> <li>• <b>Sex Offenses, Nonforcible</b></li> </ul>	<ul style="list-style-type: none"> <li>• Arson</li> <li>• Bad Checks</li> <li>• Bribery</li> <li>• Burglary/Breaking &amp; Entering</li> <li>• Counterfeiting &amp; Forgery</li> <li>• Destruction/Damage/Vandalism of Property</li> <li>• Embezzlement</li> <li>• Extortion/Blackmail</li> <li>• Fraud Offenses</li> <li>• Larceny-theft Offenses</li> <li>• Motor Vehicle Theft</li> <li>• Robbery</li> <li>• Stolen Property Offenses</li> </ul>	<ul style="list-style-type: none"> <li>• Curfew/Loitering/Vagrancy Violations</li> <li>• Disorderly Conduct</li> <li>• Driving under the Influence</li> <li>• Drug and Narcotic Offenses</li> <li>• Drunkenness</li> <li>• Family Offenses, Nonviolent</li> <li>• Gambling Offenses</li> <li>• Liquor Law Violations</li> <li>• Peeping Tom</li> <li>• Pornography &amp; Obscene Material</li> <li>• Prostitution</li> <li>• Trespass of Real Property</li> <li>• Weapon Law Violations</li> </ul>

<sup>a</sup>Traffic and welfare reports were identified as such within each dataset. Those reports not falling into any of the above five categories were coded as “other.”

# 4 THE RECORDING OF CRIME DATA AND SPATIAL SCALE: THE INFLUENCE OF LAW ENFORCEMENT RECORDING BEHAVIOR AND THE UNCERTAIN GEOGRAPHIC CONTEXT PROBLEM

---

## 4.1 SPATIAL UNCERTAINTY ASSOCIATED WITH LAW ENFORCEMENT DATA

While law enforcement agencies are evaluated on the occurrence and their response to local crime problems based on areal jurisdictions, the reality is that the occurrence of crime does not “respect” the boundaries of law enforcement jurisdictions and their internal divisions. The phenomenon of crime may be viewed as a continuous surface or “blanket” over the underlying populated landscape while the law enforcement response is stitched together from a patchwork of varying organizational jurisdictions, priorities and styles to address the problem of crime. The same problems that exist on one side of a jurisdictional boundary can also prevail on the other irrespective of the intervening jurisdictional boundary. Under certain circumstances or with a prior formalized understanding among their respective agencies, law enforcement agents may act outside their official territory or have some mutual responsibility for the same area, such as occasions when campus law enforcement patrol in areas adjacent to campus. Agencies on either side of a border may have different enforcement policies and goals. Current research has only touched upon the influence that law enforcement response might have on the recording of crime data—the same data that is used to measure underlying crime problems in an area and evaluate law enforcement’s response to these problems. Measurement and reporting are intertwined and the following paper attempts to untangle them and cast some light on the effects on spatial data analysis of law enforcement practice.

Most recorded crime data is built upon the structure of the Uniform Crime Reporting (UCR) Program. The Uniform Crime Reporting (UCR) Program came into existence in 1929 through the work of a group of police chiefs struggling to answer even basic questions about crime and crime trends. The crucial contribution of the UCR Program is an overall set of standard definitions and methods used to record comparable crime data even though local statutes and definitions may vary. However, this set of standards provides at times a false impression that law enforcement itself is uniform in its approach and response to crime. Data based upon law enforcement activity—such as UCR data—are subject to issues related to scale such as the Modifiable Areal Unit Problem (MAUP), the Uncertain Geographic Context Problem (UGCoP) and boundary edge effects.

The MAUP is “a problem arising from the imposition of artificial units of spatial reporting on continuous geographical phenomena resulting in the generation of artificial spatial patterns” (Heywood, Cornelius, & Carver, 2002, p. 8). Because of MAUP, artifacts or errors are created when data are grouped into spatial units for analysis. Take for example the crime rate of a population. A crime incident is perpetrated by a unique individual, in a unique location. Yet managers, researchers and others often

want to understand crime in a broader context. Because handling data at a 1:1 level is incredibly difficult, to gain greater insight and identify underlying patterns, data points are grouped together as a necessary condition for analysis and understanding. A dot map indicating crimes of different types is helpful for some types of analysis, such as spatial-temporal clustering, or methods of operation if available. However, a map showing different crime rates normalized by population for census tracts, municipalities or state may provide a needed perspective depending on the question being asked of the data. While aggregation is a necessary step, the analyst must carefully consider how the aggregation is introduced into the analysis. The creation of this grouping may distort or exaggerate the actual data pattern as much as it can assist in identifying patterns.

There are two distinct types of MAUP: scale (or aggregation) and zone (or grouping) (Openshaw, 1984; Openshaw & Taylor, 1979; Wong, 2009). The scale of analysis would likely produce different results for different areal units, e.g. the United States, by state, by county, or even block-by-block. The crime rate for the United States is different from that for the state of Tennessee, which is again different from that of Knox County, TN, which is again different from a neighborhood in Knoxville. As with all analyses, the ideal situation is to choose an optimal scale that matches the research question and process under investigation. However, research data often comes in specific levels of aggregation, and it is not always possible to find data at the optimal scale or even know what scale is optimal. While finer-scale data can always be aggregated to larger units if available, coarser scale data cannot easily be subdivided to finer scale to arrive at the appropriate scale for research thus possibly introducing bias to the results if the scale does not take into account context or place (Mennis & Hultgren, 2006; Wong, 2009).

As an extension, the zoning problem is primarily focused on the effects of data aggregation based upon somewhat arbitrary spatial units or zones. The zoning or grouping effect on the data has the effect of smoothing effect on existing variation, and thus the choice of a spatial unit either enhances or diminishes the spatial pattern that is in the underlying data. While this data smoothing is sometimes desirable to allow for a pattern of association to emerge more prominently, it can also lead to spurious effects or a suppression of the pattern (Wong, 2009). Zones may be based on some simple logical choice, or along preexisting physical or social divisions. However, any set of zoning rules creates potential bias. Thus, in this paper the results of two different areal schemes, census tracts and police beats, are compared. Different results may be simply artifacts of different geographical divisions and a single best solution may be illusory (Matthews S. A., Goodchild M. F. and D. G. Janelle Advanced Spatial Analysis program 2008 <http://gispopsci.org/maup/> accessed 11/30/18).

Edge effects in the MAUP also demonstrate the effects of boundaries on data. When an artificial boundary is imposed on an area, such as that of a law enforcement jurisdiction, it can lead to unintended analytical conclusions because measurements of adjacent jurisdictions are usually unaccounted for in the model. Best practice would encourage analyzing the raw point data on crime to determine appropriate break points rather than rely on the data presented within predetermined jurisdictional boundaries. However, at the local level, crime data is usually displayed and analyzed within a particular jurisdiction without regard to cases in neighboring jurisdictions, resulting in the introduction of uncertainty (O'Sullivan & Unwin, 2010). Consequently, geographic boundaries associated with the management of law enforcement can introduce uncertainty into the characteristics of the data and its location, which increases the difficulty associated with any analysis of crime data. To manage uncertainty, more information about agency responsible for collecting and reporting crime data may need to be included into statistical models using that data.

While the effects of aggregation and boundaries have been documented in the literature for decades, recently there has been greater focus on how aggregation and the use of areal units and boundaries affect the use of correlates in the modelling of geographic data. The UGCoP is a newer concept that, in some ways, more directly addresses the ecological fallacy and the MAUP. The UGCoP describes the uncertainty that can exist between the scale of covariates and the geographic scale at which the covariation exists (Kwan 2012a; 2012b). Rather than focus on the effect of aggregation or zoning on individual variables, UGCoP explores how scale comes into play when researchers are exploring the relationship among variables. As Kwan states, "...humans are active agents who construct their own geographic contexts and tie together different spatial scales through their daily activities, movements and social interactions" (Kwan 2012, 966). The challenge to researchers is to think through the mechanism being measured to arrive at the most appropriate unit of aggregation for measuring the phenomenon of interest.

While studies of crime have become increasingly focused at the scale of microplaces, the question arises how best to examine the geographic relationships of crime and collective social processes when our measures and correlates may not be available in the same units and scale (Goodchild, 2018; Groff, 2018). Is it justifiable to model crime data based solely upon the assumption that socioeconomic factors affect the levels of crime, or is there evidence that internal factors related to law enforcement practices affect the recording behavior of law enforcement? Environmental criminology has focused on the collective social processes of residents at differing scales. These processes have been referred to with different terms, such as informal social control or collective efficacy; however, these themes have been persistent in the literature. Law enforcement has its own collective functioning both within the organization and in terms of their interactions with residents (Groff, 2018). Given the known effects of aggregation on spatial data, it is also worth exploring the effects of law enforcement activity in the selection of appropriate units of spatial aggregation. While the availability of additional data at the Census tract-level makes Census tracts an attractive choice, there may be more appropriate aggregate geographies for the purpose of modelling crime data, such as law enforcement internal administrative boundaries, for example police beats or precincts, which reflect the scale of policing practice and hence crime data collection given the close connection between the dimensions of beats and precincts and the workload of policing.

The objective of this paper is to evaluate the effect of introducing law enforcement-related information into interpolation/allocation models of crime data, in the context of scale and the use of correlates to boost their performance. An example using information from a medium-sized Southern police department will apply dasymetric mapping techniques to allocate jurisdiction-wide recorded crime counts to smaller-area (or large-scale) geographies. A binary dasymetric model using street density alone will be compared with an "intelligent" dasymetric model using street density weighted by calls for service (CFS) as a measure of demand of police services. In addition, the example will evaluate the appropriate small area geographies for aggregation of crime data by comparing Census tracts to the law enforcement agency's own sub-jurisdictional divisions or beats.

## 4.2 BACKGROUND

### 4.2.1 Crime Data

As is the case with many statistical collections—especially those whose goal is a complete enumeration—the UCR Program has steadily grown in participation levels from relatively sparse beginnings. By the late 1950's, the UCR Program had participation levels that allowed for more representative statements about crime at national, state, and local levels. However, the voluntary nature of the UCR Program entails that for any given year, there are agencies that choose to either not participate at all or to provide incomplete information. This irregular compliance creates the necessity for imputing crime data for agencies that either provide incomplete data or no data at all, in order to arrive at complete national and state estimates of crime levels and trends (C. Barnett-Ryan, 2006; Cynthia Barnett-Ryan, 2019b).

UCR data is closely associated with summary-based statistics that have been collected since the early part of the 20<sup>th</sup> century. The UCR Program introduced a new incident-based collection, the National Incident-Based Reporting System (NIBRS) in the early 1990s to address the need for more detailed information about crime known to law enforcement, but its acceptance has been sporadic (Federal Bureau of Investigation, n.d.). In January 2016, the FBI announced the impending retirement of the summary-based UCR statistics by January 2021 to facilitate the complete transition to incident-level data collection. After this date, NIBRS will be the only mechanism to report data to the FBI UCR Program. (Federal Bureau of Investigation, 2016). However, the shift of national crime data to more detailed incident-based statistics will not lift the voluntary nature of the UCR Program. Modelling crime data will become more complex when these models begin to incorporate not only counts but also the characteristics of crimes.

While the current reporting levels of UCR data are well above 85 percent for most years, it is uncertain what level of reporting can be anticipated the first year that summary statistics become unavailable (Cynthia Barnett-Ryan, 2019b). At present, 6,998 agencies of the approximate 18,000 law enforcement agencies submit crime data in the incident-based format of NIBRS (Federal Bureau of Investigation, 2018). It is reasonable to assume that there will be a significant decrease in crime data reporting with the retirement of summary-based UCR statistics at least in the near-term. Missing data can be addressed through the interpolation of values between known data points though interpolation would inevitably introduce uncertainty. Managing the uncertainty of interpolation increases the necessity of gaining a more complete understanding of the qualities which factor into the identification of peer agencies, and the scale at which correlates (or potential auxiliary variables) of crime operate.

Simultaneously, interpolation can be a method for treating the uncertainty of values at non-sampled points including missing measurements. There are numerous methods available to researchers to deal with the estimation of values at non-sampled points. Some are based on global stochastic methods such as ANOVA or trend surface analysis. Global methods borrow the strength of the average effects of correlated data over the whole region or area. Local methods rely on methods that weight the effects of nearby sampled units more heavily than those further away (Burrough & McDonnell, 1998). Goodchild and Haining (2004) noted, however, that decisions regarding measurement and analysis were often a product of convenience rather than an optimal choice. These choices should be scrutinized for their impact on the final interpretation of analytical results and the certainty of those results. This scrutiny should directly address the concepts and concerns of UGCoP and MAUP.

#### 4.2.2 Uncertainty, Crime Data, and Criminality of Place

Recent criminological literature has pivoted towards a renewed focus on the role of place and criminality. Weisburd and his associates have published consistent findings that certain places have persistent problems with crime regardless of law enforcement intervention (D. Weisburd, 2015a; D. Weisburd, Groff, & Yang, 2014). This finding adds to a long lineage of research looking at crime and disorder through an environmental lens, dating to the 1920's in the US and earlier in Britain (Andresen, Linning, & Malleson, 2017; Andresen & Malleson, 2011; Braga, Papachristos, & Hureau, 2014; Friendly, 2007; Groff et al., 2010; Park, Burgess, McKenzie, & Wirth, 1925; R J Sampson & Wilson, 2012; Shaw & McKay, 1942, 1969; D. Weisburd et al., 2014). Earlier work in the 20th century concentrated on the relationship between neighborhood phenomena and potential criminogenic effects while more recently, research associated with criminality of place has focused heavily on street segments, a much higher resolution.

Weisburd (2015b) identifies the microgeography of street segments as the key geographic unit that serves as a measure of social interactions for these processes. Residents are likely to have detailed and visual familiarity with street segments, reflecting an initial layer to the activity space for the resident that is much smaller in scale compared with the typical definitions of a neighborhood. Anecdotally, the situation can be found where one block has persistent problems with crime, while the next block may not. The street segment is also a scale at which police operate daily and for which crime data are recorded. The focus on microgeographies has opened discussions about the underlying criminogenic processes and at which scale these relationships are observed.

##### 4.2.2.1 Scale and Geographic Context

Uncertainty can be introduced in a variety of ways to the study of crime data originating from many sources. For geographic questions, as described in section 4.1 above, there is the question of scale and aggregation and how these affect the interpretation of key relationships in the data. While a few researchers have investigated the potential impact of scale and aggregation of crime incidents, little has been done to measure how law enforcement practices may affect the collection of data (Boivin & Cordeau, 2011; Klinger, 1997; Phillips & Sobol, 2012; Varano, Schafer, Cancino, & Swatt, 2009). By understanding how law enforcement activity affects data collection, this study can shed light on the main avenues for improving crime data modeling and therefore help develop more robust interpolation methodologies used for estimating missing data in the UCR Program.

Individual and community associated correlates are often studied in relation to UCR data (for example, income, parents' education, in-migration). However, there are two main factors that need to occur for data to be collected that are not necessarily related to those individual or community-level processes—crime discovery and crime reporting. Crime discovery is the awareness by law enforcement of an event, either through victim reports or through law enforcement observation. Crime reporting reflects the formal recording of a crime once it has come to the attention of law enforcement. Because it is necessary for both these two factors to occur for an incident to be included in crime statistics, the possibility exists that differences in the philosophy of policing among law enforcement agencies might translate to differences in crime discovery and crime reporting. These differences may ultimately affect the UCR data. The decision over which correlates to use is subject to the UGCoP (Kwan, 2012c, 2018; Kwan & Schwanen, 2018).

UGCoP is related to challenges in detecting patterns at various scales and determining which scale captures the actor's geographic context for an issue such as crime. Kwan (2012) mentions "neighborhood effects" as related to health, but law enforcement-recorded crime data may be more closely associated with the activities of street-level policing rather than the characteristics of the community itself. It would be naïve to assume that residential characteristics of the neighborhood, including possible indicators of disadvantage, do not have a strong relationship to the criminality of place. However, there is little insight on the effect of law enforcement activities on the recorded data used as the basis for analysis. There is an implicit assumption that recorded law enforcement data only reflect the activity space of residents, neglecting the activity space of the police or criminal. It is as if the activity of the measured is accounted for but not the measurer. Research has yet to explore fully whether the opportunity structures related to crime and place are more closely aligned with the activities of the residents, the police, or both entities.

#### *4.2.2.2 Bounding Geography of Law Enforcement and Crime Data*

Boundaries take on both physical and nonphysical forms and can directly affect the activities of these entities. Physical boundaries such as streets, waterbodies, and terrain are evident, whereas nonphysical boundaries may be less so. These physical changes in the environment produced by encountering streets, land use or terrain could result in differential guardianship of an area; increase chance of victimization; or change in perception that could alter behavior. Most authors have viewed this phenomenon as either a reflection of individual's interaction with differing environments in adjacent zones or law enforcement's response to the changes from one neighborhood to the next (Y. A. Kim & Hipp, 2018; Song, Andresen, Brantingham, & Spicer, 2017).

However, it is also possible that nonphysical boundaries can affect law enforcement response and as a result introduce an artifact to the process of data collection. Examples of nonphysical boundaries include some jurisdictional boundaries, internal organizational boundaries of jurisdictions, and externally-imposed representations of the transition from one neighborhood- or community-type to another. Kim and Hipp (2018), as one example, observed a distance decay pattern for some crime types moving from centrally-located police stations to edges of jurisdictional boundaries. This pattern could be result of less patrolling of the edges of jurisdictions due to the geographic distance itself or due to ambiguity over law enforcement responsibility with an adjacent law enforcement agency's jurisdiction.

Two common representations of internal boundaries of a city's jurisdiction are Census tracts and agency precincts or beats. Census tracts are often used as a unit of aggregation and analysis, not only because of the richness of data available from the Census Bureau, but also because of the underlying methodology in the construction of those aggregate units. Census tracts are designed to be approximately homogeneous in socioeconomic characteristics at the time of creation. The tracts range in total population from 1,200 and 8,000 and may vary in area depending upon the population density. The intention of the Census Bureau was for these units to be stable over time to facilitate longitudinal comparisons; however, they may be split or merged over time when population growth or decline requires it. Census tracts are not an operational scale for police activity.

Just as with any human activity, police administration operates at multiple scales. At its most rudimentary level, the activity space of law enforcement is defined by its jurisdictional boundaries. These boundaries constrain most daily movements of law enforcement officers. To move across jurisdictional boundaries in their official capacity, officers must either rely upon existing agreements



between neighboring jurisdictions or request ad hoc permission. In addition to these “harder” boundaries between agencies, there are additional internal boundaries of precincts or beats that govern the organization of officers’ movements and activities. Their goal is optimizing the response of foot and car patrols and defining the area within which an officer’s movements are primarily constrained.

The operational activity and movement of law enforcement officers varies across the landscape of their jurisdiction based upon these sub-jurisdictional divisions. While an officer may respond to a crime in progress or a traffic accident in a neighboring precinct, he or she has primary responsibility for responding to reports of crime or disorder within his or her assigned precinct (or beat). As such, officers normally limit the number of activities that they become involved with outside of their assigned jurisdiction. A corollary is a scenario wherein police beats vary according to how police interact with the community. When foot patrols were the norm, individuals living in these communities would present their complaints directly to local precincts. However, with the advent of increased vehicle patrols and the technology associated with 911 calls, law enforcement administration has become more centralized and beats reorganized correspondingly.

The hierarchical culture of law enforcement additionally reinforces the boundaries of the precinct by fixing the first-line of command at the precinct level. Recent developments in law enforcement have applied other uses for this organizational entity. The popularity of COMPSTAT (Computer Comparison Statistics) (Bratton & Tumin, 2012; Walsh, 2001) has created an increased use of the precinct as an unit in the analysis of local crime levels and trends. In the COMPSTAT model, local precinct leaders (usually captains) are held accountable for the crime levels and trends within their area of responsibility. In the typical weekly COMPSTAT meeting, a Captain is asked to address the current reports of crime in the precinct. These reports are often displayed in a map to reflect their spatial dimension; however, the locations of crimes in neighboring precincts are not usually displayed at the same time because the goal of the COMPSTAT process is to highlight the accountability of the local law enforcement leadership rather than a neutral analysis of spatial patterns. This artifact of display can introduce an edge effect when crime hot spots span map boundaries. In addition, further analysis and aggregation from the beat level may introduce problems associated with the MAUP or the UGCoP.

The problems potentially introduced by borders can also be attributed to a greater degree of specialization of responsibility among different law enforcement agencies within the same geographic area. For example, a metropolitan area will have a primary police department that may handle policing services for most of the area. However, within that same geographic area, there could be law enforcement agencies responsible for a college campus, a hospital, or public transit. The sheriff’s department and State police operations may also have overlap. Not only do these overlapping boundaries and areas affect the movement of officers and their response to their environment, it may also affect the analysis of crime data. Much in the same way that the crime reports of a neighboring precinct may not be accounted for in crime pattern detection, the crime reports of agencies with concurrent jurisdiction may be absent from analysis or lead to conflation.

#### *4.2.2.3 The Boundary Role of the Patrol Officer*

Beyond the more concrete geographic representations of boundary and scale, boundaries can reflect social and power relations and the social construction of concepts, such as policing. The police as a

special group are given privileges that are not given to other members of the community. While the definition of policing itself can be a debated concept, there usually are three aspects at any definition's core—a right to detain, a right to use force under prescribed situations, and a right to generally demand compliance to their requests. There is a palpable sense of in-group and out-group, or us and them, when contrasting the police with private citizens, a sense more apparent in the present day; one which is amplified in the media (Klockars, 1985; Moore, Mark H. and Kelling, 1983; Vila & Morris, 1999). In this instance, the boundary between the groups is based more on each group's function. Within the police force there are individuals that serve a role at the margins between those two groups—or a boundary role. In law enforcement, it is often the patrol officer that serves in this boundary role.

The two main functions of the boundary role are information processing (ingress) and external representation (egress). The information processing function is a process by which the boundary role incumbents filter the information that is available in the environment and selects only what is deemed relevant information to be passed along to the rest of the organization. The boundary role has far-reaching implications starting with how officers scan their environment and their activity space. A patrol officers' response to their environment has direct consequences on how and which crimes are discovered and is much influenced by organizational goals and targets. Once a criminal incident has been identified, the officers are responsible for the recording of the information on the event according to organizational requirements and priorities. Which crimes are formally recorded and what information is provided on those crimes may vary from location to location and time to time.

The function of external representation is the boundary role incumbent's ability to achieve compromise between the environment and the organization's policies. A point of focus of this ability is in the interaction between community members and patrol officers. The patrol officer is responsible for adapting his or her behavior to both the policies of the agency and of the nature of the environment in which he or she is immersed. The joint effects of these interactions can affect how much crime is discovered and the willingness of victims to report crimes to the police. For example, the NYPD's decision to use a zero-tolerance policing strategy and how that is implemented by the patrol officers will result in a different reaction from the New York City community than in other cities that might be applying community policing tactics. The relationship that police maintain with the communities they serve can affect what crime is handled by police, how it is handled, and how the information about that crime is recorded.

Klinger (1997) uses the concept of the boundary role within the idea of negotiating (or triaging) the workload. Police introduce greater leniency where allowed by their administration to handle high crime areas. This is further moderated by police interpretation of the space they are policing. Policy may seek physical cues that imply a certain normative level of deviance. In addition, there are subjective factors of law enforcement perceptions of the "deserved-ness" of victims. Phillips and Sobol (2012) have been able to observe modest confirmation of Klinger's proposal for an inverse relationship where higher levels of CFS would result in lower levels of recording behavior.

### 4.3 METHODOLOGY AND DATA

Imputation of missing UCR data can be achieved by the production of a geospatial surface through interpolation facilitated by a dasymetric allocation method. Spatial interpolation can be used to infill for

locations that do not have measures, based upon the measured values of nearby locations referred to as control points (O'Sullivan & Unwin, 2010). From the measures at the control points generated by an appropriate allocation method, spatial interpolation uses a set of algorithms and calculations to arrive at an estimate for the unmeasured locations. The key to allocation methods is to identify appropriate ancillary variables to interpolate data from one set of source-date zones to a different, usually more granular set of zones. The goal of the analysis will be to assess what improvement, if any, there is to the final interpolation of missing data by using calls for service locations as a law enforcement agency correlate compared with the allocation of crime reports by population-related correlates such as streets. The analysis will test possible ancillary datasets to identify a candidate dataset to efficiently allocate crime data for a municipal law enforcement agency and a campus law enforcement agency. The results of the analysis will be compared to the actual data provided from the agencies to assess the relative accuracy of the allocated values from the comparison models.

#### 4.3.1 Dasymetric Mapping and Other Areal Interpolation Methods

Spatial interpolation "refers to the process of transferring data from one set of areas (source zones) to another (target zones)" (H. Kim & Yao, 2010, p. 5658). In the simplest of terms, interpolation comes in two varieties—simple and intelligent. Simple interpolation is the process of areal interpolation, which proportionally allocates data from the source zones into the target zones. Areal interpolation assumes a uniform distribution across the target zones. Because it does not consider the spatial distribution of underlying phenomena, it is seen as the method more prone to error. Tobler (1979) introduced an improvement on the simple areal interpolation by applying a smoothing density function over the target zone estimates. Smoothed pycnophylactic interpolation is achieved by taking the average of the neighbors of the target zone estimate and replacing the original estimate with the newly calculated average. However, to preserve the volume or counts of the original source data, the average values are proportionally adjusted up or down so that the zone's total is the same as the source data. This iterative adjustment in Tobler's method points to the main benefit of these two methods, which is the preservation of the original volume of data. Tobler's method is usually described as pycnophylactic (or volume-preserving) interpolation (Jang & Yao, 2011; H. Kim & Yao, 2010; Tobler, 1979).

Rather than assuming a uniform distribution to the data, intelligent interpolation, or dasymetric mapping, makes use of ancillary information to estimate the spatial distribution of underlying phenomena to modify the areal interpolation. The most commonly available ancillary data is land use/land cover (LULC) data that has been remotely sensed or otherwise mapped, as well as ambient lighting. The LULC data or street-lighting data is used to identify either inhabited or uninhabited areas or to assign an ordinal population density scheme to the target zones based on their suitability for inhabitation (Andresen & Brantingham, 2008; Jang & Yao, 2011; H. Kim & Yao, 2010; Sleeter & Gould, 2007; Zandbergen & Ignizio, 2010). Using this understanding of the relative distribution of the population, the source zone data can be proportionally allocated to the target zone. But rather than assume a uniform distribution, the ancillary data is used to assign different expected proportions for the source data given the likely residency patterns. While introducing an understanding of the potential spatial distribution of the source data has been shown to improve the predicted values in the target zones, there are still some problematic assumptions of dasymetric mapping. Dasymetric mapping relies on a measurement of density even though the source data may be discrete. Also, the densities can change abruptly from one zone to the next, which could be problematic if the underlying process or geography is not also subject to abrupt changes. The use of techniques, such as low pass filters, to

eliminate this problem can interfere with preserving the original count of the source data (H. Kim & Yao, 2010).

Despite these problems, dasymetric mapping is one of the more promising methods for transferring data from one zoning scheme into another, and recent scholarship has sought to improve upon the original methods or address some of the existing limitations (Jang & Yao, 2011; Mennis & Hultgren, 2006). Mennis and Hultgren (2006) tested the rather subjective determination of the functional relationship between the ancillary data and population density. Through a series of tests of three primary ways of defining the spatial relationship (containment, centroid, and percent coverage), the authors concluded that there is no one definitive spatial relationship that performs better than another across different variables. The functional relationship should be defined based upon the best understanding of the relationship between the ancillary data and the source data, which ties back to the issues raised by UGCoP. Finally, Kim and Yao (2010) proposed a hybrid method of both dasymetric and pycnophylactic methods to leverage the best aspects of each method—the volume preserving capabilities of pycnophylactic methods and the improved estimates with dasymetric mapping.

While dasymetric mapping and the future improvements that are currently in development are valuable methods towards solving the problem of analyzing data at different geographic levels or scale, it does not solve the problems described by MAUP, edge-effects and UGCoP. Ervin (2012) argues that “the MAUP (and by implication UGCoP) is an unsolvable issue”, in that “aggregating or grouping data will inevitably cause some loss of information or bias.” In each of these instances, the proper specification of the model is a critical step before any algorithm or calculation can be applied to data to predict values in a different geographic unit.

#### 4.3.2 Law Enforcement and Census Geography

The location for this analysis is the mid-Southern City of Knoxville, Tennessee. In 2014, Knoxville had a population of 187,347 and is situated on the Tennessee River in the Eastern portion of Tennessee. As the home of the main campus of the University of Tennessee, its population is diverse in comparison with the surrounding counties and can fluctuate between the summer months and the remaining portion of the year given a student population of over 28,000. While there are smaller colleges and campuses within the city of Knoxville (for example, Pellissippi State Community College), they were not included in this analysis. The city of Knoxville is primarily served by the Knoxville Police Department (KPD), which in 2014 employed 465 full time officers and civilian staff. In addition to the KPD, the University of Tennessee also maintains a police force of 93 to handle law enforcement activities on and near campus.

KPD divides its organizational space into two districts of approximately equal size. The East District comprises 51 square miles of primarily residential neighborhoods with approximately 85,000 residents. The East District is further divided into 12 beats comprised of primarily neighborhoods. The 52.25 square miles of the West District, in which an estimated 94,000 individuals reside, includes the Central Business District, the campus area of the University of Tennessee, and various shopping districts. The West District is additionally divided into 11 beats.

The city of Knoxville intersects with 89 Census tracts. Of these 89 tracts, only 36 are completely within the city boundaries because of the irregular shape of the city boundary of Knoxville. For this analysis the Census tracts were “clipped” to the city boundary to ensure comparable geography representing the

jurisdiction of the KPD between models. There are 84 complete or partial Census tracts with data used in the final model. Based upon the splitting of Census tracts since 1980, Knoxville has shown a pattern of suburban growth around the central business district.

### 4.3.3 Data

This quantitative geospatial analysis will rely upon point data from the same agencies' computer-aided dispatch (CAD) data and records management system (RMS) data for the same 12-month period. Data for this study consists of geocoded calls for service (CFS) and recorded incidents provided by the KPD for the entire calendar year of 2014. In 2014, there were 189,296 CFS recorded and geocoded, smaller set of 21,696 officer-initiated CFS, and 29,543 recorded incidents. There were higher amounts of CFS and recorded crime in the West District of KPD compared to the East District. Of the crimes recorded by KPD, the West District had four of the five beats in the top quartile and three of the six beats in the third quartile.

All CFS and recorded incidents were treated with equal weight regardless of the perceived severity of the call or crime. However, certain CFS were removed from the analytical file to include locations that were unlikely to reflect the location of a CFS. These CFS included mobile cellular towers, the police department headquarters, Knox County Communications Center, and any other record clearly marked "for record only" with generic location information, resulting in 189,296 incidents remaining to be analyzed. In addition to the crime information, the KPD also provided an ESRI shapefile constituting the boundaries of the police beats and districts in the year 2014. In addition to the law enforcement data described above, the study uses the street network as a method to identify areas of population traffic density. The road network was downloaded from the Census/TIGER website for the study year of 2014 (Census Bureau, n.d.). For this analysis, all data was projected into UTM 17N (1983).

Table 4-1. Descriptive statistics for CFS and recorded incidents

	Police Beats (n=23)			Census Tracts (n=84)		
	CFS	Officer-Initiated CFS	Recorded Incidents	CFS	Officer-Initiated CFS	Recorded Incidents
Mean	8,230.0	943.3	1,272	2,227.0	255.2	351.1
Percentile						
.05	5,463.0	403.8	905.1	78.8	5.0	4.1
.10	6,095.0	411.0	935.2	136.8	10.6	9.0
.25	6,843.0	559.8	1,087.5	612.0	39.0	57.0
.50	8,096.0	632.0	1,242.0	1,671.0	135.0	286.0
.75	9,252.0	968.0	1,461.5	3,205.0	334.0	515.0
.90	9,981.0	1,099.5	1,535.4	5,334.0	691.2	904.8
.95	11,094.0	1,441.0	1,719.6	6,680.6	892.2	1,086.4

The police beats of KPD subdivide the city into larger areas and have a larger average number of recorded crimes than Census tracts. In addition, there is less variation and a more even distribution of recorded crimes observed across police beats showing an indication of the organizational goal to manage the distribution of staff based upon crime data as well as a possible artifact of the necessity to clip certain Census tracts to fit the irregular shape of the city (See Table 4-1). While the range between the 5<sup>th</sup> and the 95<sup>th</sup> percentiles of recorded crimes in Census tracts span from 4.1 to 1,086.4 recorded crimes, the same range for police beats span from 905.1 to 1,719.6 recorded crimes. The lack of variation in recorded crime may suppress evidence of a relationship between CFS and recorded crime within this aggregate geography (See Figure 4-1 and Figure 4-2). Conversely, relationships between CFS—either total or officer-initiated—and recorded crime are strongly linear for Census tracts, but less so for the police beats of KPD (See Figure 4-3 and Figure 4-4). In both geographies of beats and tracts, there appear to be a few significant outliers. The predominant outliers in the bivariate relationship between CFS and recorded crime could be the result of major interstates or roads that could skew the relationship of CFS to recorded crime. In addition, the irregular geography of the boundary of the city of Knoxville contributes to the observed bivariate outliers. One Census tract outlier that has high levels of reported crime despite having comparatively low CFS does not fit this general case. However, recent racial shifts of this Census tract in the past five years may indicate some gentrification of the area lending further evidence to a differential law enforcement response (Times Herald, n.d.).

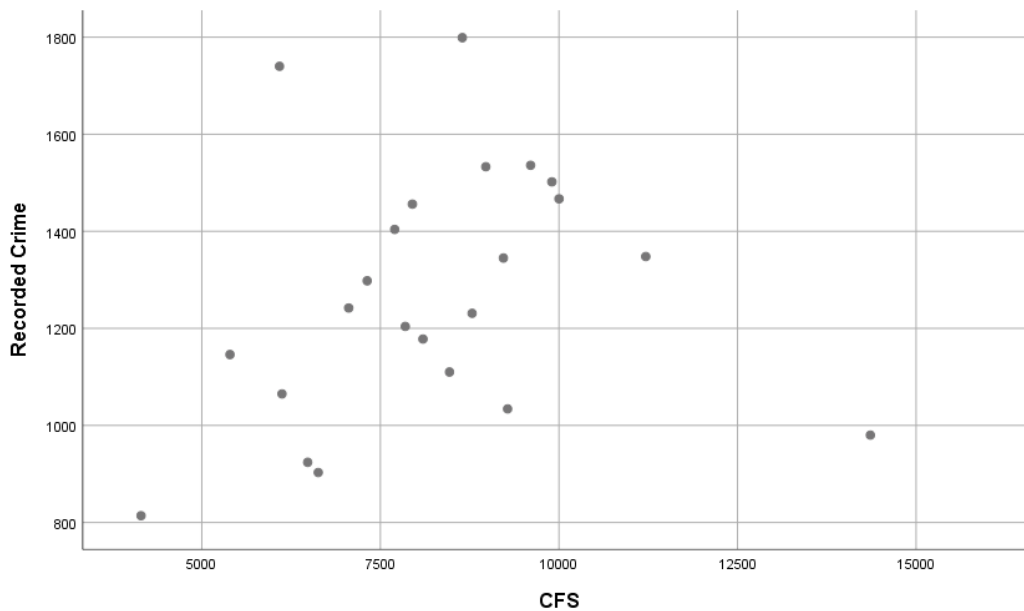


Figure 4-1. Scatterplot of recorded incidents by CFS aggregated by police beat

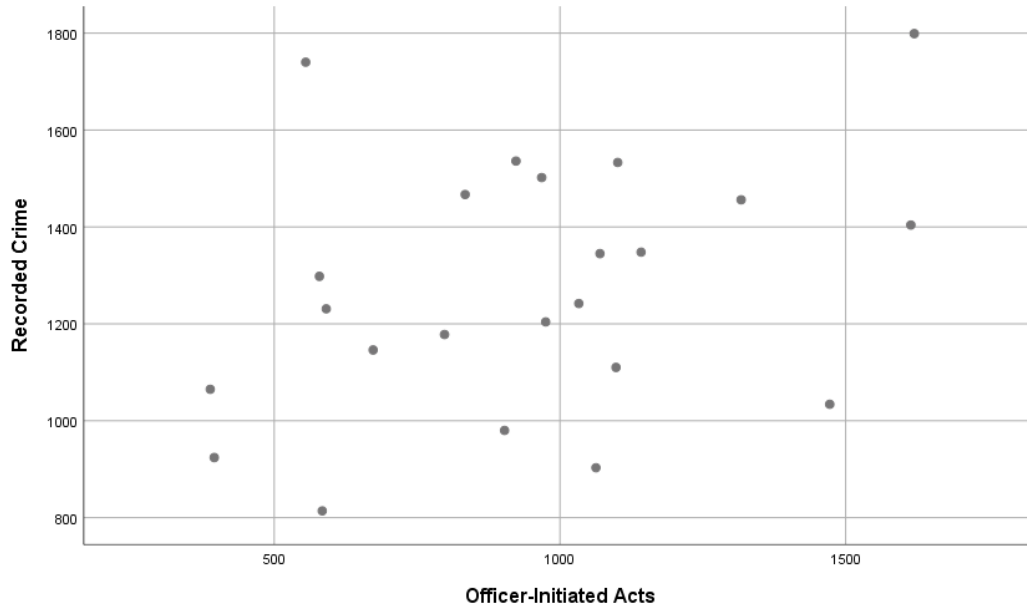


Figure 4-2. Scatterplot of Recorded Crime by Officer-Initiated Calls Aggregated by Police Beat

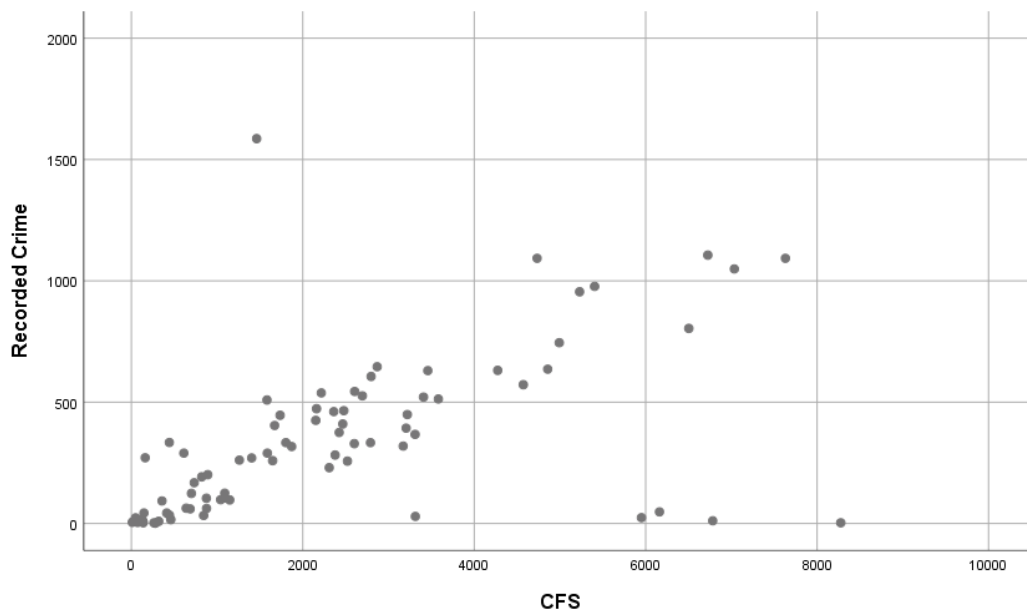


Figure 4-3. Scatterplot of Recorded Crime by CFS Aggregated by Census Tract

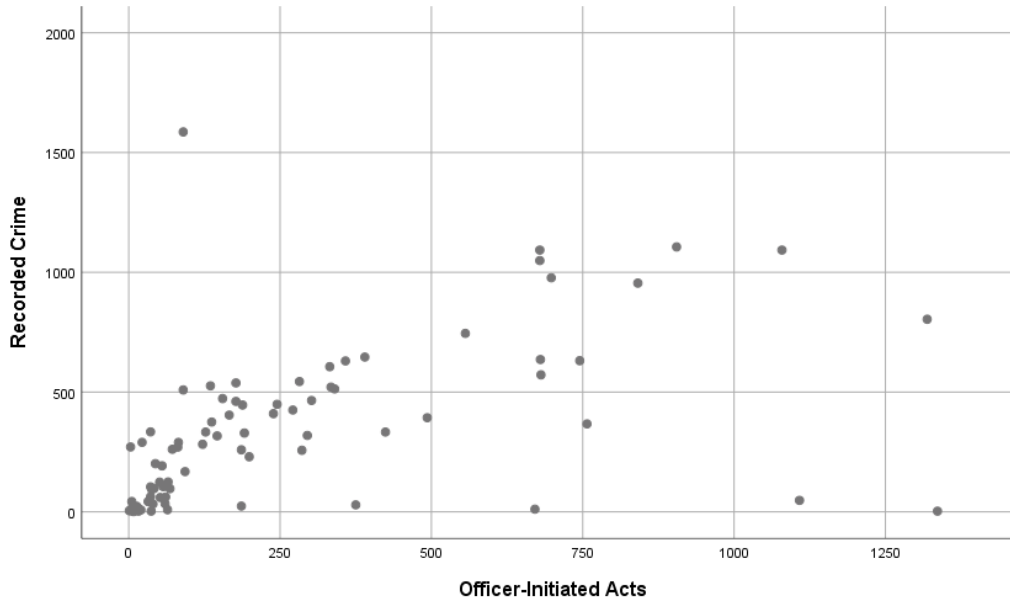


Figure 4-4. Scatterplot of Recorded Crime by Officer-Initiated Calls aggregated by Census Tract

#### 4.3.4 Study Methodology

The study compares the accuracy and sensitivity of two different ancillary datasets in their relative ability to allocate incident counts across a geographic area using a dasymetric allocation method. The dasymetric allocation was accomplished by using the Environmental Protection Agency’s Intelligent Dasymetric Mapping Toolbox for ArcGIS 10.3 (Environmental Protection Agency, 2018). This toolbox was based upon the Mennis and Hultgren (2006) “Intelligent Dasymetric Mapping” method. This method integrates domain knowledge into the model to improve the results available using a binary dasymetric methodology. Rather than assuming a uniform distribution to the data, intelligent interpolation, or dasymetric mapping, makes use of ancillary information available to estimate the underlying spatial distribution.

The three ancillary data sets compared in this study are the unweighted street network, the street network weighted through a spatial join of CFS to street segments, and the street network weighted by officer-initiated calls through a spatial join of calls to street segments. While the initial model with the unweighted street network will use a binary dasymetric method, each of the other two ancillary data sets will be allocated to either the police beat or the Census tract using an intelligent dasymetric approach where the researcher can preset expected densities associated with the ancillary data set (see Figure 4-5). These models are compared for accuracy and sensitivity using the Root Mean Squared Error (RMSE) and the Mean Absolute Percent Error (MAPE). Each measure of fit has strengths and weaknesses. The RMSE is known for its theoretical strength whereby model fit is evaluated by minimizing this distance to arrive at the expected value. However, because there is an “additional” penalty assessed for larger deviations, it would not be a robust measure against outliers. On the other, the MAPE has utility in assessing model fit across differing models and is less likely to be influenced by outliers since all deviations are weighted the same. It does not always identify the best fitting model. Both measures are presented.



#### 4.3.4.1 *Dasymetric allocation using street density*

The binary dasymetric method allocates the total crime volume for the city of Knoxville based upon the density of the street network as determined by a polyline raster of the street network downloaded from the Census/TIGER site. The dasymetric method starts by creating a raster of the feature class containing the population data—in this case, it would be the shape of Knoxville with the aggregate crime counts from the recorded data for 2014—that aligns with the ancillary dataset of the street raster file. Next, the dasymetric method will allocate crime volumes based upon preset densities. For the binary method, crimes were allocated to streets more associated with residences or businesses. No recorded crime volumes (which excludes traffic offenses) were allocated to other road types, such as interstates, major highways, off and on ramps, and off-road locations.

The dasymetric routine calculates a new crime density estimate if recorded crime only occurs where there were streets. The calculated “new density” is 3.454471 crimes per pixel. Each pixel is approximately 154.75 m<sup>2</sup>. To arrive at a new spatial density for Census Tracts and beats, the Zonal histogram tool was used to count the number of pixels in each of the other classes for either tract or beat, depending upon the case. The total pixel count was multiplied by 3.45571 (the dasymetrically-calculated spatial density) to produce an estimate for either the tract or the beat. The level of precision of the density was retained for the subsequent calculations to minimize rounding error in the final estimate.

#### 4.3.4.2 *Weighted Intelligent Dasymetric Models*

The difference between the binary unweighted dasymetric model and the “intelligent” weighted dasymetric models lies in the creation of classes for use in preset densities to account for anticipated differences in the distribution of recorded crime.

The preset densities were calculated from polyline kernel density estimates (KDE) of either CFS or officer-initiated calls joined to the underlying street segment as represented as address ranges by the Census/TIGER files. ArcGIS 10.5 uses the following formula to calculate the optimal bandwidth of the KDE.

$$\text{Search Radius} = 0.9 * \min \left( SD, \sqrt{\frac{1}{\ln 2} * D_m} \right) * n^{-0.2}$$

where:

*SD* is the standard distance


*D<sub>m</sub>* is the median distance

*n* is the number of points

Five classes were assigned to a surface raster based upon the results. The weights were assigned by examining the relative proportion of each of the classes of CFS. In other words, the predicted percent allocation weight for reported crime was set to mirror the measured percent allocation of the CFS for the same areas (see Table 4-2 for a list of the weights). The remaining dasymetric methodology remains the same as for the first model. The zonal histogram tool was used to collect the total number of pixels assigned to each of the five classes in the model. Each set of pixels was multiplied by the dasymetrically-

assigned density and summed within each tract or beat. As was the case in the binary approach, the level of precision for each of the densities was retained through the calculations to minimize rounding error.

*Table 4-2. Preset densities used in dasymetric mapping*

Class	Calls for Service	Officer initiated calls
Lowest crime density	0.036346	0.046783
	0.145147	0.150166
	0.267180	0.303051
	0.257985	0.282955
Highest crime density	0.293342	0.217045

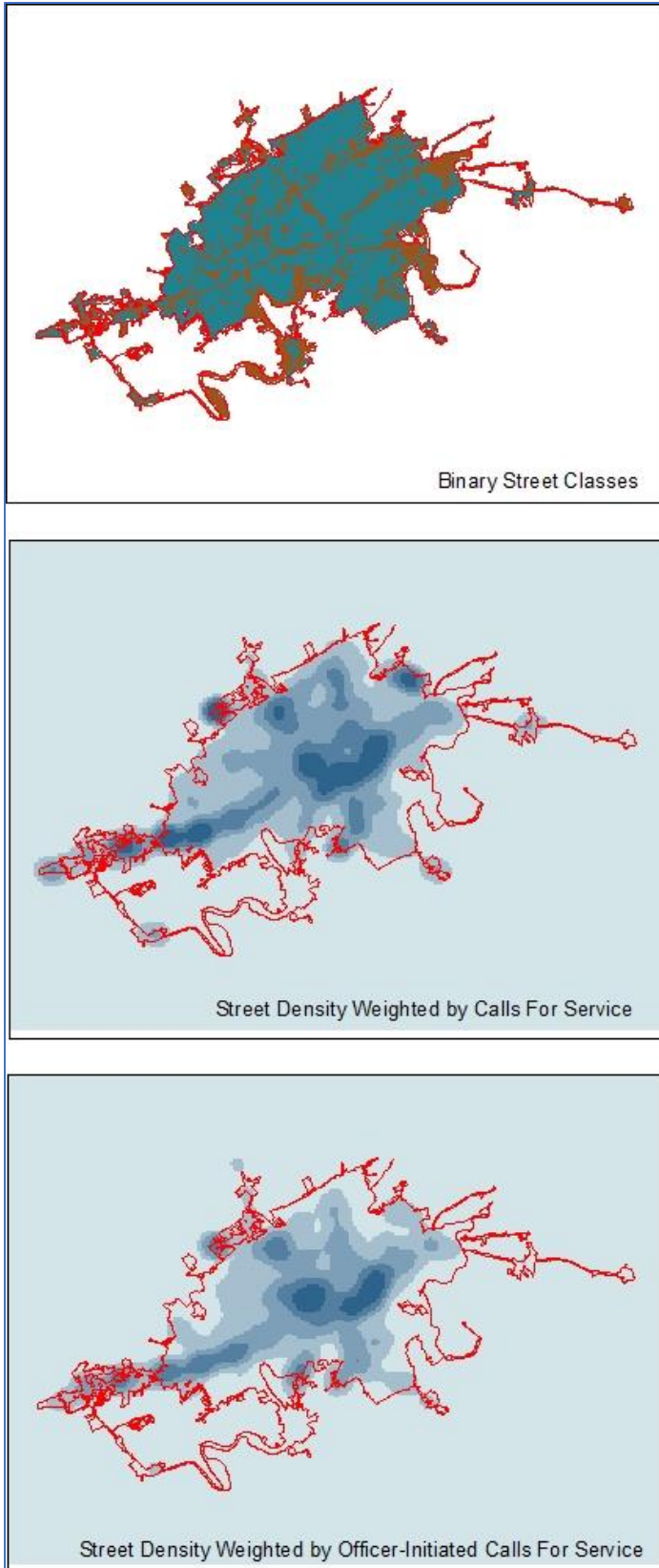


Figure 4-5. Density maps used as Ancillary Data for Dasymetric Models

#### 4.4 RESULTS

The results of the analysis illustrate the complex interplay of the ancillary data and the geographic aggregation of recorded crime data. Global measures of fit indicate that within both geographic units (Census tracts and police beats) that the use of ancillary data improved the fit over street density alone. Between the two weighted models, there is strong evidence that predictions of recorded incidents are more accurate using the locations of officer-initiated calls as opposed to CFS that would also include public requests for service. For both the MAPE and the RMSE, each value becomes progressively smaller with the introduction of the weighted models showing improved fit. The best fit was for the model using the officer-initiated acts (see Table 4-3 for measures). However, comparisons across geography using the more commonly used measure of RMSE are more difficult to interpret due to change in the change in scale and geography across models. The addition of the MAPE provides a unit-less measure of accuracy that allows for an easier comparison across models (Hyndman & Athanasopoulos, 2018).

Table 4-3. Goodness-of-fit for dasymetric-mapping produced predictions

	Beats		Tracts (Clipped)	
	MAPE	RMSE	MAPE <sup>a</sup>	RMSE
Binary	0.75	649.50	1.32	329.53
CFS-Weighted	0.53	554.52	0.67	279.26
Officer-Initiated	0.43	471.24	0.59	251.79

<sup>a</sup>The MAPE is sensitive to situations where the actual value is less than the predicted value. In those situations, a MAPE greater than 1.00 is possible.

Visual inspection of the maps of local calculations of the absolute percent error (APE) also indicates a better fit for both weighted models in comparison with the binary dasymetric model (See Figure 4-6 and Figure 4-7). Based upon the results of the weighted models, there are areas in central Knoxville that receive disproportionate number of CFS compared to recorded incidents. Conversely, areas at the edge of town appear to have fewer CFS than would be expected given their recorded crime. These results are mostly likely due to the pycnophylactic properties of dasymetric mapping. There is one area of Knoxville that attracts more CFS than would be expected based upon recorded incidents. To preserve the overall volume of the total reported crime for the city of Knoxville, recorded incidents are “borrowed” from areas that should have a higher density of reported crime. However, the RMSE certainly supports the use of the officer-initiated calls weighted model within Census tracts. However, some descriptions of RMSE caution against using this measure to compare across models. MAPE may be a better measure for comparison across models and for heteroskedastic models. Interpretation will be dependent upon the resolution of these potentially conflicting results. Based upon the APE measures, the use of the officer-initiated calls within police beats has the lowest percentages of error.

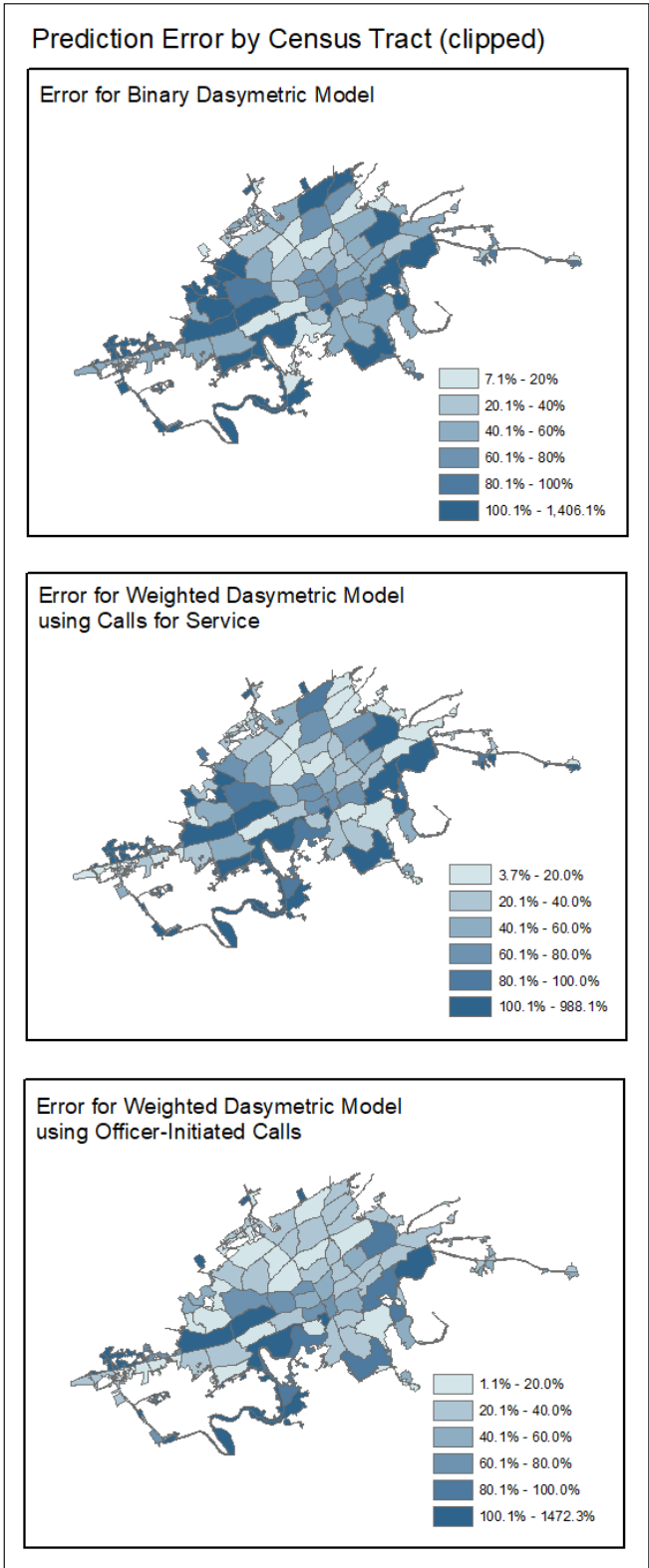


Figure 4-6. Prediction Error for Census Tracts (Clipped)

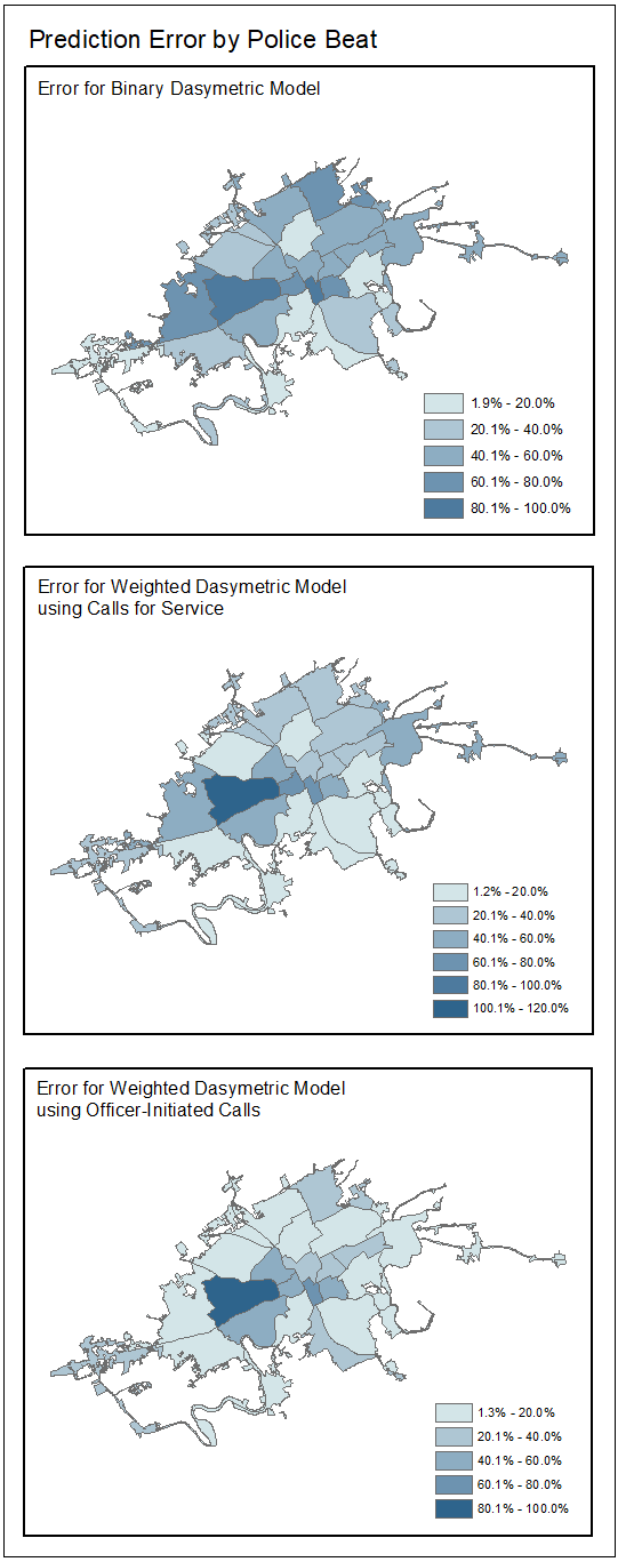


Figure 4-7. Prediction Error for Police Beats

#### 4.5 DISCUSSION AND CONCLUSION

Based upon the results of the dasymetric modeling, it appears the inclusion of law enforcement data—including the subset of officer-initiated activity—improves the fit of the dasymetric allocation of recorded crime data to both Census tracts and police beats. On both accounts there is significant evidence to support the inclusion of measures of law enforcement activity into models using recorded crime data. The best fit models for Census tracts and police beats used the subset of CFS that were clearly identified as officer-initiated. The officer-initiated activity may better represent the initial triaging of crime recording behavior proposed by Klinger (1997) and observed in subsequent research (Boivin & Cordeau, 2011; Phillips & Sobol, 2012). Given the close link between information gathering and processing by law enforcement officers in providing a boundary role, future models should focus more heavily on the spatial patterns of this subset of data when modeling data involving officer-recorded data.

The additional question of the most appropriate unit of aggregation to capture the relationship between law enforcement activity and data recording had more modest results. Both Census tracts and police beats are drawn for different purposes. Based upon their original intent, Census tracts are clearly a means to identify homogeneous communities for the measurement of key social and economic indicators (Census Bureau, 1994). On the other hand, police beats are used to organize law enforcement response and movement efficiently. In this case, police beats appear to have less error of fit in comparison to Census tracts. However, it should be noted that Census tracts do not perform badly in comparison (43.5 percent average error for beats versus 58.5 percent average error for tracts). Given the direct connection between law enforcement activity and beats/precincts and their close relationship to the processes at work, there is a slight advantage to using beats to avoid problems of the MAUP and UGCoP when using statistics based on administrative data. There may be no one “correct” unit of aggregation given that law enforcement activity is also deeply connected to their environment and the communities within them.

Finally, the study provides insight into the relative utility of dasymetric mapping as an interpolation method. While it should be noted that the best fitting model still had an average error rate of about 43 percent, most of the observed errors were likely introduced due to the pycnophylactic quality of dasymetric models because observed “borrowing” of recorded crimes from one unit of analysis to another. The approach should not be to abandon the volume preserving nature of the model given the need to keep the total crime accurate. Model fit is most likely improved by the introduction of an additional dimension to the ancillary variables or the incorporation of a nonlinear relationship between CFS and recorded crime (Klinger, 1997; Phillips & Sobol, 2012). Given the progressive improvement observed in the models by using data more closely associated with law enforcement activity, consideration should be given to the addition of ancillary variables that more directly connect law enforcement to its data uses as opposed to ancillary variables describing the population of these areas. A more process-oriented approach might be identified through a closer examination of ways in which law enforcement agencies use information to facilitate decision-making, as well as actual measures of the interaction between law enforcement and the community to improve model fit.

This study focused upon two essential geographic concepts as it relates to recorded crime data—the UGCoP and the MAUP. Both concepts seek to interrogate the data in terms of the appropriate environmental context and scale to better understand the complex relationship between law

enforcement activity and data recording behavior. The process of data collection is not a neutral process, and evidence from this analysis indicates that it has connections to the ways in which law enforcement organize themselves to manage workload, such as the subjurisdictional divisions of beats, as well as the discretionary behavior they exhibit in the act of recording information are key to improving modelling using recorded crime data. An expected strong relationship between CFS and recorded crime does not hold for this location regardless of the zoning pattern used in the analysis. However, the introduction of information on officer discretionary behavior improves the modeling and provides critical insight to future ways of modeling crime data. In addition, other methods, such as regression, could be used to improve the understanding of the relationship between CFS and recorded incidents for inclusion in spatial models.

Future research on this topic should continue to expand upon the interaction between the officer and his or her environment to identify ways to incorporate this context into crime analyses. The selection of the optimal scale to avoid issues related to MAUP or UGCoP should not just solely focus upon the perceptions of space or place by the residents, but also the perceptions of space or place by law enforcement officers responsible for recording criminal events. It is not enough to simply know how much law enforcement activity occurs in a place as evidenced by the number of CFS, it is important to understand how officers respond to a place and ultimately process and shape the information about that place and its public safety events. Understanding the interplay between law enforcement activity and place will become more important once the UCR Program shifts to incident-based data. While one study alone does not conclusively resolve the question of the impact of law enforcement activity and crime data on crime modeling, the results clearly point to a need to incorporate a fuller and more nuanced understanding of the spatial production of crime data.



## 5 CONCLUSION

---

The initial intent of this research was to explore the extent to which differences among law enforcement agencies and their discretionary practices associated with policing could be measured and accounted for in crime data and models, such as imputation and interpolation of crime data. At the start of this project, there were three primary research questions:

- What characteristics can be used to identify peer agencies for crime reporting data given the current diversity of policing? What spatial and other dimensions seem to matter and at what scale?
- Can information on the location of calls for service (CFS) improve the prediction of law enforcement recording of crime incident data for crime types identified as priorities through an assessment of analytical products? Can differences among law enforcement agency types be measured in the spatial properties of recorded incidents?
- Crime statistics are a result of police activities connected to both calls for service (CFS) and officer-initiated acts. CFS represent “demand” for police services, but what can be said about discretionary behavior as measured by officer-initiated activities? Do officer-initiated activities exhibit the same spatial patterns as CFS?

The results of three papers support the contention that there is a greater need to account for the diversity in law enforcement in terms of the communities they serve and how those differences may impact the types of data recorded and discretionary recording behavior. To improve our ability to construct robust models using law enforcement data, there needs to be greater attention to uncertainty created by the heterogeneous nature of law enforcement practices in the 21<sup>st</sup> century. Beyond the initial conclusions that can be derived from this research, the results also contribute to a larger conversation embracing “Big Geodata”, the increasing role of place in GISc. and criminology, and the role of individuals in constructing place through their roles and behaviors.

### 5.1 BIG GEODATA

In 2016, the Federal Bureau of Investigation announced that the traditional way of collecting crime data as agency-level counts would be retired in January of 2021. Beginning in 2021, the only way law enforcement agencies will report crime data to a national statistical program will be in an incident-based format (National Incident-Based Reporting System or NIBRS) as opposed to the tallies and counts most users of UCR data have become accustomed to. (Federal Bureau of Investigation, 2016). In conjunction with a modernization assessment by the National Academies of Sciences, Engineering, and Medicine, this program initiative renewed calls for more detailed georeferencing of crime data in the national crime dataset (Lauritsen & Cork, 2016). Point-level data rather than agency-level data will become the norm. Research in criminology is also increasingly focused on more granular locations or “microgeographies” of neighborhoods or street segments and crime data will properly be identified as “Big Data.”

Big Data has usually been identified by the three “Veas” of volume, velocity, and variety. These characteristics describe data that is large in size (terabytes or petabytes), available in real-time or near-real time, and often in both structured and unstructured formats (Kitchin, 2014). In recent years however, the additional “Veas” of veracity and value have been added to augment conversations about the appropriate use and interpretation of Big Data (Michael F. Goodchild, 2018b; Miller & Goodchild, 2014). Big Data is most often opportunistic which has minimal control over its collection and a great deal of unmeasured uncertainty (Wender, 2017). Certainly, local law enforcement data exhibit the characteristics as other opportunistic data. However, NIBRS data, which is a statistical collection generated from law enforcement data, also have detectable vestiges of the uncertainty of opportunistic data. While there is a general enthusiasm in the research community to have easy access to large datasets, there are new considerations necessary to be taken of the quality and the context in which opportunistic data are collected when law enforcement data is used.

Goodchild (2013) notes that most Big Data are georeferenced, that is Big Geodata. The heterogeneity of geographic data presents unique challenges for understanding the potential spatial bias and representativeness of Big Geodata (Michael F. Goodchild, 2013; Miller & Goodchild, 2014; Zhang & Zhu, 2018). The less authoritative or volunteered nature of much of Big Geodata would make it difficult to apply techniques such as spatial sampling given the underlying spatial bias introduced by the collection methods (Zhang & Zhu, 2018). To some degree UCR provides a framework of definitions and concepts to apply to crime data regardless of where it was collected. It thereby acquires authority, even if this serves to obscure a real heterogeneity. Using data provided directly from law enforcement systems removes this layer of standardization and quality control available through statistical or “smaller” data collections, such as UCR or NIBRS. In the midst of this debate of the “digital turn” in data and data collections, there are possibilities of leveraging the strengths of both “big”/“fast” data and “small”/“slow” data (Goodchild, 2013; Miller, 2017, 2018; Miller & Goodchild, 2014). However, it is important to first recognize the potential limitations associated with crime data that could be introduced from macro-level processes, such as structural societal issues in the criminal justice system (Fyfe, 1991; Lowman, 1986), or micro-level processes associated with discretionary decision making based on law enforcement agency priorities (Boivin & Cordeau, 2011; Klinger & Bridges, 1997; Phillips & Sobol, 2012). The “digital turn” with access to Big Data and Big Geodata redoubles concerns for increased critical understanding of whose needs are being met and represented through the collection of the data and how access to the data effects how this data should be viewed (Ash et al., 2018; Boyd & Crawford, 2012; Elwood & Leszczynski, 2013; Kitchin, 2014; Kwan, 2016; Mayernik, 2017; Miller & Goodchild, 2014).

The “Linus Law” states that data quality can improve and achieve faster convergence with more “eyes” that see it, but that goal may not be able achievable with crime and other geodata (especially very local data) because there may not be a sufficiently large group reviewing the data to improve its accuracy (Goodchild, 2013; Miller & Goodchild, 2014). Handling the heterogeneity of the data and finding ways to manage uncertainty will become increasingly important. In the first paper I demonstrated that end users can leverage techniques such as jackknifing and other types of simulation to address internal variation and quantify the uncertainty of results for measurements and data that have unknown distributions (Barnett-Ryan, 2019b). The analysis showed that removal of *zero-population* agencies from current classification schemes for law enforcement agencies homogenizes the average coefficient of variation across all offenses. This increased similarity means that measures of central tendency

calculated within each stratum are more precise and ultimately improve modeling of agency-level crime data. Even at aggregate levels and scales, the impact of law enforcement heterogeneity can be measured and is contextualized by the type of agency.

The open data initiatives of many law enforcement agencies provide means to acquire detailed georeferenced data collected as a part of normal administrative and operational functions. However, this opportunistic data should still be examined for a greater understanding of the political and social context in which the data is recorded as noted by earlier observers (Harries, 1975; LeBeau & Leitner, 2011; Lowman, 1986; Peet, 1975) and current researchers (Lynch, 2018). The underlying relationships between such data as calls for service (CFS) and recorded incident data that have been observed at aggregate levels may not hold for the newer microgeographies of street segment, neighborhood or point data. I found in the second paper that recorded incident data, which serves as the basis for crime statistics, of property crime showed only minimal co-location with similar types of CFS for a municipal law enforcement agency and no appreciable co-location between property crime incidents and CFS for a campus law enforcement agency (Barnett-Ryan, 2019a). In addition, in my third paper, CFS show only a modest ability to be used as an ancillary variable for dasymetric allocation models of crime data when constrained to those that are labelled as “officer-initiated” calls (Cynthia Barnett-Ryan, 2019c). In totality, these results support a finding that the societal and organizational processes that lie behind recording CFS are fundamentally different from those underlying recorded incidents. Therefore, it is unlikely that CFS would be a robust proxy for recorded incidents or current statistical measures.

## 5.2 THE PLACE OF THE INDIVIDUAL

As spatial resolution becomes finer, as in micro-geographies, the relationship between space and place is brought to the forefront. Recent scholarship on the criminology of place as exemplified by Weisburd and his colleagues (Braga, Andresen, & Lawton, 2017; Rengert & Lockwood, 2009; Tita & Greenbaum, 2009; D. Weisburd, 2015a) focuses on the experience of space from the perspective of the residents. This dovetails with consistent themes in criminology related to collective efficacy, informal social control, social disorganization, and guardianship (Fyfe, 1991; Hollis-Peel, Reynald, van Bavel, Elffers, & Welsh, 2011; Sampson & Wilson, 2012; Sampson, 2013; Schnell, Grossman, & Braga, 2018). However, crime data are not collected by the residents of a location and do not completely reflect their characteristics, if at all. The reports of crime events by residents are mediated by law enforcement officers and their discretionary decision-making choices for an event to become a matter of record. The seminal piece by Klinger (1997) describes the experience of space (or place) from a law enforcement perspective. In his ecological perspective, Klinger outlines expectations for how law enforcement officers interpret disorder or lack of disorder in the environment. His description of how law enforcement officers navigate and experience space challenge the usual assumptions regarding the standardization of recorded incident data.

At the organizational level, law enforcement agencies and officers work from different sets of assumptions about crime and the data needed for operational effectiveness. Sometimes these assumptions are based upon internally-driven priorities, such as the desire by the Chief or Sheriff to implement a COMPSTAT-style approach of accountability for mid-level managers. At other times, they can be externally-driven such as the mandates for data collection and reporting for campus law enforcement under the Clery Act. They may originate from concerns of the society at-large. These organizational priorities have the potential to create artifacts in the data collection process itself.

For example, the concept of a boundary role for an individual introduces the concept whereby an individual mediates both the collection of information from the environment and the communication of information within and between groups (Aldrich & Herker, 1977). In the case of law enforcement, patrol officers or deputies serve in the boundary role as the first responders to reports or observations of crime and disorder. In this dissertation I found preliminary evidence of differential approaches to data recording and analysis in the assessment of data and analytical products of two law enforcement agencies serving the same geographic location. The campus law enforcement agency recorded proportionally more information regarding events related to alcohol and drug-related incidents, for example, than did the municipal law enforcement agency (Barnett-Ryan, 2019c). In addition, the stronger association of officer-initiated CFS to recorded incidents lends strength to Klinger's conjectures regarding the process by which law enforcement may be deprioritizing the recording of crimes in areas where law enforcement officers expect to see high levels of disorder (Barnett-Ryan, 2019c; Boivin & Cordeau, 2011; Klinger, 1997; Phillips & Sobol, 2012).

Boundaries can also represent the "crispness" or "fuzziness" of particular social concepts—or in this case the concept of crime. In many ways, crime can be considered a boundary object. A boundary object has three main qualities—interpretive flexibility, "the material/organizational structure of different types of boundary objects and the question of scale/granularity" (Leigh Star, 2010, p. 602). In essence, boundary objects have meanings that may vary depending upon who is using the term, what purpose it is being used for, and level of detail or granularity of its use. What is deemed a "crime" is highly dependent on these three aspects and can be interpreted differently depending upon the context. The idea of boundaries comes into play when these concepts mediate between various groups. For example, the recording of crime statistics occurs in a particular law enforcement context for the use of law enforcement. Police bring their own understanding of the idea of criminal, which is deeply immersed in the law, courts, and a sense of morality. On the other hand, once crime data is recorded, non-law enforcement groups can use it for different purposes. The definitions applied at the time by law enforcement take on diverse meanings for these groups. While the alternate definitions of crime that are used can be points of contention, they are nonetheless points of boundary work between these groups.

Interpretations of crime and criminal behavior can vary in many contexts. For example, the same behavior can take on different interpretations depending upon the group performing the interpretation. This is the case when the use of controlled substances can be both seen as a crime and a health problem depending upon whether it is the courts or the medical establishment viewing this behavior. The use of drugs can be seen differently depending upon the location. In the states of Washington and Colorado, recreational marijuana use has been legalized by the state, but not for other locations. Finally, interpretations can vary from one time to another. For example, prostitution as a behavior was once classified as criminal in the sense that prostitution was a criminal act and the prostitutes were seen as criminals. However, in recent years, this same behavior—especially in situations involving juveniles—has now been reframed as human trafficking in which the same individual is now classified as a victim of a crime rather than an offender.

Finally, the results of the three studies lend insight into the construction of place from crime data. At the three scales at which these studies analyzed crime data, there is evidence of different approaches to law enforcement by law enforcement. The analysis of agency-level data was able to detail the amount of variation and heterogeneity of crime data based on geographic region and size, but most especially

for agencies serving specialized communities such as campus law enforcement, park police, or transit police. At the agency level, the aggregate data shows evidence of the differential experience of law enforcement given a specialized community or location that they serve (Barnett-Ryan, 2019b). Moving into the next smallest level of geography of subjurisdictional divisions, the findings from Barnett-Ryan (2019c) indicates that crime data appears to be more reflective of agency geographic organization. There is improvement in those dasymetric allocation models using law enforcement geographic units such as precinct or beat, compared to geographic units based upon residential populations like Census tracts. Finally, the minimal association of point CFS and recorded incident data challenges not only assumptions of the interchangeability of both datasets, but also the idea that construct of a “call for service” or “criminal incident” would be the same regardless of law enforcement agency type (Barnett-Ryan, 2019a).

### 5.3 FUTURE RESEARCH ON PLACE, SPACE, AND THE “PLACIAL TURN”

Future research on the “placial turn,” now being experienced in the disciplines of geography and criminology, not only focuses on ever increasingly high spatial resolution to improve analyses, but also on reaffirming the role of place in addition to space as a central concept to be considered and potentially formalized in analysis. Rather than captured in crisp areal units or geographic boundaries and concepts, more subjective and less distinct human and social properties come into play (Merschdorf & Blaschke, 2018). Place-based geographic information science (GISc.) runs the spectrum from questions associated with automatic modeling and the semantic web (Ballatore, 2016; Blaschke et al., 2018; McFarlane & Anderson, 2011) to critical GISc. (Elwood & Leszczynski, 2013; Leszczynski, 2012; Wilson, 2011b, 2011a, 2017). However, the digital representation of place presents considerable challenges to develop unambiguous definitions and formalizations of place.

The formalization of place in GISc. has recently delved into accounting for uncertainty when handling subjective or vague data—also noted as a concern with the use of Big Geodata—as well as the transferability of results. A consistent theme of accounting for uncertainty and the impact of data collection and data aggregation methods on the conclusions to be taken from the analysis of law enforcement data. The modifiable areal unit problem is one aspect of uncertainty-associated artifacts created in the data by either zoning or aggregation schemes. Research also continues to explore “fuzzy” boundaries or “fuzzy” topology in regards to spatial data analysis to integrate information on the local backcloth and avoid problems associated with abrupt transitions with aggregation units (Brantingham, Brantingham, Vajihollahi, & Wuschke, 2009). Future research should seek ways to incorporate these methods when modeling law enforcement data even where data are sparse.

In the case of law enforcement data, relationships that are statistically significant at the larger area of the jurisdiction (Barnett-Ryan, 2019b, 2019a) become significantly weaker at more granular levels of data collection. This is especially pronounced when looking for significant co-location of property crimes as captured through CFS and recorded incidents (Barnett-Ryan, 2019a). The findings associated with law enforcement data touch on an additional concept of uncertainty—the uncertain geographic context problem (UGCoP).

UGCoP conceptualizes the idea that the measurable relationship among different variables is dependent upon the scale and geographic area at which the process is occurring (Kwan, 2012b, 2016, 2018). Given the modest result of the dasymetric modelling using CFS as an ancillary variable and its overall

improvement when law enforcement discretionary officer-initiated events are used, the findings point to a need to identify contextual information on law enforcement engagement with their environment (Barnett-Ryan, 2019c; Klinger, 1997; Klinger & Bridges, 1997; Phillips & Sobol, 2012) or even the incorporation of quantified aspects of uncertainty (Nagle, Battenfield, Leyk, & Spielman, 2014). Findings point to the need for a closer examination of ways in which law enforcement agencies use information to facilitate decision-making could provide better insight into their discretionary decisions that have the potential to impact recorded incident data (Barnett-Ryan, 2019c).

Following this train of thought, the formalization of place also raises questions regarding the transferability of results (Blaschke, Merschdorf et. al. 2018). The transferability of results can manifest itself in a broad cultural fashion, such as when language does not precisely distinguish between place and space (Blaschke et al., 2018) or in the culture of professional or the user of the term. For example, Weisburd and his colleagues (Braga et al., 2017; Tita & Greenbaum, 2009; D. Weisburd, 2015a) use the street segment as a proxy for place, but this is based on assumptions of how residents engage with their space. It is not entirely clear that Weisburd is invoking the term, “place” in the same context as a professional geographer or patrol officer might. In addition, the analysis of even agency-level data makes assumptions regarding the typology of law enforcement agencies as a place in a state, a geographic division, or region or by the community it serves (Barnett-Ryan, 2019b). As interdisciplinary research increases between geography, criminology, and law enforcement, the concept of place

## 5.4 DISCUSSION

This study is concentrated on information in a limited geographic area, and, as such, the conclusions drawn from this study will be limited. Over recent years, the law enforcement community has been increasing its transparency and accessibility to its data. However, detailed studies of law enforcement agencies will require a level of trust that can be difficult to obtain over wide geographic areas. In order to further research in this subject, there should be an effort to replicate these results using different law enforcement data from different areas. While one research study alone cannot conclusively resolve the questions related to the impact of law enforcement activity and crime data on crime modeling, the results of these three papers confirm that spatial and statistical bias introduced into crime data is measurable and is associated with law enforcement behavior. In addition, I argue that the modest success in capturing and controlling for these systematic effects is possible with better information of law enforcement context surrounding CFS, recorded incident information, and the discretionary behavior of officers. However, these results are far from conclusive and raise questions, opening the door for more in-depth research on place-based criminology and place-based policing and the questions generated by place-based approaches. Law enforcement’s understanding of place solely influences crime recording, but the reaction to crime data through proactive policing measures also continue to fulfill expectations of measured crime patterns by concentrating the attention of officers on locations and on types of offenses.

A more nuanced understanding of the impact of agency type and their policing activities should be introduced into models using official crime statistics. Evidence increasingly makes it difficult to maintain that law enforcement is a monolithic entity acting uniformly, especially given the increased specialization across agencies. Beyond official crime reports, future research should also incorporate a more nuanced approach to understanding the context in which CFS and recorded incident data is

captured; especially if these data is gathered directly from operational systems through open or other 'Big Data' initiatives.

The concept of boundaries in their less obvious forms should be further explored as well. In addition to the influence asserted over crime data by the law enforcement perspective, variation in agency priorities should be accounted for, as the variation in agencies becomes increasingly diversified and specialized in mission. Agency engagement with data should be considered as a possible influence on the recording of incident data. Future research should explore means to formalize or quantify the uncertainty introduced into crime models by agency analytical priorities and the function of the officer in a boundary role. When functioning in the boundary role for recording crime information, patrol officers are more likely to record information of interest to their command staff than the complexity of the actual situation. Finally, the concepts associated with crime data should be investigated for differences in interpretation and for ways to expand the incorporation of "fuzzy" definitions for space, place, and even time as it relates to criminal events. This study contributes to current discussions in geography and criminology regarding ways in which uncertainty can be measured and accounted for. This will be critical as use of Big Geodata becomes more common, as well as an increased amount of georeferenced data are recorded and analyzed by non-geographers, such as with volunteered geographic data and neogeography. This study also challenges criminologists to consider the interaction of scale and context of place in the interpretation of more granular, georeferenced crime data recorded by law enforcement. Defining crime concentrations at a place or scale without engaging in serious debate about the underlying production of the data used for the analysis reduces the true complexity of the question.

## 6 CITATIONS

---

- Addington, L. A. (2004). The Effect of NIBRS Reporting on Item Missing Data in Murder Cases. *Homicide Studies, 8*(3), 193–213. <http://doi.org/10.1177/1088767904265360>
- Aldrich, H., & Herker, D. (1977). Boundary Spanning roles and Organization Structure. *Academy of Management Review, 2*(2), 217–230. <http://doi.org/10.5465/AMR.1977.4409044>
- Alpert, G. P., Macdonald, J. M., & Dunham, R. G. (2005). Police Suspicion and Discretionary Decision Making During Citizen Stops. *Criminology, 43*(2), 407–434. <http://doi.org/10.1111/j.0011-1348.2005.00012.x>
- Andresen, M. A., & Brantingham, P. L. (2008). Visualizing Ambient Population Data within Census Boundaries: A Dasyetric Mapping Procedure. *Cartographica: The International Journal for Geographic Information and Geovisualization, 43*(4), 267–275. <http://doi.org/10.3138/carto.43.4.267>
- Andresen, M. A., Curman, A. S., & Linning, S. J. (2017). The Trajectories of Crime at Places: Understanding the Patterns of Disaggregated Crime Types. *Journal of Quantitative Criminology, 33*(3), 427–449. <http://doi.org/10.1007/s10940-016-9301-1>
- Andresen, M. A., Linning, S. J., & Malleson, N. (2017). Crime at Places and Spatial Concentrations: Exploring the Spatial Stability of Property Crime in Vancouver BC, 2003–2013. *Journal of Quantitative Criminology, 33*(2), 255–275. <http://doi.org/10.1007/s10940-016-9295-8>
- Andresen, M. A., & Malleson, N. (2011). Testing the stability of crime patterns: Implications for theory and policy. *Journal of Research in Crime and Delinquency, 48*(1), 58–82. <http://doi.org/10.1177/0022427810384136>
- Arson. (2012). Retrieved February 4, 2017, from <https://ucr.fbi.gov/crime-in-the-u.s/2012/crime-in-the-u.s.-2012/property-crime/arson>
- Ash, J., Kitchin, R., & Leszczynski, A. (2018). Digital turn, digital geographies? *Progress in Human Geography, 42*(1), 25–43. <http://doi.org/10.1177/0309132516664800>
- Baddeley, A., Rubak, E., & Turner, R. (2015). *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press.
- Ballatore, A. (2016). Prolegomena for an ontology of place. *Proceedings of the Vespucci Institutes, 138*(4), 91–103. Retrieved from <https://escholarship.org/uc/item/0rw1n045>
- Barnett-Ryan, C. (2006). *Introduction to the uniform crime reporting program. Understanding Crime Statistics: Revisiting the Divergence of the NCVS and UCR*. <http://doi.org/10.1017/CBO9780511618543.003>
- Barnett-Ryan, C. (2007). Introduction to the Uniform Crime Reporting Program. In J. P. Lynch & L. A. Addington (Eds.), *Understanding Crime Statistics* (pp. 55–89). New York, NY: Cambridge University Press.
- Barnett-Ryan, C. (2019a). *Crime and Place: Differences in Spatial Relationship between Calls for Service and Recorded Incidents for Municipal and Campus Law Enforcement*. Morgantown, WV.
- Barnett-Ryan, C. (2019b). *Missing Data in Uniform Crime Reporting and Impact of Agency Type*.



Morgantown, WV.

- Barnett-Ryan, C. (2019c). *The Recording of Crime Data and Scale: The Influence of Law Enforcement Recording Behavior and the Uncertain Geographic Context Problem*. Morgantown, WV.
- Biderman, A. D., & Reiss, A. J. (1967). On Exploring the “Dark Figure” of Crime. *The ANNALS of the American Academy of Political and Social Science*, 374(1), 1–15.  
<http://doi.org/10.1177/000271626737400102>
- Blaschke, T., Merschdorf, H., Cabrera-Barona, P., Gao, S., Papadakis, E., Kovacs-Györi, A., ... Kovacs-Györi, A. (2018). Place versus Space: From Points, Lines and Polygons in GIS to Place-Based Representations Reflecting Language and Culture. *ISPRS International Journal of Geo-Information*, 7(11), 452. <http://doi.org/10.3390/ijgi7110452>
- Boivin, R., & Cordeau, G. (2011). Measuring the Impact of Police Discretion on Official Crime Statistics: A Research Note. *Police Quarterly*, 14(2), 186–203. <http://doi.org/10.1177/10986111111404179>
- Boos, D. D., & Osborne, J. (2013). *The Jackknife in Monte Carlo Studies*. Raleigh, NC.
- Boyd, D., & Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. <http://doi.org/10.1159/000477440>
- Braga, A. A., Andresen, M. A., & Lawton, B. (2017). The Law of Crime Concentration at Places: Editors’ Introduction. *Journal of Quantitative Criminology*, 33(3), 421–426. <http://doi.org/10.1007/s10940-017-9342-0>
- Braga, A. A., Papachristos, A. V., & Hureau, D. M. (2014). The Effects of Hot Spots Policing on Crime: An Updated Systematic Review and Meta-Analysis. *Justice Quarterly*, 31(4). <http://doi.org/10.1080/07418825.2012.673632>
- Brantingham, P. L., Brantingham, P. J., Vajihollahi, M., & Wuschke, K. (2009). Crime Analysis at Multiple Scales of Aggregation : A Topological Approach. In D. Weisburd, W. Bernasco, & G. J. N. Bruinsma (Eds.), *Putting Crime in its Place* (pp. 87–107). New York, NY: Springer New York. <http://doi.org/10.1007/978-0-387-09688-9>
- Bratton, W., & Tumin, Z. (2012). *Collaborate or Perish!: Reaching Across Boundaries in a Networked World*. Crown Business.
- Bromley, M. L. (2007). The Evolution of Campus Policing: Different Models for Different Eras. In B. S. Fisher & I. Sloan, John J. (Eds.), *Campus Crime: Legal, Social, and Policy Perspectives* (2nd editio, pp. 280–303). Springfield, IL: Charles C. Thomas.
- Burrough, P. A., & McDonnell, R. A. (1998). *Principles of Geographical Information Systems*. Oxford: Oxford University Press.
- Census Bureau. (n.d.). Census TIGER Geodatabases. Retrieved December 16, 2018, from <https://www.census.gov/geo/maps-data/data/tiger-geodatabases.html>
- Census Bureau. (1994). *Geographic Areas Reference Manual*. Washington, DC. Retrieved from <https://www.census.gov/geo/reference/garm.html>
- Clery Center. (n.d.). Retrieved March 2, 2018, from <https://clerycenter.org/>

- Conley, J., & Stein, R. (2014). Spatial Analysis of Fear of Crime and Police Calls for Service: An Example and Implications for Community Policing. In G. Roedl, G. A. Elmes, & J. Conley (Eds.), *Forensic GIS: The Role of Geospatial Technologies for Investigating Crime and Providing Evidence* (pp. 19–39). Dordrecht: Springer Science+Business Media B.V. <http://doi.org/10.1007/978-94-017-8757-4>
- Cooper, A. D., & Hyland, S. S. (2015). *Justice Assistance Grant (JAG) Program, 2015*. Washington, DC.
- Currie, M., Paris, B. S., Pasquetto, I., & Pierre, J. (2016). The conundrum of police officer-involved homicides: Counter-data in Los Angeles County. *Big Data & Society*, 3(2), 2053951716663566. <http://doi.org/10.1177/2053951716663566>
- Desmond, M., Papachristos, A. V., & Kirk, D. S. (2016). Police Violence and Citizen Crime Reporting in the Black Community. *American Sociological Review*, 81(5), 857–876. <http://doi.org/10.1177/0003122416663494>
- Dixon, P. M. (2002). Ripley's K function. In *Encyclopedia of Environmetrics* (Vol. 3, pp. 1796–1803). <http://doi.org/10.1002/9780470057339.var046>
- Elwood, S., & Leszczynski, A. (2013). New spatial media, new knowledge politics. *Transactions of the Institute of British Geographers*, 38(4), 544–559. <http://doi.org/10.1111/j.1475-5661.2012.00543.x>
- Environmental Protection Agency. (2018). Intelligent Dasymetric Mapping (IDM) Toolbox. Retrieved from <https://www.epa.gov/enviroatlas/dasymetric-toolbox>
- Ervin, D. (2012). Introduction to the Modifiable Areal Unit Problem. Retrieved December 12, 2018, from <http://gispopsci.org/maup/>
- Falcone, D. N., Wells, L. E., & Weisheit, R. A. (2002). The small-town police department. *Policing: An International Journal of Police Strategies & Management*, 25(2), 371–384.
- Federal Bureau of Investigation. (n.d.). National Incident-Based Reporting System. Retrieved November 17, 2018, from <https://www.fbi.gov/services/cjis/ucr/nibrs>
- Federal Bureau of Investigation. (2014). Uniform Crime Reporting Statistics: Their Proper Use. Retrieved August 1, 2015, from <https://www.fbi.gov/about-us/cjis/ucr/ucr-statistics-their-proper-use>
- Federal Bureau of Investigation. (2016). Message from the Director. Retrieved November 12, 2018, from [https://ucr.fbi.gov/nibrs/2016/resource-pages/message-from-director\\_final.pdf](https://ucr.fbi.gov/nibrs/2016/resource-pages/message-from-director_final.pdf)
- Federal Bureau of Investigation. (2018). *National Incident Based Reporting System, 2017*. Washington, D.C. Retrieved from <https://ucr.fbi.gov/nibrs/2017>
- Friendly, M. (2007). A.-M. Guerry's Moral Statistics of France : Challenges for Multivariable Spatial Analysis. *Statistical Science*, 22(3), 368–399. <http://doi.org/10.1214/07-STS241>
- Fryer Jr., R. G. (2016). An empirical analysis of racial differences in police use of force. *Working Paper*, 1–63. <http://doi.org/10.3386/w22399>
- Fyfe, N. R. (1991). The police, space and society : the geography of policing. *Progress in Human Geography*, 15(3), 249–267. <http://doi.org/10.1177/030913259101500301>
- Gatewood, G. (2001). A monograph on confidentiality and privacy in the US Census. *US Bureau of the Census, Washington*, (July). Retrieved from

<https://www.census.gov/history/pdf/ConfidentialityMonograph.pdf>

- Gibson, C. L., Walker, S., Jennings, W. G., & Miller, J. M. (2010). The impact of traffic stops on calling the police for help. *Criminal Justice Policy Review*, 21(2), 139–159.  
<http://doi.org/10.1177/0887403409344165>
- Goodchild, M. F. (2013). The quality of big (geo)data. *Dialogues in Human Geography*.  
<http://doi.org/10.1177/2043820613513392>
- Goodchild, M. F. (2018a). A GIScience Perspective on the Uncertainty of Context. *Annals of the American Association of Geographers*, 108(6), 1476–1481. <http://doi.org/10.1080/24694452.2017.1416281>
- Goodchild, M. F. (2018b). Big Geodata. In B. Huang (Ed.), *Comprehensive Geographic Information Systems* (pp. 19–25). Elsevier.
- Goodchild, M. F., & Haining, R. P. (2004). GIS and spatial data analysis: converging perspectives. *Papers in Regional Science*, 83, 363–385.
- Gove, W. R., Hughes, M., & Geerken, M. (1985). Are uniform crime reports a valid indicator of the index crimes? An affirmative answer with minor qualifications. *Criminology*, 23(3), 451–502.
- Gregory, D. E., & Janosik, S. M. (Steven M. (2003). The Effect of the Clery Act on Campus Judicial Practices. *Journal of College Student Development*, 44(6), 763–778.  
<http://doi.org/10.1353/csd.2003.0068>
- Groff, E. R. (2018). Do we really need collective social process to understand why crime occurs and offenders commit crime? In G. J. N. Bruinsma & S. D. Johnson (Eds.), *The Oxford Handbook of Environmental Criminology* (pp. 105–118). New York: Oxford University Press.
- Groff, E. R., Weisburd, D., & Yang, S.-M. (2010). Is it Important to Examine Crime Trends at a Local “Micro” Level?: A Longitudinal Analysis of Street to Street Variability in Crime Trajectories. *Journal of Quantitative Criminology*, 26(1), 7–32. <http://doi.org/10.1007/s10940-009-9081-y>
- Harries, K. D. (1975). Rejoinder to Richard Peet: “The Geography of Crime: A Political Critique.” *The Professional Geographer*, 27(3), 280–282. <http://doi.org/10.1177/0192513X12437708>
- Hart, T. C., & Rennison, C. (2003). *Reporting crime to the police, 1992-2000*. Washington, DC.
- Heywood, I., Cornelius, S., & Carver, S. (2002). *An introduction to geographical information systems*. New York: Addison Wesley Longman.
- Hickman, M. J., & Rice, S. K. (2010). Digital Analysis of Crime Statistics: Does Crime Conform to Benford’s Law? *Journal of Quantitative Criminology*, 26(3), 333–349. <http://doi.org/10.1007/s10940-010-9094-6>
- Hollis-Peel, M. E., Reynald, D. M., van Bavel, M., Elffers, H., & Welsh, B. C. (2011). Guardianship for crime prevention: A critical review of the literature. *Crime, Law and Social Change*, 56(1), 53–70.  
<http://doi.org/10.1007/s10611-011-9309-2>
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practics* (2nd editio). Melbourne, Australia: OTexts. Retrieved from <https://otexts.com/fpp2>
- Institute for Digital Research and Education, U. of C. L. A. (n.d.). FAQ: What is the Coefficient of Variation? Retrieved January 1, 2016, from

- [http://www.ats.ucla.edu/stat/mult\\_pkg/faq/general/coefficient\\_of\\_variation.htm](http://www.ats.ucla.edu/stat/mult_pkg/faq/general/coefficient_of_variation.htm)
- Jang, W., & Yao, X. (2011). Interpolating Spatial Interaction Data. *Transactions in GIS*, 15(4), 541–555. <http://doi.org/10.1111/j.1467-9671.2011.01273.x>
- Janosik, S. M., & Gregory, D. E. (2003). The Clery Act and Its Influence on Campus Law Enforcement Practices. *Journal of Student Affairs Research and Practice*, 41(1), 180–198. <http://doi.org/10.2202/1949-6605.1311>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29(4), 258–268. <http://doi.org/10.1080/10580530.2012.716740>
- Junior, D. P., & Muniz, J. (2005). “Stop or I’ll Call The Police!”: The Idea of Police, or the Effects of Police Encounters Over Time. *British Journal of Criminology*, 46(2), 234–257. <http://doi.org/10.1093/bjc/azi072>
- Kelling, B. G. L., & Moore, M. H. (1988). *The Evolving Strategy of Policing*. Washington, DC.
- Kim, H., & Yao, X. (2010). Pycnophylactic interpolation revisited: integration with the dasymetric-mapping method. *International Journal of Remote Sensing*, 31(21), 5657–5671. <http://doi.org/10.1080/01431161.2010.496805>
- Kim, Y. A., & Hipp, J. R. (2018). Physical Boundaries and City Boundaries: Consequences for Crime Patterns on Street Segments? *Crime and Delinquency*, 64(2), 227–254. <http://doi.org/10.1177/0011128716687756>
- King, J. B., Mahaffie, L., & McLarnon, G. (2016). The Handbook for Campus Safety and Security Reporting, 1–265. Retrieved from <https://www2.ed.gov/admins/lead/safety/handbook.pdf>
- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 205395171452848. <http://doi.org/10.1177/2053951714528481>
- Klinger, D. A. (1997). Negotiating order in patrol work: An ecological theory of police response to deviance. *Criminology*, 35(2), 277–306. <http://doi.org/10.1111/j.1745-9125.1997.tb00877.x>
- Klinger, D. A., & Bridges, G. S. (1997). Measurement Error in Calls-for-Service as an Indicator of Crime. *Criminology*, 35, 705–726. <http://doi.org/10.3868/s050-004-015-0003-8>
- Klockars, C. B. (1985). *The Idea of Police*. Newbury Park, CA: SAGE Publications.
- Knoxville Police Department. (2005). Standard Operating Procedure for Crime Analysis Unit. Retrieved February 1, 2019, from [www.iaca.net](http://www.iaca.net)
- Knoxville Police Department. (2014). Knoxville Police Department 2014 Annual Report. Retrieved February 1, 2019, from [http://knoxvilletn.gov/UserFiles/Servers/Server\\_109478/File/Police/kpd\\_annualreport2014.pdf](http://knoxvilletn.gov/UserFiles/Servers/Server_109478/File/Police/kpd_annualreport2014.pdf)
- Kwan, M.-P. (2012a). How GIS can help address the uncertain geographic context problem in social science research. *Annals of GIS*, 18(4), 245–255. <http://doi.org/10.1080/19475683.2012.727867>
- Kwan, M.-P. (2012b). How GIS can help address the uncertain geographic context problem in social science research. *Annals of GIS*, 18(4), 245–255. <http://doi.org/10.1177/0192513X12437708>
- Kwan, M.-P. (2012c). The Uncertain Geographic Context Problem. *Annals of the Association of American*

- Geographers*, 64(4), 37–41. Retrieved from <http://www.tandfonline.com/doi/abs/10.1111/1467-8306.00084>
- Kwan, M.-P. (2016). Algorithmic Geographies: Big Data, Algorithmic Uncertainty, and the Production of Geographic Knowledge. *Annals of the American Association of Geographers* ISSN:, 0(February 2015), 1–9. <http://doi.org/10.1080/00045608.2015.1117937>
- Kwan, M.-P. (2018). The Limits of the Neighborhood Effect: Contextual Uncertainties in Geographic, Environmental Health, and Social Science Research. *Annals of the American Association of Geographers*, 108(6), 1482–1490. <http://doi.org/10.1080/24694452.2018.1453777>
- Kwan, M.-P., & Schwanen, T. (2018). Context and Uncertainty in Geography and GIScience: Advances in Theory, Method, and Practice. *Annals of the Association of American Geographers*, 108(6), 1473–1475. <http://doi.org/10.1080/24694452.2018.1457910>
- LaRue, E., & Andresen, M. a. (2015). Spatial Patterns of Crime in Ottawa: The Role of Universities. *Canadian Journal of Criminology and Criminal Justice*, 57(2), 189–214. <http://doi.org/10.3138/CJCCJ.2013.E47>
- Lauritsen, J. L., & Cork, D. L. (Eds.). (2016). *Modernizing Crime Statistics*. Washington, D.C.: National Academies Press. <http://doi.org/10.17226/23492>
- LeBeau, J. L., & Leitner, M. (2011). Introduction: Progress in Research on the Geography of Crime. *The Professional Geographer*, 63(2), 161–173.
- Leigh Star, S. (2010). This is Not a Boundary Object: Reflections on the Origin of a Concept. *Science, Technology & Human Values*, 35(5), 601–617. <http://doi.org/10.1177/0162243910377624>
- Leslie, T. F., & Kronenfeld, B. J. (2011). The Co-Location Quotient: A New Measurement of Spatial Correlation between Categories. *Geographical Analysis*, 43(3), 306–326.
- Leszczynski, A. (2012). Situating the geoweb in political economy. *Progress in Human Geography*. <http://doi.org/10.1177/0309132511411231>
- Levitt, S. D. (1998). The Relationship Between Crime Reporting and Police: Implications for the Use of Uniform Crime Reports. *Journal of Quantitative Criminology*, 14(1), 61–82. <http://doi.org/10.1023/A:1023096425367>
- Lott, J. R., & Whitley, J. (2003). Measurement Error in County-Level UCR Data, 19(2), 185–198.
- Lowman, J. (1986). Conceptual Issues in the Geography of Crime : Toward a Geography of Social Control. *Annals of the Association of American Geographers*, 76(1), 81–94. Retrieved from <http://www.jstor.org/stable/2562874>
- Lynch, J. P. (2018). Not Even Our Own Facts: Criminology in the Era of Big Data. *Criminology*, 56(3), 437–454. <http://doi.org/10.1111/1745-9125.12182>
- Lynch, J. P., & Jarvis, J. P. (2008). Missing Data and Imputation in the Uniform Crime Reports and the Effects on National Estimates. *Journal of Contemporary Criminal Justice*, 24(1), 69–85. <http://doi.org/10.1177/1043986207313028>
- Maguire, E. R., & King, W. R. (2004). Trends in the Policing Industry. *The Annals of the American Academy of Political and Social Science*, 593(1), 15–41. <http://doi.org/10.1177/0002716204262960>

- Maltz, M. D. (1977). Crime Statistics: A Historical Perspective. *Crime & Delinquency*, 23(1), 32–40. <http://doi.org/10.1177/001112877702300103>
- Maltz, M. D. (2006). *Analysis of Missingness in UCR Crime Data*. Columbus, OH.
- Maltz, M. D. (2009). Waves, Particles, and Crime. In D. Weisburd, W. Bernasco, & G. J. N. Bruinsma (Eds.), *Putting Crime in its Place* (pp. 123–142). New York, NY: Springer New York. <http://doi.org/10.1007/978-0-387-09688-9>
- Maltz, M. D., & Targonski, J. (2002). A Note on the Use of County-Level UCR Data. *Journal of Quantitative Criminology*, 18(3), 297–318. <http://doi.org/10.1023/A:1016060020848>
- Maltz, M. D., & Targonski, J. (2003). Measurement and Other Errors in County-Level UCR Data : A Reply to Lott and Whitley, 19(2), 199–206.
- Maltz, M. D., & Targonski, J. (2004). *Making UCR Crime Data Useful and Accessible*. National Criminal Justice Reference Service. Chicago.
- Manning, P. K. (2010). *Policing contingencies*. University of Chicago Press.
- Marvell, T. B., & Moody, C. E. (1996). Specification Problems, Police Levels, and Crime Rates. *Criminology*, 34(4), 609–646.
- Mayernik, M. S. (2017). Open data : Accountability and transparency. *Big Data & Society*, (July-December), 1–5. <http://doi.org/10.1177/2053951717718853>
- Mazerolle, L., Bennett, S., Davis, J., Sargeant, E., & Manning, M. (2013). Procedural justice and police legitimacy: A systematic review of the research evidence. *Journal of Experimental Criminology*, 9(3). <http://doi.org/10.1007/s11292-013-9175-2>
- McCarthy, C. (2016). Avoid common Clery Act compliance pitfalls, challenges. *Campus Legal Advisor*, 15(9), 1–3. <http://doi.org/10.1002/cala>
- McFarlane, C., & Anderson, B. (2011). Thinking with assemblage. *Area*, 43(2), 162–164. <http://doi.org/10.1111/j.1475-4762.2011.01012.x>
- Mennis, J., & Hultgren, T. (2006). Intelligent Dasyetric Mapping and Its Application to Areal Interpolation. *Cartography and Geographic Information Science*, 33(3), 179–194. <http://doi.org/10.1559/152304006779077309>
- Merschdorf, H., & Blaschke, T. (2018). Revisiting the Role of Place in Geographic Information Science. *ISPRS International Journal of Geo-Information*, 7(9), 364. <http://doi.org/10.3390/ijgi7090364>
- Messner, S. F., Beaulieu, M., Isles, S. N., & Mitchell, L. (2014). Revisiting the Quality and Use of Race-Specific Homicide Data: Exploring Substantive Implications. *Homicide Studies*, 18, 151. <http://doi.org/10.1177/1088767905277144>
- Miller, H. J. (2017). Geographic information science I: Geographic information observatories and opportunistic GIScience. *Progress in Human Geography*, 41(4), 489–500. <http://doi.org/10.1177/0309132517710741>
- Miller, H. J. (2018). Geographic information science II: Mesogeography: Social physics, GIScience and the quest for geographic knowledge. *Progress in Human Geography*, 42(4), 600–609. <http://doi.org/10.1177/0309132517712154>

- Miller, H. J., & Goodchild, M. F. (2014). Data-driven geography. *GeoJournal*, 80(4), 449–461. <http://doi.org/10.1007/s10708-014-9602-6>
- Molloy, J. C. (2011). The open knowledge foundation: Open data means better science. *PLoS Biology*, 9(12), 1–4. <http://doi.org/10.1371/journal.pbio.1001195>
- Monthly Report of Offenses Known by Law Enforcement (Return A) Master File. (1960). Washington, DC: United States Justice Department, Federal Bureau of Investigation.
- Monthly Report of Offenses Known by Law Enforcement (Return A) Master File. (2012). Washington, DC: United States Justice Department, Federal Bureau of Investigation.
- Moore, Mark H. and Kelling, G. L. (1983). To Serve and Protect - Learning from Police History. *Public Interest*, 70, 49–65.
- Nagle, N. N., Battenfield, B. P., Leyk, S., & Spielman, S. (2014). Dasymetric Modeling and Uncertainty. *Annals of the Association of American Geographers*, 104(1), 80–95. <http://doi.org/10.1080/00045608.2013.843439>
- Nesbary, D. K. (1998). Handling emergency calls for service: organizational production of crime statistics. *Policing: An International Journal of Police Strategies & Management*, 21(4), 576–599. <http://doi.org/10.1108/13639519810241638>
- Nolan, J. J., Haas, S. M., & Napier, J. S. (2011). Estimating the Impact of Classification Error on the “Statistical Accuracy” of Uniform Crime Reports. *Journal of Quantitative Criminology*, 27(4), 497–519. <http://doi.org/10.1007/s10940-011-9135-9>
- O’Connor Shelley, T., & Crow, M. S. (2009). The nature and extent of conservation policing: Law enforcement generalists or conservation specialists? *American Journal of Criminal Justice*, 34(1–2), 9–27. <http://doi.org/10.1007/s12103-008-9057-8>
- O’Sullivan, D., & Unwin, D. (2010). *Geographic information analysis*. John Wiley & Sons.
- Openshaw, S. (1984). The modifiable areal unit problem. *Concepts and Techniques in Modern Geography*, 38, 41–53.
- Openshaw, S., & Taylor, P. J. (1979). A Million or So Correlation Coefficients: Three Experiments on the Modifiable Areal Unit Problem. *Statistical Applications in the Spatial Sciences*, 21, 127–144.
- Paoline III, E. A., & Sloan III, J. J. (2003). Variability in the organizational structure of contemporary campus law enforcement agencies: A national-level analysis. *Policing: An International Journal of Police Strategies & Management*, 26(4), 612–639. <http://doi.org/10.1108/13639510310503541>
- Park, R. E., Burgess, E. W., McKenzie, R. D., & Wirth, L. (1925). *The City*. University of Chicago Press. Retrieved from <https://books.google.com/books?id=cOnH6O3CvXgC>
- Peet, R. (1975). The Geography Of Crime: A Political Critique. *The Professional Geographer*, 27(3), 277–280. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.0033-0124.1975.00277.x/abstract>
- Pelfrey, W. V. J. (2007). Style of Policing Adopted by Rural Police and Deputies. *Policing: An International Journal of Police Strategies & Management*, 30(4), 620–636. <http://doi.org/10.1525/sp.2007.54.1.23>.

- Pepper, J., Petrie, C., & Sullivan, S. (2010). Measurement error in criminal justice data. In A. R. Piquero & D. Weisburd (Eds.), *Handbook of quantitative criminology* (pp. 353–374). Springer.  
<http://doi.org/10.1007/978-0-387-77650-7>
- Phillips, S. W., & Sobol, J. J. (2012). Police decision making: An examination of conflicting theories. *Policing, 35*(3), 551–565. <http://doi.org/10.1108/13639511211250794>
- Police Data Initiative. (n.d.). Retrieved February 1, 2019, from [www.policedatainitiative.org](http://www.policedatainitiative.org)
- Quenouille, M. H. (1949). Approximate Use of Correlation in Time Series. *Journal of the Royal Statistical Society, Series B, 11*, 18–84.
- Quenouille, M. H. (1956). Notes on Bias Estimation. *Biometrika, 43*, 353–360.
- Ratcliffe, J. H. (2008). Intelligence-led Policing. In R. Wortley & L. Mazerolle (Eds.), *Environmental criminology and crime analysis*. Willan.
- Reichman, O J ; Jones, M.B.; Schildhauer, M. P. (2011). Challenges and Opportunités of Open Data in Ecology. *Science, 331*(February), 703–705. <http://doi.org/10.1126/science.1197962>
- Renauer, B. C. (2012). Neighborhood Variation in Police Stops and Searches: A Test of Consensus and Conflict Perspectives. *Police Quarterly, 15*(3), 219–240. <http://doi.org/10.1177/1098611112447746>
- Rengert, G. F., & Lockwood, B. (2009). Geographical Units of Analysis and the Analysis of Crime. In D. Weisburd, W. Bernasco, & G. J. N. Bruinsma (Eds.), *Putting Crime in its Place* (pp. 109–122). New York, NY: Springer New York. <http://doi.org/10.1007/978-0-387-09688-9>
- Rosenfeld, R. (2007). Explaining the divergence between UCR and NCVS aggravated assault trends. *Understanding Crime Statistics: Revisiting the Divergence of the NCVS and the UCR*, 251–268.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581–592.
- Sampson, R. J. (2013). The Place of Context: a Theory and Strategy for Criminology’s Hard Problems. *Criminology, 51*(1), 1–31. <http://doi.org/10.1111/1745-9125.12002>
- Sampson, R. J., & Wilson, W. J. (2012). *Great American City: Chicago and the Enduring Neighborhood Effect*. University of Chicago Press. Retrieved from <https://books.google.com/books?id=eEHpu8iMq7UC>
- Schnell, C., Braga, A. A., & Piza, E. L. (2017). The Influence of Community Areas, Neighborhood Clusters, and Street Segments on the Spatial Variability of Violent Crime in Chicago. *Journal of Quantitative Criminology, 33*(3), 469–496. <http://doi.org/10.1007/s10940-016-9313-x>
- Schnell, C., Grossman, L., & Braga, A. A. (2018). The routine activities of violent crime places: A retrospective case-control study of crime opportunities on street segments. *Journal of Criminal Justice, 47*(1), 1–15. <http://doi.org/10.1016/j.jcrimjus.2018.10.002>
- Shaw, C., & McKay, H. (1942). *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.
- Shaw, C., & McKay, H. (1969). *Juvenile Delinquency and Urban Areas* (rev.). Chicago: University of Chicago Press.
- Skogan, W. G. (2009). Concern About Crime and Confidence in the Police: Reassurance or Accountability? *Police Quarterly, 12*(3), 301–318. <http://doi.org/10.1177/1098611109339893>



- Sleeter, R., & Gould, M. (2007). *Geographic Information System Software to Remodel Population Data Using Dasymetric Mapping Methods*. Reston, VA.
- Song, J., Andresen, M. A., Brantingham, P. L., & Spicer, V. (2017). Crime on the edges: patterns of crime and land use change. *Cartography and Geographic Information Science*, *44*(1), 51–61. <http://doi.org/10.1080/15230406.2015.1089188>
- State University of New York. (n.d.). 2011-2016-Compare-of-Handbook-for-Campus-Safety-Security-Reporting. Retrieved January 19, 2019, from <https://system.suny.edu/counsel/2016-Clery-Changes/>
- Steenbeek, W., & Weisburd, D. (2016). Where the Action is in Crime? An Examination of Variability of Crime Across Different Spatial Units in The Hague, 2001–2009. *Journal of Quantitative Criminology*, *32*(3), 449–469. <http://doi.org/10.1007/s10940-015-9276-3>
- Storch, J. (n.d.). The 2011 Clery Handbook: New Developments and Important Changes. Retrieved February 5, 2017, from <http://counsel.cua.edu/fedlaw/nacuanotescleryact.cfm>
- Tabarrok, A., Heaton, P., & Helland, E. (2010). The measure of vice and sin : a review of the uses , limitations and implications of crime data. In and P. R. Z. Benson, Bruce L. (Ed.), *Handbook on the economics of crime* (pp. 53–81). Edward Elgar Publishing.
- Telep, C. W., Mitchell, R. J., & Weisburd, D. (2014). How Much Time Should the Police Spend at Crime Hot Spots? Answers from a Police Agency Directed Randomized Field Trial in Sacramento, California. *Justice Quarterly*, *31*(5), 905–933. <http://doi.org/10.1080/07418825.2012.710645>
- Telep, C. W., Weisburd, D., Gill, C. E., Vitter, Z., & Teichman, D. (2014). Displacement of crime and diffusion of crime control benefits in large-scale geographic areas: a systematic review. *Journal of Experimental Criminology*, *10*(4), 515–548. <http://doi.org/10.1007/s11292-014-9208-5>
- Times Herald. (n.d.). American Community Survey. Retrieved March 1, 2019, from <https://data.thetimesherald.com/american-community-survey/census-tract-68-knox-county-tennessee/poverty-status/total-population/yty/14000US47093006800/>
- Tita, G. E., & Greenbaum, R. T. (2009). Crime, Neighborhoods, and Units of Analysis: Putting Space in Its Place. In D. Weisburd, W. Bernasco, & G. J. N. Bruinsma (Eds.), *Putting Crime in its Place* (pp. 145–170). New York, NY: Springer New York. <http://doi.org/10.1007/978-0-387-09688-9>
- Tobler, W. R. (1979). Smooth Pycnophylactic Interpolation for Geographical Regions. *Journal of the American Statistical Association*, *367*(367), 519–536.
- Tukey, J. (1958). Bias and Confidence in not Quite Large Samples (abstract). *The Annals of Mathematical Statistics*, *29*, 614.
- Tyler, T. R. (2004). Enhancing Police Legitimacy. *The Annals of the American Academy of Political and Social Science*, *593*(1), 84–99. <http://doi.org/10.1177/0002716203262627>
- UCAD Project Committee. (2012). *Unified CAD Functional Requirements*.
- US Department of Justice. (2012). *Crime in the United States*. Washington, DC. Retrieved from [www.fbi.gov](http://www.fbi.gov)
- US Department of Justice. (2013a). National Incident-Based Reporting System (NIBRS) User Manual. Retrieved February 1, 2019, from [www.fbi.gov](http://www.fbi.gov)

- US Department of Justice. (2013b). Summary Reporting System (SRS) User Manual. Retrieved February 1, 2019, from [www.fbi.gov](http://www.fbi.gov)
- Varano, S., Schafer, J., Cancino, J., & Swatt, M. (2009). Constructing crime: Neighborhood characteristics and police recording behavior. *Journal of Criminal Justice*, 37(6), 553–563. <http://doi.org/10.1016/j.jcrimjus.2009.09.004>
- Vila, B., & Morris, C. (1999). *The Role of Police in American Society: A Documentary History*. (B. Vila & C. Morris, Eds.). Westport, CT: Greenwood Press.
- Walsh, W. F. (2001). Compstat: an analysis of an emerging police managerial paradigm. *Policing: An International Journal of Police Strategies & Management*, 24(3), 347–362. <http://doi.org/10.1108/13639510110401717>
- Walsh, W. F., & Vito, G. F. (2004). The Meaning of Compstat: Analysis and Response. *Journal of Contemporary Criminal Justice*, 20(1), 51–69. <http://doi.org/10.1177/1043986203262310>
- Wang, F., Hu, Y., Wang, S., & Li, X. (2017). Local Indicator of Colocation Quotient with a Statistical Significance Test: Examining Spatial Association of Crime and Facilities. *Professional Geographer*, 69(1), 22–31. <http://doi.org/10.1080/00330124.2016.1157498>
- Weisburd, B. D., Mastrofski, S. D., Greenspan, R., & Willis, J. J. (2004). The Growth of Compstat in American Policing. Retrieved February 1, 2019, from [https://www.policefoundation.org/wp-content/uploads/2015/06/Weisburd-et-al.-2004-The-Growth-of-Compstat-in-American-Policing-Police-Foundation-Report\\_0.pdf](https://www.policefoundation.org/wp-content/uploads/2015/06/Weisburd-et-al.-2004-The-Growth-of-Compstat-in-American-Policing-Police-Foundation-Report_0.pdf)
- Weisburd, D. (2015a). The law of crime concentration and the criminology of place. *Criminology*, 53(2), 133–157. <http://doi.org/10.1111/1745-9125.12070>
- Weisburd, D. (2015b). The law of crime concentration and the criminology of place. *Criminology*, 53(2). <http://doi.org/10.1111/1745-9125.12070>
- Weisburd, D., & Amram, S. (2014). The law of concentrations of crime at place: the case of Tel Aviv-Jaffa. *Police Practice and Research*, 15(2), 101–114. <http://doi.org/10.1080/15614263.2013.874169>
- Weisburd, D., Groff, E. R., & Yang, S. M. (2013). *The Criminology of Place: Street Segments and Our Understanding of the Crime Problem*. *The Criminology of Place: Street Segments and Our Understanding of the Crime Problem*. Oxford University Press. <http://doi.org/10.1093/acprof:oso/9780195369083.001.0001>
- Weisburd, D., Groff, E. R., & Yang, S. M. (2014). The Importance of Both Opportunity and Social Disorganization Theory in a Future Research Agenda to Advance Criminological Theory and Crime Prevention at Places. *Journal of Research in Crime and Delinquency*, 51(4). <http://doi.org/10.1177/0022427814530404>
- Weisburd, D., Mastrofski, S. D., McNally, A. M., Greenspan, R., & Willis, J. J. (2003). Reforming to Preserve: COMPSTAT and Strategic Problem Solving in American Policing. *Criminology and Public Policy*, 2(3), 421–456.
- Weiss, T. (2011). The blurring border between the police and the military: A debate without foundations. *Cooperation and Conflict*, 46(3), 396–405. <http://doi.org/10.1177/0010836711416961>
- Wender, B. A. (Ed.). (2017). *Refining the Concept of Scientific Inference When Working with Big Data*.

- Washington, D.C.: National Academies Press. <http://doi.org/10.17226/24654>
- Wheeler, A. P., Worden, R. E., & McLean, S. J. (2016). Replicating Group-Based Trajectory Models of Crime at Micro-Places in Albany, NY. *Journal of Quantitative Criminology*, 32(4), 589–612. <http://doi.org/10.1007/s10940-015-9268-3>
- Williams, B. H., Murphy, P. V, Zinn, H., Stewart, J. K., & Moore, M. H. (1990). *The Evolving Strategy of Police: A Minority View*. Washington, DC.
- Wilson, M. W. (2011a). Data matter(s): legitimacy, coding, and qualifications-of-life. *Environment and Planning D: Society and Space*, 29(5), 857–872. <http://doi.org/10.1068/d7910>
- Wilson, M. W. (2011b). “Training the eye”: Formation of the geocoding subject. *Social and Cultural Geography*. <http://doi.org/10.1080/14649365.2010.521856>
- Wilson, M. W. (2017). *New lines: Critical GIS and the trouble of the map*. U of Minnesota Press.
- Wong, D. (2009). The Modifiable Areal Unit Problem (MAUP). In A. S. Fotheringham & P. A. Rogerson (Eds.), *The Sage Handbook of Spatial Analysis* (pp. 105–121). SAGE Publications.
- Zandbergen, P. A., & Ignizio, D. A. (2010). Comparison of Dasymetric Mapping Techniques for Small-Area Population Estimates. *Cartography and Geographic Information Science*, 37(3), 199–214. <http://doi.org/10.1559/152304010792194985>
- Zhang, G., & Zhu, A.-X. (2018). The representativeness and spatial bias of volunteered geographic information: a review. *Annals of GIS*, 24(3), 151–162. <http://doi.org/10.1080/19475683.2018.1501607>
- Zhao, J. S., & Hassell, K. D. (2005). Policing Styles and Organizational Priorities: Retesting Wilson’s Theory of Local Political Culture. *Police Quarterly*, 8(4), 411–430. <http://doi.org/10.1177/1098611104269078>