Despite an abundance of research on the subject, it is still difficult to predict where and when crime will happen. The amount of effort already thrown at this issue is indicative of its importance to number of different stakeholders. First and foremost, police departments are looking for the best guidance possible on how to allocate their limited resources. Public officials who are in charge of scheduling festivals, conferences, or demonstrations have a vested interest in making those events as crime-free as possible.

Business owners, real estate professionals, and developers also have significant amounts of money on the line when they decide on an area in which to invest. Being able to predict areas that are at a high risk for crime ensures that investment is well placed. Additionally, school systems and social services departments have a serious stake in implementing policies which will lead to a decline in future crime. Knowing what factors correlate predictively to crime would give them valuable guidance in that mission.

As mentioned, there have been many studies attempting to address this problem looking at a wide range of variables. In our literature review, we have taken a sample of these studies and broken them down into three broad categories: Who, When, and Where. To elaborate, studies that attempt to predict crime based on characteristics of the people involved in the crime; studies focusing on when the crime happened and what else was going on at that time; and studies concerned with attributes of the geographic location of the crime. Below we summarize some of the most recent and relevant research in these three approaches and identify some potential ways forward in this vital area of study.

In addition to looking at who offenders are and when crimes take place, crime research must also examine where crimes occur. Studies of geographical locations as they relate to crime events began as early as the 1800’s, but police departments and researchers did not have a sustained focus on crime mapping until the late 1960’s when computers could be engaged in automating the process (Weisburd and McEwen 8). The relatively recent advent of sophisticated modeling computing tools such as ArcGIS maximizes the opportunity to map crime to location, and as might be expected, there has been a shift in recent years from examination of the larger environment, such as the neighborhood, to a narrowed focus on crime events at places within that neighborhood, such as intersections, addresses, or individual businesses (Deryol et al. 306; Han et al. 1111). For example, one recent study featured an analysis of impact on crime rates within 250 feet of individual foreclosed residential properties; the authors discovered a 19% increase in violent crimes for foreclosed properties that became vacant (Cui and Walsh 84). Although a number of different interesting analyses have been performed for crime frequency around different types of specific locations, such as liquor stores and drug treatment centers (Han et al. 1118-1119; Furr-Holden et al. 23), there are still many types of business properties and subsets of properties to be examined to enhance our ability to use location to predict criminal activity. We also must seek a better understanding of the geographical context for crimes at specific locations; Deryol, et al. note the importance of other nearby destinations (nodes), travel routes (paths), and the broader environmental landscape in close proximity to crime hotspots (locations of high criminal activity) to provide contextual information to crime events (306).

Many factors contribute to whether or not a crime occurs at a specific location. Eck and Weisburd note the importance of controllers, or those whose presence and effectiveness can prevent a crime event (5-6). These controllers include: intimate handlers, or those who have influence over the offender, such as a parent or a teacher; place managers, responsible for oversight of a location, such as a building superintendent; and guardians, or those who protect the criminal targets, such as a friend or police officer. As noted above, the greater geographical context for a location is also critical in influencing a crime event; for example, if one location is less proximal to a bus stop than another, it may be more difficult to reach by criminal offenders as well as targets and thus characterized by a lower frequency of crime. Critical infrastructure such as street lighting and public transportation and the socioeconomic status of the location’s neighborhood will also be factors to consider when examining the likelihood of criminal activity at a particular location.

As we stated previously, many cities make data on criminal activity publicly available, and these typically include the category of offence, at least some type of location information, and the date of the event, as well as other markers. Many of these datasets are available at Data.gov. To protect the identity of victims, the exact address of the crime event is frequently masked and only the block-level information is provided. There do appear to be some cities, such as Austin, Texas, that provide exact addresses, and these will be more useful for examination of crime locations. Also of note, many cities use the Uniform Crime Report (UCR) format, and one of the issues with this data is that if a single event involves multiple crimes, only the most serious crime will be listed. For example, if someone breaks into a house and murders an occupant, only the murder will be indicated (Cui and Walsh 74); this means that less serious crimes will be underreported in the dataset.

Our desired alternative related to crime places is to improve the ability to predict whether or not a crime will occur at a specific location, especially for crime hotspots. This will in turn be useful for the ultimate goal of reducing crime both at those locations and overall. Also, with a better understanding of what specific locations have higher frequency of crime, police departments and city leaders can potentially partner with business owners to reduce crime at those locations, including offering incentives or punitive measures as appropriate to effect the desired change. While many business owners will likely offer resistance to being held responsible for crimes that occur on their property, a recent article at Bloomberg.com highlighting the problem of crime incidents at Wal-Marts across the country noted the efforts of mayor Dennis Buckley of Beech Grove, Indiana, who successfully used social and television media to put pressure on his local Wal-Mart to better partner with the city in managing criminal activity at their store (Pettypiece and Voreacos).

In order to transition to our desired alternative situation and support the above approach, our objective is to have a better understanding of how to accurately and fairly identify crime hotspots. Merely identifying locations with more crimes will not be enough if there is to be a potentially punitive element in marking a location as a hotspot; we must ensure that other factors are taken into consideration. For example, comparing the number of crimes that occur at a Wal-Mart with those that occur at a Target would perhaps not be an appropriate comparison due to the potential differences in customer base, size of store, location, and access to store. Once an appropriate measure has been identified to determine hotspots that have crime events above the norm and that data has been shared with police departments, a metric to evaluate progress would be to examine month-to-month changes in crime reports for targeted hotspot areas where interventions have been implemented. If crime at those locations grows or is stagnant, follow-up actions with business owners could be planned.

A representative sample of data sources, data mining methods, and some of the evaluation methods and their outcomes for studies on crime places is summarized in the table below. Given the previous work on crime places, the problem remains unsolved because of the multitude of factors influencing whether a crime occurs at a particular location or not. Although we are better able to narrow our focus to the microenvironment with today’s technology and statistical methods, a crime event is characterized by a confluence of factors and our models must attempt to represent this accurately. In addition, knowing whether a crime is more or less likely to occur at a particular location does not necessarily translate into effective mitigation of the risk.

|  |  |  |  |
| --- | --- | --- | --- |
| **Data sources** | | | |
| Cui and Walsh | Furr-Holden et al. | Deryol et al. | Han et al. |
| Crime data, Police Department of the City of Pittsburgh | Publicly funded outpatient drug treatment centers, Baltimore Substance Abuse Systems, Inc. (BSAS) | Commercial density – Cincinnati Area Geographic Information System’s open website | Liquor stores, Pennsylvania Liquor Control Board (PLCB) |
| **Data mining methods** | | | |
| Classification (identifying different periods associated with foreclosures: pre-foreclosure, foreclosure, vacancy, reoccupation) | Negative binomial regression modeling | Poisson-based hierarchical regression analysis | Difference-in-difference-in-differences (triple difference) modeling |
| **Evaluation methods and outcomes** | | | |
| Tested for potentially treated controls by defining larger control areas. For violent crime, the resulting reduction in magnitude indicated that the effective treatment area is smaller than 500 feet, thus validating the original approach. | Reran regression models only using drug treatment centers, convenience stores, liquor stores, and corner stores that had no other businesses of a similar type within the buffer range. The results were similar to the original findings, thus validating the approach. | Assessed whether there was a relative improvement across different explanatory models by re-running the analysis within a linear model specification. The decrease in deviance statistics were only significant when moving toward the model looking at the additive effects of the variables under consideration; the decrease was not significant for a model looking at a product of the variable values. | Used a falsification model, arbitrarily examining a different day (Tuesday) than the target day (Sunday). The falsification model did not demonstrate a statistically significant increase in crime for the false variable value, thus validating the original model. |

Works Cited

Cui, Lin and Walsh, Randall. “Foreclosure, Vacancy and Crime.” *Journal of Urban Economics* 87 (2015): 72-84. Print.

Deryol, Rustu, et al. “Crime Places in Context: An Illustration of the Multilevel Nature of Hot Spot Development.” *Journal of Quantitative Criminology* 32.2 (June 2016): 305-325. Print.

Eck, John E. and Weisburd, David. “Crime Places in Crime Theory.” *Crime and Place*. Cincinnati, OH: Criminal Justice Press, 1995. Print.

Furr-Holden, C. Debra M., et al. “Not in My Back Yard: A Comparative Analysis of Crime Around Publicly Funded Drug Treatment Centers, Liquor Stores, Convenience Stores, and Corner Stores in One Mid-Atlantic City.” *Journal of Studies on Alcohol and Drugs* 77.1 (2016): 17–24. Print.

Han, SeungHoon, et al. “The Effect of a Sunday Liquor-Sales Ban Repeal on Crime: A Triple-Difference Analysis.” *Alcoholism: Clinical and Experimental Research* 40.5 (May 2016): 1111-1121. Print.

Pettypiece, Shannon and Voreacos, David. *Walmart’s Out-of-Control Crime Problem Is Driving Police Crazy*. Bloomberg Businessweek, 17 Aug. 2016, <http://www.bloomberg.com/features/2016-walmart-crime/>. Accessed 27 August 2016.

Weisburd, David and McEwen, Tom, eds. *Crime Mapping and Crime Prevention*. Cincinnati, OH: Criminal Justice Press, 1997. Print.

Wolfe, Scott E. and Pyrooz, David C. “Rolling Back Prices and Raising Crime Rates? The Walmart Effect on Crime in the United States.” *British Journal of Criminology* 54 (2014): 199-221. Print.

Unfamiliar Terms

attenuation bias

clustered standard error

robust standard error

difference-in-differences research design

detrended

counterfactual

fixed effect

composition effects

triple-difference study

permutation test

Huber-White sandwich estimator of variance

autocorrelation

Euclidian distance

negative binomial regression models

conjunctive analysis

transformation using log base 10 function due to skewed distributions

grand-mean centering

principal component analysis (with Direct Oblimin Rotation)

Global Moran’s I statistics

queen contiguity

over-dispersion