THE SEARCH FOR THE BROKEN WINDOWS TIPPING POINT: A DOSE-RESPONSE PROPENSITY SCORE ASSESSMENT OF THE RELATIONSHIP BETWEEN DISORDER AND VIOLENT CRIME

By

Alaina De Biasi Podges

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Criminal Justice – Doctor of Philosophy

2020

ABSTRACT

THE SEARCH FOR THE BROKEN WINDOWS TIPPING POINT: A DOSE-RESPONSE PROPENSITY SCORE ASSESSMENT OF THE RELATIONSHIP BETWEEN DISORDER AND VIOLENT CRIME

By

Alaina De Biasi Podges

Wilson and Kelling (1982) provide a simple instruction for the implementation of order-maintenance policing: direct limited police resources to the *broken windows tipping point*. In doing so, they imply a certain functional form of the relationship between disorder and violent crime. That is, Wilson and Kelling's (1982) description of the tipping point suggests that the disorder-crime relationship is best captured as a threshold effect. If this is indeed the case, then a proper test of the validity of broken windows theory should accommodate nonlinearity. To this end, this study empirically examined the functional form of the relationship between physical disorder and violent crime rate in Detroit, Michigan utilizing a dose-response propensity score methodology. To facilitate its analysis, this study utilized block-group level data on physical disorder, violent crime, as well as socioeconomic and land use characteristics from the Detroit Police Department's record management system, Motor City Mapping project, and Census. Despite its comprehensive analysis, the functional form of the disorder-crime relationship remains unclear. That being said, the bulk of the evidence favors a nonlinear relationship, with partial support for Wilson and Kelling's (1982) interpretation of the broken windows tipping point. Several directions for future research are identified in an effort to spur the cultivation of this undeveloped avenue of research.

This dissertation would not have been possible without the support of James Podges, husband and editor extraordinaire.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Statement of the Problem

In a seminal article published in *The Atlantic Monthly*, James Q. Wilson and George L. Kelling (1982) first described broken windows theory (BWT) as a developmental sequence of events in which unattended minor issues produce deleterious consequences for neighborhoods. These issues encompass physical conditions (e.g., abandoned structures/lots, graffiti, trash, and overgrown vegetation) and social nuisances (e.g., panhandling, loitering, and public drinking) that signify neighborhood decline and inspire fear within residents, hereon referred to collectively as disorder. The perpetuation of violent crime - a controversial topic of study in broken windows research – is foremost among the consequences of unattended disorder. However, disorder has also been linked to high levels of fear (e.g., Taylor & Shumaker, 1990; Covington & Taylor, 1991; LaGrange, Ferraro, & Supancic, 1992; Ross & Jang, 2000; Markowitz et al., 2001; Spelman, 2004), low levels of informal social control and collective efficacy (e.g., Foster-Fishman et al., 2007; Kleinhans & Bolt, 2014; Wickes & Hipp, 2018; Wickes, Broidy, & Hipp, 2018), as well as poor mental and physical health (e.g., Cutrona et al., 2000; Ross & Mirowsky, 2001; Hill, Ross, & Angel, 2005). In light of these consequences, efforts to reduce disorder within neighborhoods are of great interest to residents, police, and policymakers, alike. That being said, Wilson and Kelling's (1982) simple solution to the problem of neighborhood disorder - stop small problems before they become much larger - has incited much debate, resulting in a voluminous and mixed body of research which questions the validity of BWT (e.g., Skogan, 1990; Sampson & Raudenbush, 1999; Kelling & Sousa, 2001; Braga et al., 2015; Weisburd et al., 2015).

While the mechanisms through which disorder affects violent crime are hotly debated, order-maintenance (or broken windows) policing is still a popular means of disrupting the broken windows cycle and restoring order within neighborhoods. In keeping with Wilson and Kelling's (1982) original conceptualization, order-maintenance policing addresses threatening behaviors and physical aspects of the environment that are determined by negotiated rules for street-level order realized through police-community partnerships. In the most comprehensive evaluation to-date, Braga, Welsh, and Schnell (2015) utilized meta-analytical techniques to evaluate 30 randomized experimental and quasi-experimental tests of policing strategies that addressed disorder (i.e., policing disorder strategies). Overall, they found these strategies to have a significant, yet modest effect on crime reduction (Braga et al., 2015). The programs that had the strongest impact were those that utilized community and problem-solving interventions, reflective of order-maintenance policing (Braga et al., 2015). Drawing from a subset of studies from this review, Weisburd et al. (2015, p. 591) later revealed that "there is little evidence that the model proposed in broken windows policing is driving such crime reductions." Rather, they argue that the mechanisms underlying criminal opportunity theories may better help explain its crime control gains (Weisburd et al., 2015).

Notwithstanding Braga et al.'s (2015) and Weisburd et al.'s (2015) findings, broken windows research has failed to heed a critical instruction provided by Wilson and Kelling (1982, para. 47): to identify neighborhoods *at the tipping point* ‘‘where the public order is deteriorating but not unreclaimable, where the streets are used frequently but by apprehensive people, where a window is likely to be broken at any time, and must quickly be fixed if all are not to be shattered." This failure is significant. Efforts to identify the broken windows tipping point hold important implications for how police resources should be directed and for evaluations of the effectiveness of policing disorder initiatives. Knowledge of the location of the broken windows tipping point can also be used to inform evaluations that seek to validate the theoretical pathways of BWT.

According to Wilson and Kelling (1982), the tipping point lies somewhere between two stable neighborhood equilibria existing at opposing extremes: low disorder, low crime neighborhoods, and high disorder, high crime neighborhoods. They recognize that some neighborhoods are so demoralized and crime-ridden that the best the police can do is react to calls for service (CFS), while some neighborhoods are so stable and serene that order-maintenance policing is not needed (Wilson & Kelling, 1982). In other words, the implementation of order-maintenance policing initiatives at these extremes is a poor use of police resources. Rather, Wilson and Kelling (1982) argue that the best use of limited police resources is to stabilize neighborhoods located at the tipping point. Unfortunately, they provide limited detail on how to identify such neighborhoods (Wilson & Kelling, 1982).

A logical first step toward identifying neighborhoods at the tipping point is to examine the functional form of the relationship between disorder and violent crime. A focus on functional form is necessary in order to generate a more holistic theoretical understanding of the disorder-crime relationship. Unfortunately, within the social sciences, generally, and Criminology, specifically, very few theorists have specified the functional form of the causal relationships explicated in their theories. In the absence of this information, many researchers are quick to assume a linear relationship. To this point, Wilson and Kelling's (1982) description of neighborhood extremes leaves open the possibility of a *linear* relationship between disorder and violent crime. That is, for every unit increase in disorder, violent crime increases by a constant amount. In fact, the vast majority of broken windows research has assumed a linear relationship (e.g., Skogan, 1990; Sampson & Raudenbush, 1999; Harcourt, 2001; Eck & Maguire, 2005; Steenbeek & Kreis, 2015; Wheeler, 2018; Konkel, Ratkowski, & Tapp, 2019).

In one such study, Steenbeek and Kreis (2015) developed a method for identifying areas with “lukewarm” levels of disorder; areas that neither have too little, nor too much disorder. As a general rule, they argue that it is reasonable to assume that “areas at the tipping point rank near the middle of the range of values on the disorder scale” (Steenbeek & Kreis, 2015, p. 527). Indeed, this assumption is supported by Wilson and Kelling’s (1982) description of the broken windows tipping point. Overlooked by Steenbeek and Kreis (2015), however, is that Wilson and Kelling’s (1982) description of a tipping point - at which small increases in disorder have a large impact on violent crime - is suggestive of a *nonlinear* relationship between disorder and violent crime. For this reason, Steenbeek and Kreis’ (2015) approach, if applied, may result in the misdirection of police resources, a consequence that holds especially severe implications for police departments that have limited resources and high levels of violent crime.

A nonlinear effect can be described in terms of a dose-response relationship, wherein the causal variable - represented as a level (or dose) - does not have a proportional effect on the outcome variable of interest across the range of its distribution (see Galster, 2014, 2018). Threshold effects are a special kind of nonlinear effect in which the impact of the causal variable dramatically changes once some critical level is surpassed (see Galster, 2014, 2018). Wilson and Kelling’s (1982) description of a tipping point suggests that disorder maintains a threshold effect on violent crime. To elaborate, their theory suggests that neighborhood disorder will not have a large impact on violent crime so long as it does not exceed too far beyond the lower equilibrium (Wilson & Kelling, 1982). In an ideal scenario, informal social control would work to decrease disorder and return the neighborhood back to a low disorder, low crime state. However, if disorder were to go unaddressed for too long and reach too high a level, informal social control would become ineffective. Without intervention, the neighborhood would eventually be propelled into a high disorder, high crime state.

In a competing view, Crane (1991) suggests in his “Epidemic Theory of Urban Ghettos” that the impact of disorder on violent crime should be the greatest in the worst quality neighborhoods: urban ghettos. Thus, Crane (1991) places the location of the tipping point not in the middle of the disorder distribution, but near its end. Ultimately, policing initiatives directed to disorder and/or crime hot spots that achieve significant crime control gains make alternative interpretations of the broken windows tipping point, such as Crane’s (1991), more attractive (see Braga et al., 1999; Braga & Bond, 2008; Braga, Hureau, & Papachristos, 2012, 2014).

In the only study that considers a nonlinear relationship between disorder and violent crime, Geller (2007) used a first-difference model to evaluate Wilson and Kelling’s (1982) and Crane’s (1991) interpretations of the tipping point. She found no support for either (Geller, 2007). Instead, she identified “an increasing concave relationship, in which the disorder-crime relationship is strongest in lower-disorder neighborhoods” (Geller, 2007, p. 87). While supportive of a nonlinear effect, this finding is inconsistent with the traditional understanding of a tipping point as a threshold effect.

In summary, evaluations of the theoretical foundation and effectiveness of order-maintenance policing initiatives cast doubt on whether BWT contributes a unique and valuable framework to the field of Criminology. That being said, broken windows research has failed to heed Wilson and Kelling’s (1982) key instruction for the implementation of such initiatives: to identify neighborhoods at the tipping point. Their description of a tipping point suggests a threshold effect, whereby the impact of disorder on violent crime dramatically increases past a critical level of disorder located somewhere in the middle of the disorder distribution (Wilson & Kelling, 1982). However, in the only examination that models the relationship between disorder and violent crime as nonlinear, Geller (2007) failed to find support for either Wilson and Kelling's (1982) or Crane's (1991) interpretation of the broken windows tipping point. Instead, she identified a nonlinear relationship that is not consistent with a threshold effect (Geller, 2007). Ultimately, more research must be conducted on the disorder-crime relationship before any conclusions can be drawn regarding the broken windows tipping point. Importantly, this research must utilize methods that minimize concerns about selection bias and account for the possibility of nonlinearity.

Research Aims

This study examines the functional form of the relationship between disorder and violent crime in an effort to shed light on the broken windows tipping point. This knowledge is especially beneficial for cities that have limited police resources to combat violent crime and contain neighborhoods in desperate need of revitalization. For this reason, this study focuses on the city of Detroit, Michigan. At an area of approximately 139 square-miles, Detroit contains a population of approximately 672,662 (as of 2018). It is a predominantly African American city characterized by high levels of poverty and violent crime.

Detroit was once considered the affluent capital of the Great Lakes region; a city with a vibrant culture, developed infrastructure, and strong auto-manufacturing economy. However, this label of affluence slowly began to disintegrate. Reliance on auto-manufacturing and competition with foreign adversaries proved to be a catastrophic combination, resulting in the eventual deindustrialization of Detroit and pervasive joblessness which reached its peak in 2009 during the Great Recession (2007– 2009). Detroit's experience, however, was not unlike other rust belt cities (e.g., Cleveland, OH; Gary, IN; and Youngstown, OH) which also relied heavily on auto-manufacturing. During this time, out-migration and financial insecurity resulted in an explosion in the number of abandoned structures, driving the creation of the Detroit Demolition Program which organizes their sale or demolition.[1] To this point, a report released in 2014 by the city of Detroit found that nearly a third of its land parcels (84,641) had been abandoned by their owners.[2] The majority of these parcels contained structures in very poor condition, indicated by broken or boarded windows/doors, fire damage, and/or a collapsed porch/roof (Blight Removal Task Force Plan, 2014).

Given its deteriorated urban landscape and high level of violence, the implications of the tipping point for police operations in Detroit has received much attention. When asked about adopting order-maintenance policing in Detroit neighborhoods, Kelling (as cited in Williams, 2012) responded: "You have a cluster of nice homes and then you get three blocks of homes that may make it or may not make it…The policing issue is how much do you invest in those areas that are at the tipping point and how much do you invest in the areas that have been largely destroyed? I think those are policy decisions Detroit is facing up to." Kelling's (as cited in Williams, 2012) statement underscores the importance of efforts to identify the tipping point for police operations in Detroit and motivates this study.

To this end, this study examines the relationship between physical disorder, measured at t_1, = 2014 on violent crime rate, measured at t_2 = 2015, while also addressing the possibility that factors that are associated with different levels of physical disorder are not proportional across levels. Importantly, this study's focus on physical disorder complements the city of Detroit's

---

[1] Emerging from the Great Recession, Detroit experienced a 16.18% decrease in its population, Gary a 17.04% decrease, Cleveland a 10.10% decrease, and Youngstown a 9.83% decrease.
[2] Land parcel is a term used in real-estate to define a plot of land that is owned (or intended to be owned) by a person or entity.

interest in efforts that address and mitigate the effects of blight – physical indicators of neighborhood decline - in order to spur neighborhood revitalization. Furthermore, this study utilizes an extension of propensity score matching - the generalized dose-response propensity score (GPS) approach - to conduct its assessment. This approach is uniquely suited for the identification of tipping points because it "explicitly model(s) the functional form of the level of a causal variable and a given outcome" while also addressing selection effects, when possible, through covariate balancing across matched levels of the causal variable (Mears et al., 2013, p. 460). In particular, this approach is superior to traditional regression – which addresses selection effects through the addition of control variables - because covariate balance must be achieved across matched levels of the causal variable before they can be reliably compared. To facilitate this analysis, this study collects census block-group level data on physical disorder, violent crime, as well as socioeconomic and land use characteristics from a variety of sources, including the Detroit Police Department's (DPD) record management system, Motor City Mapping (MCM) project, and Census.

In summary, this study seeks to advance knowledge on the functional form of the disorder-crime relationship. Evidence that supports a *nonlinear relationship* between disorder and violent crime rate - whether consistent or not with Wilson and Kelling's (1982) interpretation of the broken windows tipping point – would suggest that a proper test of the validity of BWT must accommodate nonlinearity. It would also support the idea that the impact of order-maintenance policing on violent crime may be more optimal in some neighborhoods than in others. The peculiarities of the disorder-crime relationship would provide some indication of where to focus these initiatives in order to achieve the greatest crime control benefits. Alternatively, evidence that supports a *linear relationship* between disorder and violent crime rate would also be of value. This finding would provide validation for research that has modeled the disorder-crime relationship as linear. It would also lessen concerns regarding where to implement order-maintenance policing initiatives. Regardless of which functional form is supported, police-community relations and police resources are two factors that stand to mitigate the effectiveness of order-maintenance policing initiatives, and therefore deserve the utmost consideration.

Study Outline

Moving forward, 'Chapter 2: Literature Review' consists of four sections. The first section – 'Broken Windows Theory' – summarizes BWT and its empirical support, as well as interprets the functional form of the disorder-crime relationship based upon Wilson and Kelling's (1982) description of the tipping point. It also highlights the failure of broken windows research to consider the tipping point and discusses the implications of this failure. With this foundation set, the second section – 'A Closer Look: Alternative Perspectives' – reviews research that challenges BWT, including alternative perspectives of the tipping point. Shifting gears, the discussion turns to order-maintenance policing. The third section - 'Policing Disorder & The Broken Windows Tipping Point' – describes the central tenants of order-maintenance policing and strategies used to address disorder, with a particular focus on Detroit. It also reviews the empirical evidence of policing disorder strategies before proceeding to a discussion of the implications of the tipping point for police operations. The final section – 'Current Study' - presents this study's primary research question and hypotheses, setting the stage for the following chapter.

'Chapter 3: Project Design and Implementation' consists of three sections. The first section – 'Measures and Data Sources' – describes the measures used to evaluate this study's hypotheses and their respective data sources. In particular, it highlights the strengths of this study's causal (i.e., physical disorder) and outcome (i.e., violent crime rate) measures. The second section – 'Analytical Strategy' – describes the GPS approach and its merits as they relate to the ability to make causal inferences. The final section – 'Sensitivity Checks' – describes a series of additional precautions that were taken to inspire confidence in this study's findings. These precautions include both parametric and nonparametric techniques of estimating the dose-response function.

'Chapter 4: Analysis & Results' consists of three sections. The first section – 'Parametric Method' – describes the outcome of each step of the parametric GPS approach, with particular attention given to the estimation of the dose-response function. The second section – 'Semiparametric Methods' - describes the estimation of the dose-response function produced from three types of semiparametric methods: penalized spline, radial spline, and inverse weighting kernel function. The final section – 'Summary of Findings' –summarizes this study's key research findings.

'Chapter 5: Discussion & Conclusion' consists of three sections. The first section – 'An Overview: The Search for the Broken Windows Tipping Point' – briefly summarizes the motivation behind this study, as well as its design, implementation, and results. The second section – 'Directions for Future Research' – identifies and discusses four domains in which research can be developed to advance knowledge of the functional form of the disorder-crime relationship: 1) Measures of disorder; 2) Neighborhood context; 3) Confounding factors; and 4) Longitudinal data analysis. The final section – 'Closing Remarks' – discusses the ways in which this study contributes to theory, practice, and policy.

CHAPTER 2: LITERATURE REVIEW

Broken Windows Theory

Theoretical Foundation of the Broken Windows Theory

In their seventeen-page article in *The Atlantic Monthly*, Wilson and Kelling (1982) described a developmental sequence in which unattended disorder produces deleterious consequences for communities, most outstanding among them the proliferation of violent crime. They offered the police a simple solution to prevent this sequence of events from unfolding: address disorder before it escalates (Wilson & Kelling, 1982). This solution is based on the idea that minor issues will eventually lead to a breakdown of informal social control within neighborhoods. Informal social control reflects the ability of a neighborhood to exert control over the behavior of its residents and its capacity to socialize them conventionally (Bursik, 1988; Bursik & Grasmik, 1993; Sampson & Groves, 1989). When strong, informal social control helps thwart the proliferation of violent crime within neighborhoods (Sampson & Groves, 1989; Bursik & Grasmick, 1993; Wilson, 1996; Warner & Rountree, 1997; Sampson, Raudenbush, & Earls, 1997; Bursik, 1999; Morenoff et al., 2001; Kubrin & Weitzer, 2003). Furthermore, Wilson and Kelling (1982) reason that more severe problems can be avoided if disorder is quickly addressed. However, if disorder is left unchecked, the broken windows developmental sequence will unfold over time.

According to Wilson and Kelling (1982), unattended disorder lowers the benchmark for expected and acceptable behaviors within neighborhoods, signaling a breakdown of informal social control. In doing so, it invites more disorder to occur. However, the presence of disorder alone is not enough to trigger the next step in the broken windows developmental sequence. Residents must perceive disorder to be a problem within their neighborhoods and interpret it as a consequence of failing social controls. As a response to worsening neighborhood conditions, Wilson and Kelling (1982) contend that residents will eventually become fearful and withdraw from community life, and may leave the neighborhood altogether. This effect further undermines informal social control. As informal social control weakens and disorder and minor crimes increase, criminals become emboldened to commit more severe criminal acts, interpreting disorder as a cue of neighborhood disinvestment. Accordingly, Wilson and Kelling (1982, para. 25) state, "[i]f the neighborhood cannot keep a bothersome panhandler from annoying passersby, the thief may reason, it is even less likely to call the police to identify a potential mugger or to interfere if the mugging actually takes place." Following this logic, neighborhoods with high levels of disorder are more likely to experience increases in crime than neighborhoods in which informal social control is effectively exercised to constrain and/or eliminate it. Once rising crime rates are noticed by residents, their fear and isolation from community life deepens. This acknowledgement serves to further entrench the neighborhood in a cycle of disorder and decline. Unfortunately, the time-frame in which this process is expected to unfold is unclear and has yet to be fully explored.

Missing from this overview of the broken windows developmental sequence is Wilson and Kelling's (1982) description of the tipping point. In their article, they devote a a meager paragraph toward describing it:

> Some neighborhoods are so demoralized and crime-ridden as to make foot patrol useless; the best the police can do with limited resources is respond to the enormous number of calls for service. Other neighborhoods are so stable and serene as to make foot patrol unnecessary. The key is to identify neighborhoods at the tipping point—where the public order is deteriorating but not unreclaimable, where the streets are used frequently but by apprehensive people, where a window is likely to be broken at any time, and must quickly be fixed if all are not to be shattered. (Wilson & Kelling, 1982, para. 47)

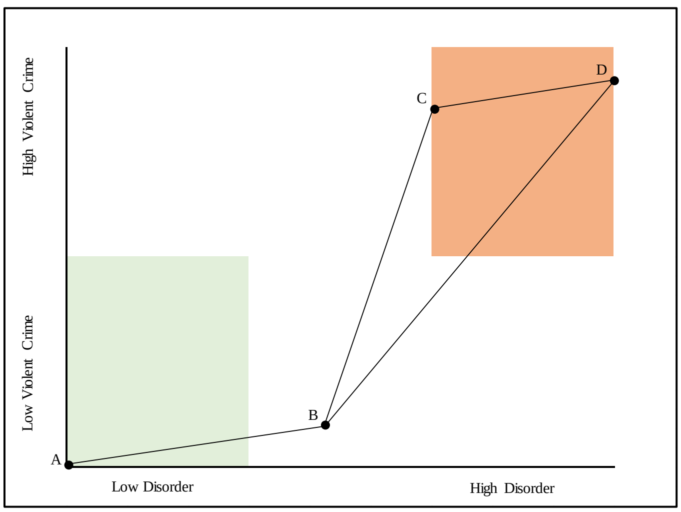
From what little information they provide, it can be determined that the tipping point is located somewhere between two opposing neighborhood extremes: low disorder, low crime neighborhoods, and high disorder, high crime neighborhoods (Wilson & Kelling, 1982). Notwithstanding its implications for policing, later discussed in detail, the tipping point affects our understanding of the relationship between disorder and violent crime in two important ways.

First, Wilson and Kelling's (1982) description of a tipping point suggests that the impact of disorder on violent crime is nonlinear. Specifically, it suggests that this relationship can be best captured as a threshold effect. As previously mentioned, threshold effects are a special kind of nonlinear effect in which the impact of the causal variable dramatically changes once some critical level is surpassed (see Galster, 2014, 2018). According to Wilson and Kelling (1982), this critical level – the tipping point - is located somewhere between two neighborhood extremes: low disorder, low crime neighborhoods, and high disorder, high crime neighborhoods. For this reason, efforts that seek to shed light on the broken windows tipping point must accommodate nonlinearity.

Second, Wilson and Kelling's (1982) vague description of the broken windows tipping point requires that we make several *theoretically-informed* assumptions regarding the exact functional specifications of the relationship between disorder and violent crime. To this point, two competing interpretations of this relationship emerge from their description of the tipping point (Wilson & Kelling, 1982). These interpretations are visually depicted in Figure 1 by the line segments $\overline{ABCD}$ and $\overline{ABD}$. Both line segments share the same origin (at point $\dot{A}$) and location of the broken windows tipping point (at point $\dot{B}$). Importantly, the *simplest* construction of each line segment was selected to reflect each hypothesized functional form.

Recall Wilson and Kelling (1982, para. 47) argue that some places are “so stable and serene as to make foot patrol unnecessary.” This description suggests that efforts to address disorder in low disorder, low crime neighborhoods will not be a worthwhile investment of police resources. In other words, we can expect the strength of the relationship between disorder and violent crime in these neighborhoods to be such that efforts to decrease disorder will not produce large enough crime reduction gains (i.e., the amount by which crime is reduced) to warrant police efforts. This relationship is depicted by the line segment $\overline{AB}$.

Figure 1. Theoretically-derived Interpretations of the Relationship between Disorder and Violent.



Beyond the tipping point, however, there are two viable interpretations of the relationship between disorder and violent crime. Recall Wilson and Kelling (1982, para. 47) argue that “some neighborhoods are so demoralized and crime-ridden as to make foot patrol useless” and that “the best the police can do with limited resources is respond to the enormous number of calls for service.” What motivates the following competing interpretations is how we come to understand the factors that render foot patrol useless. For example, we can apply a similar interpretation as before to inform our understanding of the relationship between disorder and violent crime in high disorder, high crime neighborhoods: The strength of the relationship between disorder and violent crime in these neighborhoods is such that efforts to decrease disorder will not produce large enough crime reduction gains to warrant police efforts. This relationship is captured by the line segment $\overline{CD}$ and completes the line segment $\overline{ABCD}$.

Alternatively, we can interpret Wilson and Kelling's (1982) description to suggest that the dosage of police response needed to address disorder in high disorder, high crime neighborhoods is extraordinary, and beyond what police departments with limited resources are equipped to provide. Thus, the ineffectiveness of foot patrol in high disorder, high crime neighborhoods is now an issue of inadequate police response dosage. Therefore, if police resources are plentiful or highly focused on small areas (i.e., hot spots), there is a potential to achieve significant crime reduction gains. This relationship is captured by the line segment $\overline{BD}$ and completes the line segment $\overline{ABD}$.

Three sets of neighborhood characteristics emerge from this assessment. These characteristics are based upon a neighborhood's location relative to the broken windows tipping point: 1) Before the tipping point; 2) At the tipping point; and 3) Beyond the tipping point. In neighborhoods that are located *before* the tipping point, disorder, fear of crime, and violent crime are low, and residents have the opportunity to build and exercise informal social control. If disorder does not exceed too far beyond the lower equilibrium, informal social control is expected to decrease disorder, helping return the neighborhoods back to a low disorder, low crime state (Wilson & Kelling, 1982).

In neighborhoods located *at* the tipping point, disorder is mounting and about to reach a level that will elicit a significant fear response, resulting in social withdrawal followed by a breakdown of informal social control and an uptick in violent crime (Wilson & Kelling, 1982). Public order is deteriorating and along with it the ability of informal social control to constrain and/or eliminate disorder (Wilson & Kelling, 1982).[3] Wilson and Kelling (1982) advocate the use of formal mechanisms in neighborhoods located at the tipping point; these neighborhoods are those that are at the greatest risk of being propelled into a high disorder, high crime state. To this point, they state that "[t]hough citizens can do a great deal, the police are plainly the key to order-maintenance" (Wilson & Kelling, 1982, para. 46).

In neighborhoods located *beyond* the tipping point, disorder, fear of crime, and violent crime are high, and social isolation prevents residents from contributing to the development of informal social control. Without resident involvement, informal social control will disintegrate. This description paints a relatively bleak image of neighborhoods located beyond the tipping point. However, there is a path forward. Problem-oriented policing initiatives directed to disorder and/or crime hot spots have been shown to be particularly effective at reducing crime levels without significant displacement or damage to police-community relations (Braga & Bond, 2008; Braga et al., 1999; Braga et al., 2012, 2014). Often a part of problem-oriented policing initiatives, neighborhood revitalization efforts also have been shown to reduce crime levels in declining neighborhoods, although they may come at the cost of gentrification (see MacDonald & Stokes, 2019). Examples of these efforts include the demolition/rehabilitation of vacant housing (e.g., Kondo et al., 2016; Spader et al., 2016; Wheeler et al., 2018; Jay et al., 2019; Larson et al., 2019) and transformation of vacant lots (e.g., Garvin et al., 2013; Kondo et al., 2018; Branas et al., 2018), as well as the creation of defensible spaces (e.g., Jeffery, 1971;

[3] Outside of the broken windows perspective, there are a variety of factors that may undermine the effectiveness of informal social control, such as the strength, density, and type of social ties (Sampson & Groves, 1989; Bursik & Grasmick 1993; Bursik 1999; Sampson, 2012; Browning et al., 2017), as well as the level of neighborhood attachment (Kitts, 1999; Rohe & Stegman 1994; Silver & Miller, 2004; Xiao & McCright, 2014), and shared local exposure (Jacobs, 1961; Small, 2009; Browning et al., 2017).

Newman, 1972, 1996; Brown & Altman, 1983; Taylor, Gottfredson, & Brower, 1984; Ratcliffe, 2003; Eck & Guerette, 2012).

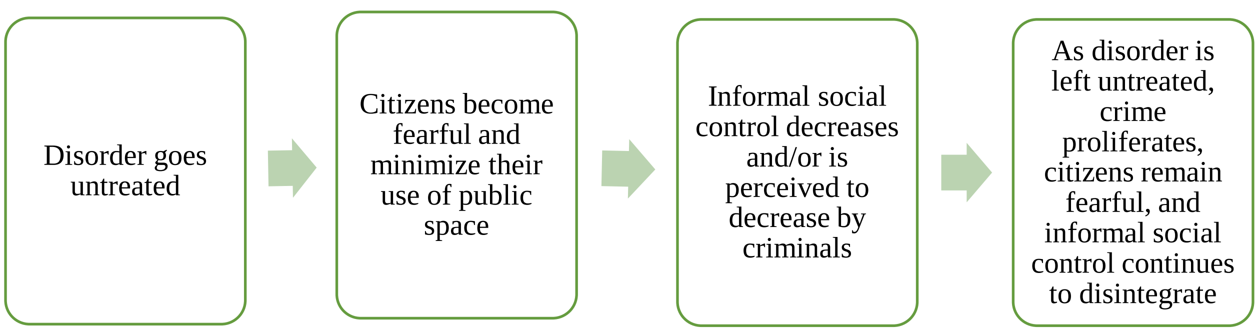
Ultimately, it is possible for neighborhoods that are located beyond the tipping point to experience reductions in both disorder and crime levels. However, the question remains whether the efforts discussed can change neighborhoods *enough* to transition them from a high disorder, high crime state, to a low disorder, low crime state. In order to begin to address this question, it is first important to recognize that not all neighborhoods start off on equal footing. Group-based trajectory models (GBTMs) of crime suggest as much (Weisburd et al., 2004; Yang, 2010; Weisburd, Groff, & Yang, 2012; Curman, Andresen, & Brantingham, 2014; Wheeler et al., 2016; Andresen, Curman, & Linning, 2017; Gill, Wooditch, & Weisburd, 2017). This modeling technique is able to capture the developmental patterns of crime as they unfold over time, revealing patterns of stability and change.

Neighborhood structural features help shed light on the crime patterns revealed by GBTMs. Drawing from social disorganization theory, these features traditionally include physical disorder, poverty, residential instability, ethnic heterogeneity, and concentrated disadvantage. Furthermore, a large body of research has found neighborhood structural features to be positively associated with crime (Shaw & McKay, 1942; Sampson et al., 1997; Boggess & Hipp, 2010; Steenbeek & Hipp, 2011; Hipp, Kim, & Kane, 2019), and, in particular, high crime trajectories (Weisburd et al., 2012; Gill et al., 2017; Krivo et al., 2018). Of these features, physical disorder is arguably the easiest feature to alter. In light of these findings, the effort required to facilitate a neighborhood's transition from a high disorder, high crime state to a low disorder, low crime state is likely influenced by the degree of flexibility afforded by its trajectory, as well as the systemic features that contribute to it. Thus, neighborhoods that are on a high and stable crime trajectory will likely be the most difficult to change.

Empirical Support for Broken Windows Theory: Implications for the Broken Windows Tipping Point

The broken windows developmental sequence, as described above, is depicted in Figure 2. As can be clearly seen, the path from disorder to violent crime comprises three core theoretical propositions. These propositions must be supported for BWT to be in a position to meaningfully contribute to the field of Criminology. First, unattended disorder must lead to an increase in residents' fear of crime. Second, fear of crime must lead residents to withdraw from community life, resulting in a decrease in informal social control. Third, violent crime must increase in response to declining levels of informal social control.

Figure 2. The Broken Windows Developmental Sequence.



There is considerable evidence supporting these linkages. Research has largely identified a positive association between disorder - including both systematically observed and perceived levels - and fear of crime (e.g., Taylor & Shumaker, 1990; Covington & Taylor, 1991; LaGrange, et al., 1992; McGarrell, Giacomazzi, & Thurman, 1997, 1999; Markowitz et al., 2001; Spelman, 2004; Hinkle & Weisburd, 2008). Furthermore, there has been a comparatively smaller body of research with a greater degree of mixed findings that examines the relationship between fear of crime and community withdrawal and informal social control, primarily captured as collective efficacy.[4] That being said, this pathway still garners a fair amount of support (e.g., Garofalo, 1981; Markowitz et al., 2001; Crank, Giacomazzi, & Heck, 2003). Lastly, a rich line of criminological inquiry has tied reductions in collective efficacy to an increase in crime (e.g., Kasarada & Janowitz, 1974; Sampson, 1988; Sampson & Groves, 1989; Bursik & Grasmick, 1993; Wilson, 1996; Sampson et al., 1997; Warner & Rountree, 1997; Bursik, 1999; Morenoff et al., 2001; Browing, 2002; Kubrin & Weitzer, 2003; Lowenkamp, Cullen, & Pratt, 2003; Sabol, Coulton, & Korbin, 2004; Armstrong, Katz, & Schnebly, 2015), securing the final link of the broken windows developmental sequence.

Despite this large body of evidence, the majority of broken windows research lends support to only one step of the broken windows developmental sequence and are challenged - to varying degrees - by competing findings. To this point, a recent meta-analysis of 96 studies on the effect of disorder on aggressive behaviors and fear of crime conducted by O'Brien, Farrell, and Welsh (2019) failed to find consistent support for the relationships laid out in BWT. However, it is important to consider that BWT details a longitudinal process of neighborhood decline. Yet, most studies that seek to validate BWT utilize cross-sectional data and, therefore, are unable to capture the dynamics of the broken windows developmental sequence. In fact, all but six studies included in O'Brien et al.'s (2019) review utilized cross-sectional data. Furthermore, very few studies evaluate the cyclic nature of BWT by accounting for reciprocal effects between crime and neighborhood conditions (Sampson & Raudenbush, 1999; Markowitz et al., 2001; Robinson et al., 2003; Steenbeek & Hipp, 2011; Boggess & Maskaly, 2014; O'Brien & Sampson, 2015). In arguably the most complete examination of BWT to date, Steenbeek and

[4] Traditionally defined, collective efficacy includes two key components: social cohesion and informal social control (Sampson et al.,1997). In general, collective efficacy requires trust and solidarity amongst resident, as well as their willingness to intervene to maintain order within neighborhoods (Sampson et al., 1997).

Hipp (2011) examined 10 years of neighborhood data in a series of sophisticated longitudinal cross-lagged models and concluded:

> [T]he results suggest a cyclical model in which neighborhoods have relatively stable levels of disorder overtime, and the processes that lead to disorderly neighborhoods are difficult to turn around. Neighborhoods with high levels of disorder cause more people to move out, and higher residential instability leads to a lower percentage of people taking action to improve the livability and safety of the neighborhood. Neighborhood disorder thus has cumulative effects over and above the direct effect on residential instability by reinforcing itself via a weakening of community processes of social control. (p. 864)

Overall, they found considerable support for the longitudinal process of neighborhood decline hypothesized by Wilson and Kelling (1982) (Steenbeek & Hipp, 2011).

It should be clear that Wilson and Kelling (1982) argue that the primary pathway through which disorder affects violent crime is through fear and social withdrawal, leading to lower levels of informal social control within neighborhoods (see also review in Gualt & Silver, 2008). Despite this argument, many researchers have interpreted BWT to suggest a direct relationship between disorder and violent crime (e.g., Skogan, 1990; Sampson & Raudenbush,1999; Harcourt, 2001; Eck & Mcguire, 2005). Indeed, the role of disorder as a cue which signals to offenders that no one cares, in turn inspiring them to commit crime, is consistent with this understanding. To this point, in a later article Wilson and Kelling (1989, p. 47) imply a direct relationship between disorder and crime:

> A rash of burglaries may occur because drug users have found a back alley or an abandoned building in which to hang out. In their spare time, and in order to get money to buy drugs, they steal from their neighbors. If the back alleys are cleaned up and the abandoned buildings torn down, the drug users will go away.

Skogan (1990) was the first to seriously consider the direct relationship between disorder and violent crime. In *Disorder and Decline*, he identified a significant positive relationship between disorder and robbery, controlling for poverty, residential stability, and racial composition (Skogan, 1990). Using the same data, Harcourt (2001) applied a “corrected” approach that addressed several serious issues surrounding Skogan's (1990) consideration of missing values, construction of independent variables, and narrow focus on only one crime outcome. After removing neighborhoods with strong disorder-crime ties, Harcourt (2001) failed to identify any significant relationships between disorder and crime. However, Eck and Maguire (2005) later argue that Harcourt's (2001) study did not disprove Skogan's (1990) findings in support of the disorder-crime link. Rather, Harcourt (2001) discovered that the data were affected by outliers.

Sampson and Raudenbush (1999) utilized data from the Project on Human Development in Chicago Neighborhoods (PHDCN) to investigate the disorder-crime link. Through weighted least squares regression and variable path analysis, they identified a positive direct link between disorder and violent crime (Sampson & Raudenbush, 1999). In all cases except for robbery, however, this link disappeared when collective efficacy was introduced into the model, defined as "the linkage of cohesion and mutual trust with shared expectations for intervening in support of neighborhood social control" (Sampson & Raudenbush, 1999, p. 612-613). They also identified a reciprocal relationship between collective efficacy and crime, where collective efficacy negatively affected crime and crime negatively affected collective efficacy (Sampson & Raudenbush, 1999). Ultimately, Sampson and Raudenbush (1999, p. 627) argue that their results "point to a spurious association of disorder with predatory crime." Later, Bratton and Kelling (2006) denounced Sampson and Raudenbush's (1999) study in an article published in the *National Review*:

> They [Sampson and Raudenbush (1999)] claimed that broken windows posits a *direct* link between disorder and serious crime. From the first presentation of broken windows we have argued, to the contrary, that the link, while clear and strong, is *indirect*. Citizen fear, created by disorder, leads to weakened social controls, thus creating the conditions in which crime can flourish. (para. 9)

Supporting Bratton and Kelling's (2006) argument, Xu, Fielder, and Flaming (2005) argue that Sampson and Raudenbush's (1999) discovery of a reciprocal relationship between collective efficacy and crime in fact supports an indirect link between disorder and crime. Utilizing a different data source, they demonstrated that disorder has strong direct, indirect, and total effects on crime even while controlling for collective efficacy (Xu et al., 2005).

Overall, studies that evaluate the direct relationship between disorder and crime have done little to produce a clearer image of this relationship. At their worst, they fail to find a significant relationship (e.g., Harcourt, 2001; Sampson & Raudenbush, 1999). While disorder has been found to have a strong direct effect on crime (e.g., Xu et al., 2005), it is more often the case that a modest effect is identified (e.g., Boggess & Maskaly, 2014; Wheeler, 2018; Konkel et al., 2019). It is also common for this effect to vary by crime type and/or type of disorder (e.g., Sampson & Raudenbush, 1999; Taylor, 1999, 2001).

This review of empirical evidence of BWT clearly demonstrates its highly contentious standing within the field of Criminology and the need for more complete tests of the broken windows developmental sequence that draw upon longitudinal data and consider reciprocal effects between crime and neighborhood conditions, such as collective efficacy. The current study acknowledges the mixed body of findings revealed from its review that provide sufficient grounds on which to question the validity of BWT. However, it has no intention of directly addressing these findings. Rather, it seeks to advance broken windows research in another way: by exploring the functional form of the relationship between disorder and violent crime in an effort to shed light on the broken windows tipping point.

Studied directly, evaluations of the disorder-crime link ignore the social-psychological underpinnings of BWT and are incomplete tests of the theory. However, a direct evaluation is an appropriate starting place for efforts that seek to evaluate the function form of the disorder-crime relationship given Wilson and Kelling's (1982) description of the tipping point as located somewhere between low disorder, low crime neighborhoods and high disorder, high crime neighborhoods. If this study finds evidence in support of a nonlinear relationship, then it would suggest that future evaluations of the validity of BWT must accommodate the possibility of nonlinearity, and that past evaluations which failed to do so my have over- or under-stated the effect of disorder on violent crime based upon the nuances of this relationship.

A Closer Look: Alternative Perspectives

Unfortunately, there is a surprising dearth of studies on the broken widows tipping point. The vast majority of studies do not explicitly evaluate the broken windows tipping point, nor do they consider how the tipping point may impact their findings. As previously discussed, Wilson and Kelling's (1982) interpretation of the tipping point suggests a threshold effect of disorder on violent crime. Nonetheless, most studies that examine a direct relationship between disorder and violent crime assume a linear trend in disorder. If the disorder-crime relationship has been misspecified, however, regression estimates and assumptions of statistical tests which assume linearity will produce misleading findings. Quite obviously, evaluations that misspecify the relationship between disorder and violent crime are unable to advance our understanding of the disorder-crime relationship or the broken windows tipping point, for that matter.

Without having conducted formal evaluations, a small number of studies *suggest* that the relationship between disorder and violent crime may in fact be nonlinear (Taylor & Shumaker, 1990; Gau & Pratt, 2010; Yang, 2010). For example, Gau and Pratt (2010) utilized an ordinary least squares regression model to evaluate the interaction effect between perceptions of disorder and a disorder-crime difference score. This score was constructed by taking the absolute value of the difference between scores obtained from scales that measured perceptions of neighborhood crime and disorder problems. Thus, higher disorder-crime scores represent a larger disparity between perceptions of disorder and crime problems. Furthermore, Gau and Pratt (2010) divided their sample of respondents into two. One sample consisted of respondents who perceived low levels of disorder, and the other those who perceived high levels of disorder. Running a regression analysis for each sample, Gau and Pratt (2010) found that respondents who lived in orderly neighborhoods could not distinguish disorder from crime, but respondents who lived in disorderly neighborhoods could make this distinction. They argue that their findings suggest a nonlinear trend in disorder (Gau & Pratt, 2010). Beyond some critical threshold of disorder, respondents are better able to differentiate between disorder and crime.

In another study, Yang (2010) utilized group-based trajectory and joint trajectory analyses to evaluate the longitudinal relationship between disorder and crime. She found that while the absence of disorder ensured that a place would be free of violence, high levels of disorder only predicted violence problems 30% of the time (Yang, 2010). Furthermore, Yang (2010, p. 158) suggests that "perhaps the current results can be explained by the fact that violence only occurs in places where disorder has passed the 'tipping point.'" She instructs that future research should "focus on examining the possible existence of a *threshold* which must be surpassed in order for disorder to have impacts on crime" (Yang, 2010, p. 158).

Beyond Wilson and Kelling's (1982) description of the broken windows tipping point, nonlinear effects have been found to drive a number of other neighborhood-level processes (see review in Galster, 2018). In one such process pertinent to our understanding of the broken windows tipping point, Crane (1991) proposes a contagion model to understand the spread of social problems within communities. As implied, the model assumes that social problems are contagious. If they are kept below a critical threshold, their frequency and prevalence will eventually return to low levels. Beyond this threshold, however, social problems will spread like an epidemic, as increasing numbers of individuals engage in problematic behaviors. Crane (1991) identified two factors that determine the susceptibility of a community to an epidemic: 1) Residents' susceptibility to deviant peer influence; and 2) Residents' overall risk of developing social problems. Ultimately, he hypothesized that "[t]he relationships between neighborhood quality and the incidence of particular social problems should be nonlinear. Social problems should increase as neighborhood quality declines, but not at a constant rate. Somewhere near the bottom of the distribution of neighborhood quality, there should be a jump in the rate of increase" (Crane, 1991, p. 1228).

To explore his hypothesis, Crane (1991) examined the effect of neighborhood quality on high school dropout rates and teenage childbearing. In particular, he captured neighborhood quality as the percentage of individuals in a neighborhood that held either a managerial or professional job (Crane, 1991). Neighborhoods that were on the low range of this measure were considered to be of low quality, while neighborhoods on the high range were considered to be of high quality. Crane (1991) found the effects of neighborhood quality were the largest in the lowest-quality neighborhoods, otherwise referred to as urban ghettos (Crane, 1991). Insofar as disorder is an indicator of neighborhood quality and crime a social problem, Crane's (1991) contagion model offers an alternative perspective to Wilson and Kelling's (1982) interpretation of the broken windows tipping point. In particular, his perspective suggests that the impact of disorder on violent crime will be the most severe in urban ghettos (Crane, 1991). In other words, Crane's (1991) contagion model moves the tipping point from the middle of the disorder distribution towards its end.

Furthermore, Crane's (1991) assessment complements findings from a Detroit study conducted by Raleigh and Galster (2015) which explores the relationship between neighborhood disinvestment and violent crime rate. They utilized several attributes of Detroit neighborhoods to simulate five stages of neighborhood disinvestment, with the fifth stage representing the greatest level of disinvestment (i.e., highest levels of vacant land, vacant housing units, and renters; the lowest median incomes, employment rates, and population density) (Raleigh & Galster, 2015). Raleigh and Galster (2015) found the transition from one stage of disinvestment to the next to have a disproportional effect on violent crime rate. While violent crime rate increased at each transition, the final transition (from stage 4 to stage 5) experienced the largest increase in the growth rate. Unlike Crane's (1991) measure, Raleigh and Galster's (2015) measure of neighborhood quality (i.e., neighborhood disinvestment) captured indicators of physical disorder (vacant land and vacant housing units). For this reason, their findings are especially compelling in support of a nonlinear relationship between disorder and violent crime, with a tipping point located at the high end of the disorder distribution (Raleigh & Galster, 2015).

Adding to these findings, in areas with high levels of social problems consistent with urban ghettos, efforts to address disorder have been largely successful at reducing crime (Braga & Bond, 2008; Braga et al., 1999; Braga et al., 2012, 2014). In fairness, however, policing strategies that target disorder in hot spots are often embedded within problem-oriented and situational crime prevention strategies which draw from competing theoretical mechanisms to explain crime reduction, an issue that will be discussed later. For this reason, it is difficult to disentangle the effects of these complementary strategies. Nonetheless, these studies open the possibility of the broken windows tipping point being located toward the end of the disorder distribution. They also lend support to the interpretation of the relationship between disorder and violent crime depicted by the line segment $\overline{ABD}$ in Figure 1. Given sufficient police resources or a high spatial dosage of these resources, this relationship supports significant crime reduction gains in high disorder, high crime neighborhoods.

Motivated by Wilson and Kelling's (1982) description of a tipping point, Geller's (2007) evaluation is the only study that explores the functional form of the disorder-crime relationship. To elaborate, she used a first-difference model to capture the relationship between physical disorder and violent crime rate (Geller, 2007). The inclusion of a squared-term of physical disorder provided the model some flexibility, imposing a global structure on the relationship between physical disorder and violent crime rate. Geller (2007) identified a concave relationship in which the disorder-crime link was the strongest in low disorder neighborhoods (Geller, 2007). Setting aside for now issues regarding how nonlinearity was captured, her finding does garner support. To this point, Taylor and Shumaker (1990) argue in favor of the idea of inoculation, whereby the severity of previous experiences of an adverse phenomenon are lessened over-time as individuals adapt to their surroundings. They apply this idea to the relationship between fear and disorder: "slippage between fear and local disorder levels will be greater in locales where the level of disorder is higher, because residents in the higher threat contexts are experiencing a greater degree of perception adaptation" (Taylor & Shumaker, 1990, p. 629). Supporting inoculation, Taylor and Shumaker (1990) identified a concave relationship between disorder and fear, where the slope flattens then declines at high levels of disorder (see Figure 3). Furthermore, the idea of inoculation has been adapted to explain disparities in perceptions of disorder among residents living in the same neighborhood (see Sampson & Raudenbush, 2004; Franzini et al., 2008; Hipp, 2010; Sampson, 2012). In particular, Sampson and Raudenbush (2004) suggest that the greater past exposure residents have to disorder, the greater the amount of disorder they will need to be exposed to for them to perceive it to be a problem within their neighborhood.

Figure 3. Quadratic Effect of Disorder on Fear of Crime. Adapted from "Local Crime as a Natural Hazard: Implications for Understanding the Relationship between Disorder and Fear of Crime," by R.B. Taylor and S.A. Shumaker, 1990, *Environmental/Ecological Psychology, 18*(5), p. 631.

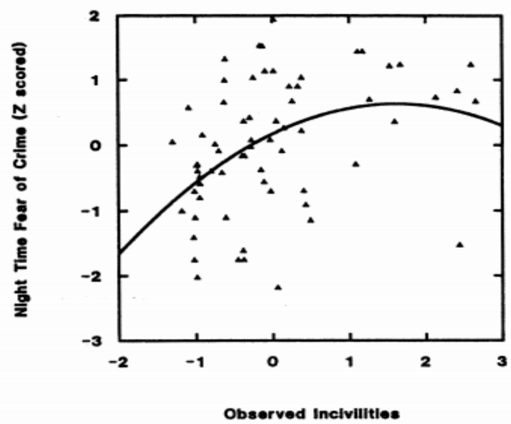


Fig. 1. Quadratic effect of social and physical incivilities on fear of crime. *Note.* Fear of crime item: "How safe do you feel or would you feel walking alone in your neighborhood at night?" Neighborhood average on this item has been z scored. Incivilities measure = principal components score based on several items measured by teams of raters. (See Taylor, Shumaker, and Gottfredson (1985) for details on this scale.) n = 66 neighborhoods. If the outlier neighborhood at bottom right is dropped, the quadratic effect is reduced but still noticeable and marginally significant.

Fear is a crucial response to disorder that is needed for the broken windows developmental sequence to unfold. As suggested by Taylor and Shumaker (1990), this sequence will be affected if residents become desensitized to disorder at high levels. There are three likely ways in which inoculation stands to affect the broken windows developmental sequence at high levels of disorder. First, the relationship between disorder and violent crime at high levels of disorder may flatten as residents adapt to disorder. In other words, an increase in disorder will not result in a proportionate (or greater) fear response which is needed to drive neighborhood decline. With levels of fear at (or nearly at) a constant, increases in disorder will no longer positively contribute to the perpetuation of violent crime within neighborhoods. Second, the relationship between disorder and violent crime at high levels of disorder may be negative. This relationship indicates a reversal of the broken windows developmental sequence. In this case, adaptation to disorder will result in a decrease in levels of fear. Less fearful, residents will be more likely to reclaim public space and seize opportunities to develop and exercise informal social control, resulting in a decrease in violent crime within neighborhoods. Third, a combination of these two outcomes is also possible. As disorder rises, levels of fear will plateau and eventually decline, resulting in a decrease in violent crime within neighborhoods through the previously described mechanisms.

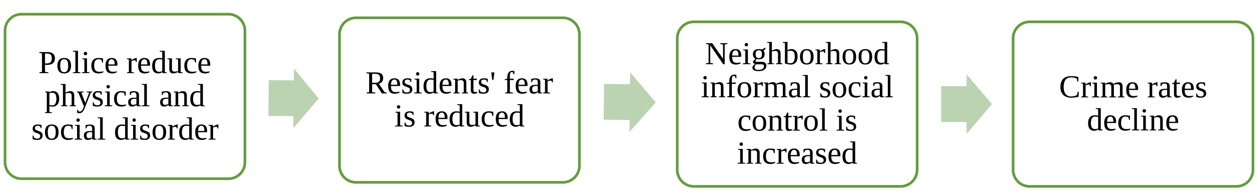
In summary, the mechanisms through which disorder affects crime are hotly debated and generate an understandable amount of skepticism regarding the value of BWT to the field of Criminology. That being said, research on BWT has failed considerably - with few exceptions - on two fronts: 1) It has failed to investigate whether disorder has a nonlinear effect on violent crime; and 2) More specifically, it has failed to consider how the tipping point – if it exists - may impact study findings. To-date, only one study investigates the nonlinear relationship between disorder and violent crime. Geller’s (2007) evaluation revealed a concave relationship which suggests that the disorder-crime link is the strongest in low disorder, low crime neighborhoods, a finding that is inconsistent with both Wilson and Kelling’s (1982) and Crane’s (1991) interpretations. However, this single study is far from conclusive. A topic that remains to be discussed is the impact of these failures for police operations. The following section provides an overview of order-maintenance policing with a particular focus on Detroit, followed by a focused discussion on the significance of the broken windows tipping point for policing strategies that address disorder.

Policing Disorder & The Broken Windows Tipping Point

What is Order-Maintenance Policing?

Originating from BWT, order-maintenance policing is a community-driven approach that seeks to reduce violent crime by addressing physical aspects of an environment and threatening behaviors within the public domain that inspire fear and upset community life. The primary goal of order-maintenance policing is to reinforce informal social control by reducing residents' fear associated with these physical and social phenomena. The logic being that with disorder no longer driving the broken windows cycle, residents will be less fearful and more inclined to use public space, in turn providing them more opportunities to engage in behaviors that fortify informal social control (see Figure 4). Thus, a clear gauge of whether order-maintenance policing is operating through the pathways laid out in BWT is whether it affects residents' levels of fear.

Figure 4. The mechanisms of Broken Windows Policing. Adapted from "Understanding the Mechanisms Underlying Broken Windows Policing: The Need for Evaluation Evidence," by D. Weisburd, J.C. Hinkle, A. Braga, and A. Wooditch, 2015, *Journal of Research in Crime and Delinquency, 52*(4), p. 594.



In its intended form, order-maintenance policing is shaped by negotiated rules for street-level order realized through police-community partnerships (Wilson & Kelling, 1982). In this way, it is compatible with a procedural justice framework since the manner in which the police exercise their legal authority is shaped by the community and presumed fair. However, this is not to say that order-maintenance policing cannot go awry. It is often unclear whether individuals who violate community-specific standards for public conduct are violating the law. For example, many social disorder violations are termed “soft crimes” and classified into ambiguous legal categories, such as disturbing the peace, loitering, and vagrancy. Due to the legal ambiguity of disorder violations, officers are not prompted to resolve such violations with arrests (Kelling & Coles, 1996). Rather, they are encouraged to use discretion to resolve issues (Kelling & Coles, 1996). To this point, non-arrest solutions are able to address issues before they escalate, and also serve to protect community and police relationships by discouraging aggressive enforcement (Bittner, 1967; Brown, 1981; Kelling & Coles, 1996; Gau & Brunson, 2010; Todak & James, 2018).

It is worth emphasizing that aggressive enforcement of minor offenses is a characteristic of *zero-tolerance policing*. Unlike order-maintenance policing, zero-tolerance policing damages police-community relations and is inherently in opposition with procedural justice, as epitomized by research which exposes it as a racially biased tactic (Harris, 1993; Gelman et al., 2007; Ridgeway, 2007; Gau & Brunson, 2010; Hanink, 2013; Gau, 2014; Rengifo & Fratello, 2015; Rengifo & Folwer, 2016). The rigidity of zero-tolerance policing also denies officers from using discretion in enforcing minor offenses. A staple of order-maintenance policing, this crucial tool is needed to safeguard police-community relations (Kelling & Coles, 1996). Kelling and Coles (1996, p. 9) make these distinctions clear and attack zero-tolerance policing as an unsustainable approach: “It's not a credible policy that the police are going to be able to implement for any length of time and offenders know that.”

Furthermore, order-maintenance policing seeks to reduce the physical and psychological distance between the police and residents in an effort to mount an appropriate and leveled response to address disorder and secure the cooperation and compliance of residents (Pate et al., 1985). Aligned with this objective, Wilson and Kelling’s (1982) original vision of order-maintenance policing was strongly focused on the implementation of foot patrol. This policing strategy is well known for its ability to increase perceptions of safety, as well as decrease fear of crime, and, to a lesser degree, crime (Kelling, 1981; Trojanowicz & Branas, 1985; Esbensen & Taylor, 1984; Cordner, 1986; Bowers & Hirsch, 1987; Skogan & Frydl, 2004; Ratcliffe, Taniguchi, Groff, & Wood, 2011; Piza & O'Hara, 2014; Groff et al., 2015; Andresen & Hodgkinson, 2018). As compared to car patrol officers, foot patrol officers are more likely to address disorder incidents, engage in public service activities and information gathering, and initiate pedestrian stops (Trojanowicz, 1986; Groff et al., 2012).

In addition, one-to-one contact with police officers has the potential to improve police-community relations by providing opportunities for residents and officers to become more familiar with one another and establish trust (Trojanowicz, 1986; Groff et al., 2015; Cowell & Kringen, 2016). To this point, the Newark Foot Patrol Experiment and the Flint Neighborhood Foot Patrol Program – early seminal studies - both identified foot patrol to have a positive effect on residents' satisfaction with the police (Kelling, 1981; Trojanowicz & Baldwin, 1982). Complementing this finding, more recent studies have garnered support for the ability of foot patrol to increase perceptions of the police as approachable, friendly, fair, accountable, and respectful (Cowell & Kringen, 2016; Simpson, 2017). These perceptions have been shown to encourage resident involvement in police efforts and strengthen existing police-community partnerships (Hinds, 2007; Reisig, 2007; Tyler & Fagan, 2008; Leroux & McShane, 2017).

Outside of foot patrol, another tactic to address disorder is through problem-oriented policing. In fact, order-maintenance policing can be understood to be a branch of problem-oriented policing which focuses on a particular type of problem: disorder. To this point, order-maintenance policing often involves strategic partnerships with local community groups, businesses, social services, and city agencies in order to develop solutions to address disorder (e.g., Weisburd & Green, 1995; McGarrell et al., 1997, 1999; Baker & Wolfer, 2003; Skogan, 2006; Braga, 2010; Braga, Hureau, & Papachristos, 2011; Taylor, Koper, & Woods, 2011; Weisburd et al., 2012). These solutions may or may not directly involve the police. For example, civil remedies have become a popular means through which to address disorder and are often used in conjunction with criminal penalties (see Figure 5). In particular, code enforcement and nuisance abatement are the most commonly used civil remedies. Unlike traditional order-maintenance policing, these tactics address both public and private displays of disorder.

Figure 5. Shorthand descriptions of some property-related civil remedies. From "Using civil actions against property to control crime problems," by M. J. Smith and L. Mazerolle, 2013, Center for Problem-Oriented Policing, 11, p.10.

Property-Related Civil Remedy	Target	Short-Hand Description of the Civil Action Outcome
Code enforcement	Landlord/owner	Enforces health and safety rules
Zoning	Landlord/owner	Limits activities and structures to particular areas or locations
Nuisance abatement	Landlord/owner (usually)	Returns to (or seeks to achieve) quiet enjoyment
Eviction	Tenant	Removes the tenant
Trespass	Uninvited persons	Removes the non-tenant
Civil injunction	Various	Orders someone to do or stop doing something immediately
Receivership	Landlord/owner	Gets someone else to manage the property
Condemnation	Landlord/owner	Locks it up and tears it down

Code enforcement refers to "the legal action taken by an enforcement body in response to a violation of one or more municipal health and safety codes" (Smith & Mazerolle, 2013, p.10). Nuisance abatement is considered to be a broader, more formalized version of code enforcement. As a municipal ordinance, nuisance abatement allows legal action to be taken in situations "in which a person is being deprived of his or her right to 'quiet enjoyment' by some existing condition, or by actions being carried out by another person, group, or business" (Worrall & Wheeler, 2019, p. 14). As such, nuisance abatement ordinances can take many forms.

Property-owners are motivated to comply to the standards established by a code or ordinance through a civil injunction. Consequences for noncompliance range in severity and may include a fine, jail-time, eviction, or forced closure or sale of the property (see Smith & Mazerolle, 2013). Property-owners may also be held civilly liable for illegal activities that occur on their properties (see Smith & Mazerolle, 2013). Outside of the police, enforcement relies on a broad range of actors (e.g., building, health, electrical, plumbing, and fire inspectors) that encompass a variety of agencies (see Smith & Mazerolle, 2013). The police may collaborate with these agencies by bringing problem properties to their attention, assisting on inspections, issuing notices of violations (e.g., excessive alcohol consumption, over-crowding, litter, overgrown foliage, unkempt properties, and abandoned/derelict buildings), and/or enforcing the consequences of noncompliance.

Around the time-frame of interest to this study (2014-2015), the DPD engaged in several efforts that aligned with the central tenants of order-maintenance policing. In 2012, the city of Detroit, on the verge of bankruptcy, enlisted the assistance of the Manhattan Institute and Bratton Group to facilitate the DPD's adoption of policing tactics inspired by BWT. This collaboration resulted in a community policing pilot program which launched in June 2012 in Detroit's Grandmont-Rosedale neighborhood. The pilot program consisted of three main components:

> A focus on individuals who commit home invasions; an increase in what is known as the "felt presence" of police by having officers proactively engage citizens to fix Detroit's equivalent of "broken windows"; leveraging the community as the eyes and ears to report suspicious/criminal activity. (Detroit Public Safety Foundation, 2013)

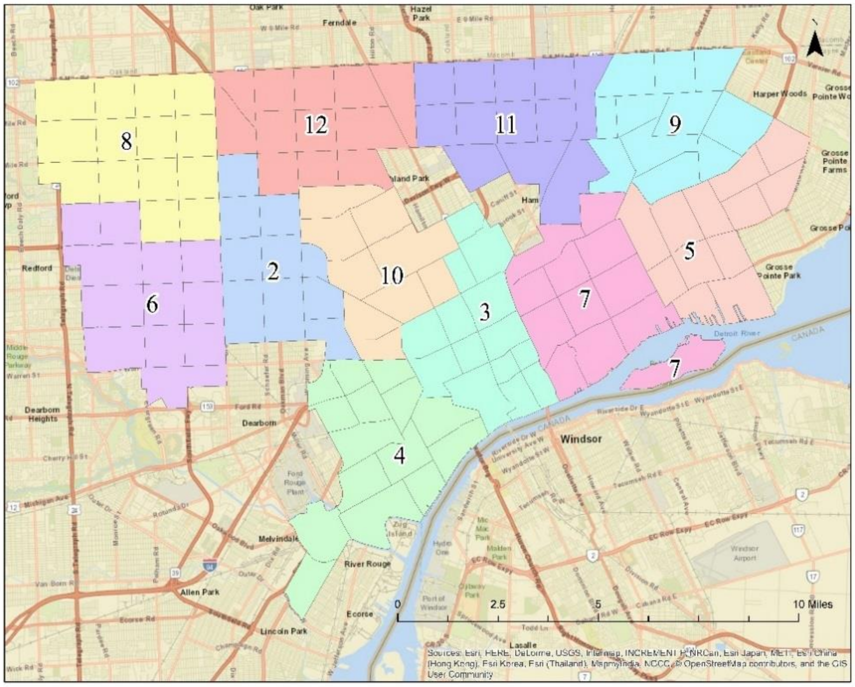
United by a shared purpose of creating a safer community, the DPD forged partnerships with residents and business-owners within Grandmont-Rosedale, as well as the criminal courts, Wayne County Sheriff Department, Michigan Department of Corrections, Greater Detroit Centers for Working Families, and Detroit Public Safety Foundation (Detroit Public Safety Foundation, 2013).

During the year-long pilot program, the DPD made over 1,200 proactive contacts with residents and conducted home visits with individuals who were previously arrested for serious crimes (Detroit Public Safety Foundation, 2013). Due to its low density, foot patrol was not implemented in Grandmont-Rosedale.[5] At the completion of the pilot program in June 2013, the DPD announced a 26% reduction in home invasions (Detroit Public Safety Foundation, 2013). Kelling - a senior fellow at the Manhattan Institute - announced, "The results demonstrate that if you increase the felt presence of police and conduct proactive outreach, the police and community together can prevent crime" (Detroit Public Safety Foundation, 2013, p.1).

Following the success of the pilot program, the city hired a new police chief: Chief James R. Craig. Soon after his arrival to Detroit, Chief Craig launched the Neighborhood Police Officers program, a comprehensive strategy aimed at improving communication and collaboration between the police, residents, and local businesses in an effort to create safer neighborhoods (City of Detroit, 2020). This program is currently on-going. For each precinct, three to five officers are designated to serve in the long-term position of neighborhood police officer (NPO) (City of Detroit, 2020). On average, an NPO is responsible for two scout car areas (SCAs), consisting of an area of approximately 2.09 square-miles. Together, the long-term nature of the position and responsibility for a smaller, more manageable geographic area enables NPOs to become more familiar with the community dynamics of the SCAs to which they are assigned (see Figure 6).

[5] In 2013, the author interviewed George Kelling regarding Detroit's community policing pilot program in Grandmont-Rosedale. He stated that foot patrol was not appropriate given the area's low density.

Figure 6. Scout Car Areas within Precincts.



Furthermore, NPOs play a non-adversarial role within their assigned SCAs, primarily addressing non-emergency and quality of life issues. Aligned with this role, NPOs receive additional training aimed at promoting positive interactions between the police and the community. Importantly, the program seeks to increase one-to-one contact with NPOs in settings outside crime in an effort to develop two-way relationships of trust with the community. To this end, NPOs are provided personal cellphones to communicate directly with residents and business-owners, and are encouraged to engage in playful interactions with youth and attend community events. Residents and business-owners are also provided the opportunity to meet NPOs at monthly Community Relations Council meetings in which they can raise issues for discussion. Thus, while NPOs still rely on their patrol vehicles to get from place-to-place, it does not hinder them from directly interacting with individuals within their assigned SCAs.

The initiation of the NPO program coincided with the resurrection of Detroit's 311 service request program, which had ceased operations on June 30th, 2012. Almost two years later, the program was rebranded "Improve Detroit." This program is currently on-going. The development of a mobile application (app) and online reporting system provide alternative ways to report non-emergency issues related to the physical environment (e.g., abandoned vehicles, potholes, and illegal dumping). These upgrades provide easier, more streamlined alternatives to report and track issues. By downloading the mobile application onto their cellphones, NPOs can easily report issues as they come to their attention on assignment. Currently, NPOs are amongst the most active users of the Improve Detroit app.

Meanwhile, Detroit's downtown was also undergoing several changes of its own. In an effort to improve its appearance and safety, the city invested in the installation of more and better street lighting, beautification efforts (e.g., planting trees, plants, and flowers), and property development. In addition, the Downtown Detroit Partnership (DDP) combined the efforts of the DPD and more than 20 businesses to maintain order. In particular, the DPD focused its efforts on "increasing the uniformed presence of officers on downtown streets, and the perceptions and realizations of public order created by clean streets and sidewalks, well-maintained landscaping, public spaces, and streetscape elements" (Clean Downtown, 2013).

Empirical Support for the Effectiveness of Policing Disorder Strategies

A host of studies have evaluated the impact of policing disorder strategies on violent crime. Most controversial among them are those that sought to identify the contributions of order-maintenance policing to the crime drop in New York City during the 1990s. Controlling for a host of socio-demographic variables, Kelling and Sousa (2001) found a significant negative relationship between misdemeanor arrests – a proxy for order-maintenance policing activities - and violent crime. They interpreted these results as supporting order-maintenance policing and discrediting explanations that focus on the role of "root causes" of crime. However, in a re-analysis of the evidence Harcourt and Ludwig (2006) failed to find an association between misdemeanor arrests in New York City and violent crime. Several evaluations conducted in other cities have also failed to find evidence of a crime reduction effect associated with policing disorder strategies (e.g., Katz, Webb, & Schaefer, 2001; Pace, 2010; Weisburd, Hinkle, Famega, & Ready, 2011).

In a later study, Rosenfeld, Fornango and Rengifo (2007) addressed several limitations in both Kelling and Sousa (2001) and Harcourt and Ludwig's (2006) analyses, such as the failure to account for the effects of spatial autocorrelation, and simultaneity between order-maintenance policing and serious crimes. Their re-analysis suggests that order-maintenance policing contributed to small but significant declines in homicide and robbery in New York City. However, unlike Kelling and Sousa (2001), they found several root causes of crime, such as low socio-economic status, racial composition, and immigrant concentration, to have a positive and significant effect on crime (Harcourt & Ludwig, 2006). In another study, Messner et al. (2007) found misdemeanor arrests to be associated with significant reductions in homicide rates, with the greatest impact on gun homicide rates. Several evaluations conducted in other cities also lend support to the ability of policing disorder strategies to produce significant crime reduction gains (e.g., Braga et al., 1999; McGarrell et al., 1999; Braga & Bond, 2008; Berk & MacDonald, 2010)

Ultimately, early evaluations of New York City's crime drop during the 1990s and evaluations conducted elsewhere provided no clearer understanding of the effectiveness or significance of policing disorder strategies. In the wake of these mixed evaluations, Braga et al. (2015) conducted a systematic review of published and unpublished empirical evidence on the effectiveness of policing disorder strategies. This review consisted of 30 randomized experimental and quasi-experimental evaluations (Braga et al., 2015). Using meta-analytical techniques, Braga et al. (2015) found that policing disorder strategies had a significant modest effect on crime reduction. The strategies that had the greatest impact were those that utilized community and problem-solving interventions consistent with the central tenants of order-maintenance policing, while aggressive strategies had no significant effect (Braga et al., 2015).

While Braga et al.'s (2015) review sheds light on the effectiveness of policing disorder strategies, it tells us nothing about the validity of BWT. More specifically, it tells us nothing about whether the crime control gains associated with policing disorder strategies are achieved by disrupting the cycle of disorder and decline described by Wilson and Kelling (1982). The observed crime control gains may be partially or wholly achieved by competing theoretical mechanisms. To the extent that this is true, BWT would fail to provide a unique and valuable framework to the field of Criminology.

In a follow up review, Weisburd et al. (2015) addressed this issue, casting doubt on Braga et al.'s (2015) findings. They argue that if policing disorder strategies indeed disrupt the broken windows process, then they should be associated with significant reductions in fear of crime (Weisburd et al., 2015). Using meta-analytical techniques, they evaluated six studies on the effect of policing disorder strategies on fear (Weisburd et al., 2015). Overall, Weisburd et al. (2015) failed to find evidence to suggest that policing disorder strategies yield significant reductions in fear, and one evaluation on its effect on collective efficacy also found no significant impact. They conclude that "the evidence does not indicate that broken windows policing mechanisms are behind the crime control gains of disorder policing programs observed by Braga et al. (2015)" (Weisburd et al., 2015, p. 598). Instead, Weisburd et al. (2015) argue that the mechanisms underlying criminal opportunity theories may help to better explain the crime control gains observed in policing disorder evaluations.

Implications for the Broken Windows Tipping Point

Four separate but related questions emerge from the review provided on the broken windows tipping point and policing disorder strategies. First, how may the tipping point impact the effectiveness of policing disorder strategies? Second, what, if anything, do evaluations of the effectiveness of policing disorder strategies tell us about the location of the broken windows tipping point? Third, what implications does the tipping point have for the allocation of police resources? Fourth, what does it mean for efforts to identify the tipping point if policing disorder strategies are not found to disrupt the broken windows cycle?

> Q1: How may the broken windows tipping point impact the effectiveness of policing disorder strategies?

As previously mentioned, evaluations of the effectiveness of policing disorder strategies fail to heed Wilson and Kelling's (1982) key instruction: to identify neighborhoods at the tipping point. The implications of this failure are multifaceted. Broadly speaking, the nature of the relationship between disorder and violent crime encourages certain expectations regarding the potential crime control gains of policing disorder strategies. Without this knowledge, we are unable to accurately judge the effectiveness of policing disorder strategies. We are also unable to know how best to direct police resources. For these reasons, it is very likely that policing disorder strategies have been directed to the wrong locations, and have, as a result, not realized their full potential. When the tipping point is considered, Wilson and Kelling's (1982) simple instruction to police - stop small problems before they become much larger – is not so simple after all.

Q2: What, if anything, do evaluations of the effectiveness of policing disorder strategies tell us about the location of the broken windows tipping point?

Ultimately, it is unclear what the empirical findings from evaluations of the effectiveness of policing disorder strategies tell us about the tipping point. To start, evaluations that find no effect on crime do not suggest that policing disorder strategies were implemented in the wrong neighborhoods (i.e., neighborhoods that were not located at the tipping point). As will later be discussed, there are several issues associated with evaluations of policing disorder strategies that may lead researchers to incorrectly conclude that they have no effect on crime. In addition to these issues, research has shown that neighborhood context plays a critical role in the effectiveness of policing strategies (Kelling & Coles, 1996; Kane & Cronin, 2009). Recall a goal of order-maintenance policing is to aid residents in regaining control over their communities. Strategic police-community partnerships are instrumental to achieving this goal; they require that residents trust in the police and are committed to police efforts to improve their neighborhood. However, certain neighborhood conditions may undercut the effectiveness of these partnerships or prevent them from occurring altogether. For example, deeply embedded negative attitudes toward the police may prove to be an insurmountable obstacle toward establishing police-community partnerships. In fact, research has shown that individuals who harbor negative attitudes toward the police are less willing to utilize formal mechanisms of social control (Scaglion & Condon, 1980; Dunham & Alpert, 1988; Silver & Miller, 2004). Research has also shown that residents' neighborhood attachments predict their willingness to collectively engage in informal social control, as well as partner with the police (Silver & Miller, 2004; Long & Perkins, 2007). Thus, in neighborhoods in which neighborhood attachments are low, such as in highly transient neighborhoods, the potential of order-maintenance policing to drive positive neighborhood change may not be fully realized.

Furthermore, empirical findings which indicate that policing disorder strategies *do* reduce crime also provide no insight on whether they were implemented in neighborhoods at the tipping point. This is because Wilson and Kelling (1982) never argue that to have an effect on crime policing disorder strategies must be implemented in such neighborhoods. Instead, they argue that the best use of limited police resources is to target neighborhoods at the tipping point (Wilson and Kelling, 1982). That being said, how we interpret Wilson and Kelling's (1982) description of the tipping point (and the nature of the relationship between disorder and violent crime it suggests) sets up certain expectations for the effectiveness of order-maintenance policing initiatives.

Recall the two competing interpretations of Wilson and Kelling's (1982) description of the broken windows tipping point depicted by the line segments $\overline{ABCD}$ and $\overline{ABD}$ in Figure 1 and replicated in Figure 7. In neighborhoods at or leading up to the tipping point, efforts to decrease disorder should not produce very large crime reduction gains (see Zone 1 depicted in Figure 7). What distinguishes these neighborhoods are their (1) ability to address disorder via informal social control and (2) risk of inflaming the broken windows developmental sequence. In neighborhoods located at the tipping point, informal social control is faltering. Without intervention, it is unable to return the neighborhoods back to a low disorder, low crime state. Given their position, these neighborhoods are at great risk of being propelled into a high disorder, high crime state.

Policing disorder strategies implemented in neighborhoods at the tipping point are oriented towards *preventing* disorder from extending beyond tipping point levels, resulting in an uptick in violent crime. Police intervention should be *minimal*, just enough to supplement and restore informal social control within neighborhoods. If possible, however, citizen action should be the primary mechanism through which disorder is addressed (Wilson & Kelling, 1982; Kelling & Coles, 1996). To this point, Wilson and Kelling (1982, para. 44) argue that "[e]ven in areas that are in jeopardy from disorderly elements, citizen action without substantial police involvement may be sufficient." They provide some examples of what citizen action may entail:

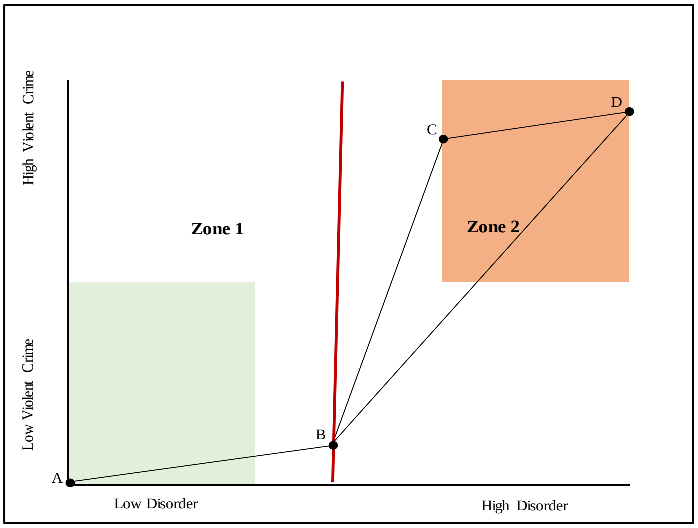
> Meetings between teenagers who like to hang out on a particular corner and adults who want to use that corner might well lead to an amicable agreement on a set of rules about how many people can be allowed to congregate, where, and when. Where no understanding is possible—or if possible, not observed—citizen patrols may be a sufficient response. (Wilson & Kelling, 1982, para. 44-45)

Beyond the tipping point, however, the primary purpose of policing disorder strategies is geared towards returning disorder to pre-tipping point levels (see Zone 2 depicted in Figure 7). In other words, policing disorder strategies are oriented toward achieving significant crime reduction gains. However, our expectations regarding whether we think this goal can be easily achieved depends, in part, on which interpretation of the nature of the relationship between disorder and violent crime we place stock in.

Wilson and Kelling's (1982) description of the broken windows tipping point can be understood as a statement regarding the appropriate dosage of police response needed to address disorder in neighborhoods located beyond the tipping point. According to their description, we should not expect policing disorder strategies to elicit significant crime reduction gains if they are implemented with limited police resources (Wilson & Kelling, 1982). In such a scenario, Wilson and Kelling (1982) argue that the best police can do is to respond to calls for service. That being said, there is another option that Wilson and Kelling (1982) failed to consider: police resources can be directed to small geographic areas that contain heightened levels of disorder and/or crime (i.e., hot spots). In this scenario, a high *spatial* dosage of policing disorder activities can be achieved with limited police resources (Trojanowicz, 1986; Ratcliffe et al., 2011).

Therefore, assuming the proper dosage of police response was utilized, we should expect policing disorder strategies directed to neighborhoods that fall along the $\overline{BC}$ or $\overline{BD}$ line segments depicted in Figure 7 to return significant crime reduction gains given the strong relationship between disorder and violent crime in such places. Alternatively, Wilson and Kelling's (1982) description of the tipping point can be understood as an assessment of the strength of the relationship between disorder and violent crime in high disorder, high crime neighborhoods. In such neighborhoods, the relationship between disorder and violent crime is modest. For this reason, beyond point $\dot{C}$ efforts to address disorder will not return significant crime reduction gains. Upholding this competing interpretation, we should not expect significant crime reduction gains for neighborhoods that fall along the $\overline{CD}$ line segment depicted in Figure 7. Regardless of which interpretation we place stock in, the strength of police-community partnerships, as well as the degree of flexibility afforded by a neighborhood's crime trajectory and the systemic features that contribute to it are additional factors that reasonably affect the effort required to facilitate a neighborhood's transition from a high disorder, high crime state to low disorder, low crime state.

Figure 7. The Broken Windows Tipping Point: Crime Prevention vs. Crime Reduction.



There are several reasons to question the expectations described here. To start, policing disorder strategies that have been effective at reducing crime in hot spots may indicate that the tipping point is located further down the disorder distribution in "urban ghettos" (Braga & Bond, 2008; Braga et al., 1999; Braga et al., 2012, 2014). Further complicating matters, Geller's (2007) finding of a concave relationship in which the effect of disorder on violent crime is the strongest at low levels of disorder suggests that the broken windows tipping point may not resemble a threshold effect as suggested by Wilson and Kelling (1982). While much more research needs to be conducted on this phenomenon, this small handful of studies provide reason to critically re-consider Wilson and Kelling's (1982) interpretation of the tipping point.

Q3: What implications does the broken windows tipping point have for the allocation of police resources?

A prime metric by which to assess the success of a policing strategy is the effect it has on crime. To this point, Wilson and Kelling (1982, para. 51) suggest that "[w]e may have encouraged [the police] to suppose, however, on the basis of our oft-repeated concerns about serious, violent crime, that they will be judged exclusively on their capacity as crime-fighters." With this metric in mind, if the police want to have a large observable impact on crime, research suggests that it focus its efforts on crime hot spots (see Braga et al., 2019). However, BWT requires that this metric of success be reconsidered. In actuality, BWT suggests that policing disorder strategies should be directed to neighborhoods that are at the brink of decline: neighborhoods at the broken windows tipping point. As compared to neighborhoods located beyond the tipping point, these strategies are expected to achieve much smaller crime reduction gains given the nature of the disorder-crime relationship suggested by Wilson and Kelling (1982). Adding a layer of complexity, the identification of a valid treatment effect will be much harder in neighborhoods located at the tipping point than neighborhoods located beyond it given lower base rates of disorder and crime (see Hinkle et al., 2013).

Furthermore, policing disorder strategies focused on neighborhoods located at the tipping point are oriented toward preventing *future* increases in violent crime. This increase is anticipated by virtue of being at the tipping point. The logic being that future expenditures of police resources, as well as the consequences of violence for neighborhoods and their residents, can be avoided by providing a minimal police presence in neighborhoods located at the tipping point until neighborhood informal social control is able to re-establish and sustain public order unaided. Thus, the benefits of police efforts oriented toward *crime prevention*, defined as actions taken to prevent future crime emergence, are overlooked by gauging success primarily in terms of crime reduction gains, defined as the amount by which crime is reduced.

In an ideal scenario, the police should work to reduce overall crime incidents in neighborhoods, as well as prevent future crime emergence. Police efforts geared towards the former objective are relatively easy to assess. For example, a simple comparison of crime levels before and after a police intervention can be used to identify a treatment effect. However, it is exceedingly more difficult to identify a treatment effect for efforts geared toward the latter objective, as their aim is to avoid a future potentiality: crime that has not yet occurred. Thus, evaluations of these efforts require knowledge of what would have likely happened in the neighborhood had the police never intervened.

According to Wilson and Kelling (1982), the best use of limited police resources is to direct them to neighborhoods at the tipping point. However, the proven effectiveness of hot spot policing provides reason to question this approach. Therefore, a better approach is to conduct a comprehensive assessment of the costs and benefits associated with how resources should be allocated, differentiating policing strategies aimed at crime reduction from those aimed at crime prevention. To this point, Wilson and Kelling (1982) seemingly support such an assessment:

> But the most important requirement is to think that to maintain order in precarious situations is a vital job. The police know this is one of their functions, and they also believe, correctly, that it cannot be done to the exclusion of criminal investigation and responding to calls. (para. 51)

Factors worthy of consideration in such cost-benefit assessments might include the ability of the police strategy to improve residents' quality of life, reduce fear of crime, strengthen police-community relations, and avoid the financial and social costs that would likely accompany future violence. Of course, this short list of factors is far from complete. The identification of the tipping point will surely advance this list by facilitating the identification of appropriate performance outcomes associated with police efforts geared toward crime prevention at the tipping point. Ultimately, even with limited police resources one strategy need not be completely abandoned to support the other. Rather, resources should be differentially allocated. In such a scenario, neighborhoods that stand to experience the largest net benefit from a policing disorder strategy should be prioritized.

Q4: What does it mean for efforts to identify the broken windows tipping point if policing disorder strategies are not found to disrupt the broken windows cycle?

What remains to be discussed are the implications of efforts to identify the broken windows tipping point in the event that policing disorder strategies are not found to disrupt the broken windows cycle. As previously mentioned, Weisburd et al.'s (2015) review suggests that the mechanisms underlying criminal opportunity theories may better help explain the crime reduction gains observed in policing disorder evaluations. Importantly, this suggestion is motivated by their failure to find these strategies to have a significant effect on fear of crime and, in one case, collective efficacy (Weisburd et al., 2015). While our theoretical understanding of the mechanisms underlying the tipping point would no longer hold merit, the idea still holds relevance for efforts to address disorder. To elaborate, even in the extreme case in which policing disorder strategies have no roots in BWT it is still reasonable to direct limited police resources to neighborhoods located at the tipping point in an effort to avoid future crime emergence. While we may anticipate significant crime reduction gains if policing disorder strategies are implemented in high disorder, high crime neighborhoods, the dosage of police resources that would be needed to return them to a low disorder, low crime state in which residents are able to exercise informal social control would likely come at too high a cost. At the opposing extreme, policing disorder strategies would be inappropriate in low disorder, low crime neighborhoods for the simple reason that disorder and crime are not issues in these areas. For these reasons, the allocation of limited police resources to neighborhoods at the tipping point emerges as a completely defensible approach in order to avoid the future potentiality of increased violence. Nevertheless, this approach should be weighed in light of the benefits associated with policing efforts geared toward crime reduction.

Four critical issues of broken windows research impair evaluations of policing disorder strategies and may lead researchers to incorrectly conclude that they do not disrupt the broken windows cycle. First, the time-frame in which the broken windows developmental sequence is expected to unfold is unknown. Consequently, it is unclear how quickly the effects of order-maintenance policing are expected to impact fear of crime, eventually leading to a reduction in crime. That being said, it is reasonable to suspect that policing disorder activities should have a relatively immediate effect on potential offenders through deterrence-based processes. An increase in police activities should send a clear signal to potential offenders that criminal acts will likely be detected, which should in turn heighten their risk of apprehension and deter offending.

Second, it has been suggested that some policing disorder strategies may increase residents' fear of crime, commonly referred to as a "backfire effect" (Rosenbaum, 2006; Hinkle & Weisburd, 2008). In particular, the implementation of heightened police activities may signal to residents that disorder and/or crime has risen, triggering public levels of fear to rise. Thus, any reduction in fear of crime that had been achieved by addressing disorder will be diminished by increases in fear associated with the strategy itself. As a result, researchers may incorrectly conclude that policing disorder strategies have no effect on fear of crime. However, more recent research has questioned this phenomenon (see Weisburd et al., 2011; Ratcliffe et al., 2015). Ultimately, much more research needs to be conducted that explores this phenomenon across various target populations and crime levels, as well as types of hot spots and policing strategies.[6]

Third, the appropriate dosage of policing needed to effectively address the issue of disorder is unknown (Wilson & Kelling, 1982). This issue is not unique to order-maintenance policing and has been explored in applications of other policing strategies, such as hot spot policing, that are anchored in a deterrence-based understanding of crime (Kelling, 1974; Koper, 1995; Telep et al., 2014; Groff et al., 2015; Santos & Santos, 2015). As previously discussed, heightened policing disorder activities may be beneficial in terms of leveraging crime reduction gains through deterrence-based pathways. However, it has been suggested that these activities may inspire fear within residents, which in turn may prevent the crime control gains associated with disrupting the broken windows cycle from being fully realized. Thus, a key issue is whether

[6] As previously mentioned, aggressive tactics to address disorder, such as those employed in zero-tolerance, violate the central tenants of order-maintenance policing as described by Wilson and Kelling (1982). For this reason, the implications of these tactics on fear of crime were excluded from this discussion.

there exists an appropriate dosage of order-maintenance policing activities that is able to jointly harness deterrence-based and broken windows processes to produce maximum crime control benefits.

Last, another reason to consider why policing disorder strategies may have little effect on fear of crime is that such strategies may be implemented in places in which the disorder-fear connection is weak or non-existent. To elaborate, research suggests that the disorder-fear connection may be weak or non-existent in neighborhoods in which disorder is either low or high (Taylor, Shumaker, & Gottfredson, 1985; Taylor & Shumaker, 1990; Innes, 2004; Millie, 2008; Sampson & Raudenbush, 2004). In low disorder neighborhoods, residents are unlikely to perceive disorder as a problem. In high disorder neighborhoods, residents may become inoculated to its presence. The explanation for why the disorder-fear connection may be weak or non-existent in both neighborhood types is the same: if residents are unaware or unbothered by the presence of disorder, then it will likely not result in a fear response and the process of neighborhood decline - as hypothesized by Wilson and Kelling (1982) – will not be spurred. For this reason, in such neighborhoods it would be unreasonable to expect policing disorder strategies to reduce neighborhood levels of fear by addressing disorder.

In summary, evaluations of policing disorder strategies are unable to provide us with a clear understanding of the location of the broken windows tipping point, or the extent to which the crime control benefits associated with these strategies are driven by broken windows or alternative processes. The lack of attention that has been given to the tipping point is extremely surprising given its potential impact on the effectiveness of policing disorder strategies. Importantly, research that seeks to identify the location of the tipping point not only has significant implications for the effectiveness of police operations that address disorder, but also for evaluations that examine the theoretical mechanisms underlying such strategies.

Moving forward, it is the purpose of this study to empirically examine the relationship between disorder and violent crime rate as a *first step* toward identifying the broken windows tipping point. Research that seeks to identify the tipping point should not implicitly assume that disorder maintains a linear relationship with violent crime, nor should it take Wilson and Kelling's (1982) description of the tipping point at face-value. To this point, the present study makes great strides to advance research on the tipping point by adopting a methodological approach that allows for flexibility in modeling decisions.

Current Study

This study empirically assesses the validity of four hypothesized functional forms of the relationship between physical disorder and violent crime rate that emerged from its literature review. In the absence of detail provided by Wilson and Kelling (1982), disorder has been hypothesized to have a positive linear effect on violent crime (H1). However, Wilson and Kelling's (1982) description of the tipping point implies a threshold effect of disorder on violent crime. Importantly, their description gives rise to two competing interpretations of this relationship (Wilson & Kelling, 1982). First, disorder maintains a threshold effect on violent crime such that small variations in disorder exert a modest positive effect on violent crime at low levels of disorder, and a dramatic positive effect past a critical level located somewhere between low and high levels (H2a). Second, disorder maintains a threshold effect on violent crime such that small variations in disorder exert a modest positive effect on violent crime at low and high levels of disorder, and a dramatic positive effect past a critical level located somewhere between these extremes (H2b). Alternatively, Crane's (1991) epidemic theory of ghettos – supported by

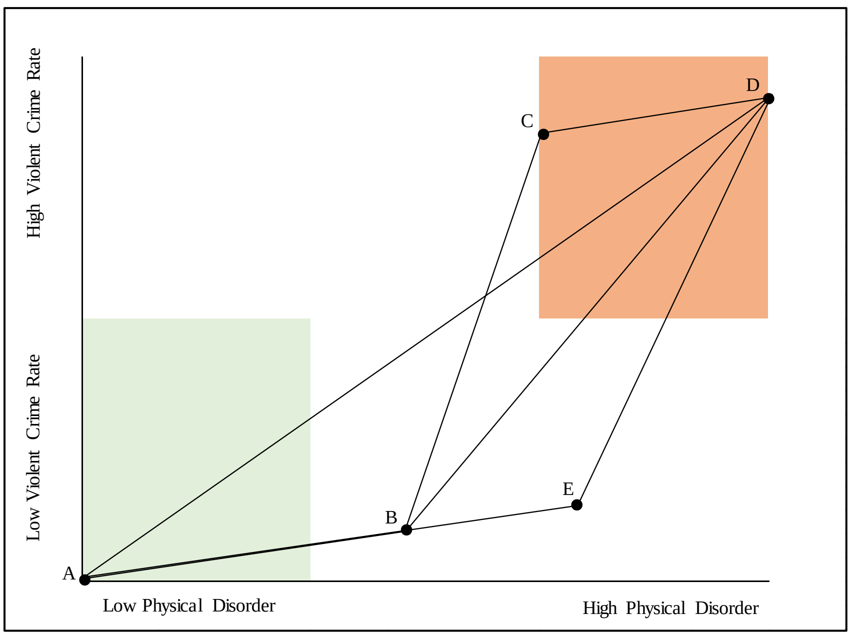
Raleigh and Galster (2015) - pushes the tipping point toward the end of the disorder distribution. In particular, it suggests that small variations in disorder exert a modest positive effect on violent crime at low and mid-range levels of disorder, and a dramatic positive effect past a critical level located somewhere at high levels of disorder (H3). As seen in Figure 8, hypotheses H1, H2a, H2b, and H3 are captured by the $\overline{AD}$, $\overline{ABD}$, $\overline{ABCD}$, and $\overline{AED}$ line segments, respectively. Furthermore, there is also reason to believe that disorder may not exhibit a threshold effect on violent crime, but rather a nonlinear effect that may take one of many forms, such as the concave relationship between disorder and violent crime rate identified by Geller (2007) (H4). These hypotheses are adjusted to reflect this study's focus on the effect of *physical disorder* on *violent crime rate*, and are explicitly stated in Table 1.

Table 1. Key Research Question & Hypotheses.

Research Question: What is the functional form of the relationship between physical disorder and violent crime rate?
H1: Physical disorder maintains a positive linear effect on violent crime rate.
H2a: Physical disorder maintains a threshold effect on violent crime rate such that small variations exert a modest positive effect on violent crime rate at low levels, and a dramatic positive effect past a critical level located somewhere between low and high levels.
H2b: Physical disorder maintains a threshold effect on violent crime rate such that small variations exert a modest positive effect on violent crime rate at low and high levels, and a dramatic positive effect past a critical level located somewhere between these two extremes.
H3: Physical disorder maintains a threshold effect on violent crime rate such that small variations exert a modest positive effect on violent crime rate at low and mid-range levels, and a dramatic positive effect past a critical level located somewhere at high levels.
H4: Physical disorder maintains a nonlinear effect on violent crime rate.

To assess these hypotheses, this study utilizes a dose-response propensity score method, an appropriate and rigorous evaluation design that minimizes concerns about selection bias and allows for causal inferences. This approach estimates the average treatment effect of various levels of physical disorder, measured at t_1 = 2014, on violent crime rate, measured at t_2 = 2015, explicitly models the functional form of these variables, and allows for covariate balancing across matched levels of physical disorder. Importantly, this method is well-suited for the identification of tipping points because it allows for nonlinear threshold effects to be estimated. To facilitate its analysis, this study utilizes block-group level data on physical disorder, violent crime, as well as socioeconomic and land use characteristics from the DPD's record management system, MCM project, and Census. These sources provide the data necessary to create theoretically relevant variables of key interest to this study.

Figure 8. Hypotheses H1, H2a, H2b, and H3.

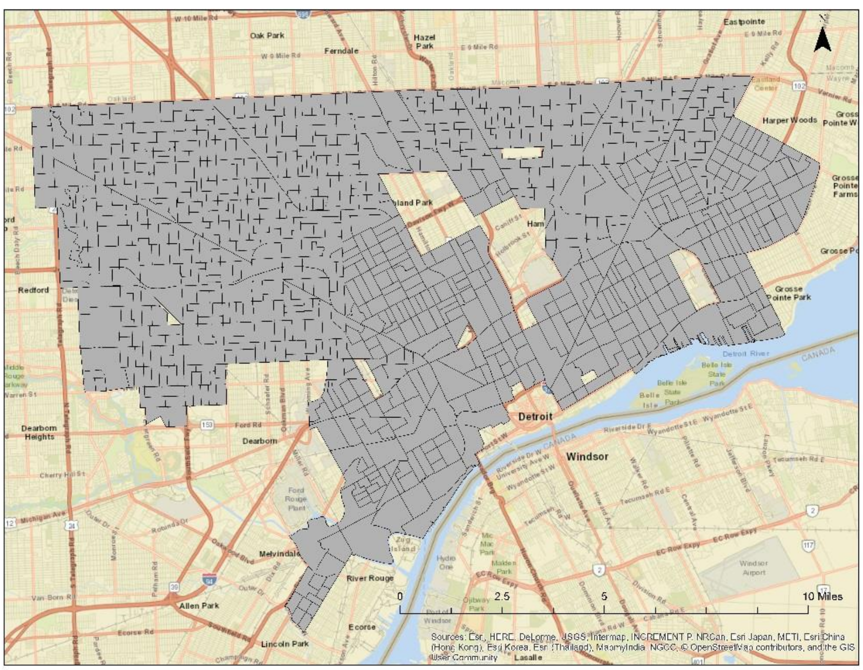


CHAPTER 3: STUDY DESIGN AND IMPLEMENTATION

Measures and Data Sources

This study collected data on physical disorder, violent crime, as well as socioeconomic and land use characteristics from the DPD's record management system, MCM project, and Census. These data were aggregated to the block-group level, a common unit of analysis in research on the relationship between disorder and crime (see Sampson & Raudenbush, 1999, 2004; Yang, 2010; O'Brien et al., 2015; O'Brien & Sampson, 2015; Wheeler, 2018). Reflecting its focus on neighborhoods, this study excluded block-groups that fell within Detroit's downtown area, contained no population, and/or did not contain properties zoned for residential use. This selection procedure excluded 22 block-groups, resulting in a total sample size of 857 (see Figure 8).

Figure 9. Selected Block-groups.



Physical Disorder. A traditional method to capture disorder in neighborhoods is through physical audits of the environment, referred to as systematic social observation (SSO). SSO is an appealing approach to measuring disorder because it relies on independent and structured observations of the environment by trained surveyors. As a result, it avoids many measurement issues common to alternative approaches, such as community surveys and 311 service requests. To elaborate, community survey measures of disorder may be compromised by residents' inability to distinguish between disorder and crime (Gau & Pratt, 2008, 2010), and have been found to be affected by individual- and neighborhood-level characteristics (Taylor, Shumaker, & Gottfredson, 1985; Sampson, 2009, 2012; Sampson & Raudenbush, 2004; Wickes, et al., 2013) As self-reported data, 311 service requests are affected by both under- and over-reporting (Sherman, Gartin, & Buerger, 1989; Klinger & Bridges, 1997). These data may also be biased if residents systematically differ in their likelihood to request city services (O'Brien, 2015; O'Brien & Sampson, 2015; O'Brien, Sampson, & Winship, 2015; White & Trump, 2016).

The use of SSO to measure disorder is not without its limitations. To start, it is an extremely timely and costly approach. For these reasons alone, SSO may be out of reach for economically-strained communities. Virtual audits of the environment - made possible by geospatial technologies like Google Street View - are a less-costly alternative to SSO. However, much more research is needed to assess the reliability and validity of this method for capturing disorder (e.g., Clarke et al., 2010; Rundle et al., 2011; Odgers et al., 2012; Mooney et al., 2016). Furthermore, it is well known that SSO is likely to vary depending on weather conditions, time of day, and day of the week, and is also dependent on unity in inter-rater reliability (see Skogan, 2012, 2015).

Perhaps the greatest limitation of SSO involves the argument that disorder is a *social construct*, rather than an objective condition that is similarly perceived across individuals (Harcout, 2001; Sampson & Raudenbush, 2004; Hinkle & Yang, 2014). Indeed, research suggests that what residents perceive as disorder may not align with how disorder is captured by outsiders, such as those conducting SSOs (e.g., Perkins et al., 1993; Franzini et al., 2008; Hinkle & Yang, 2014). Perceptions of disorder play a critical role in the broken windows developmental sequence. If residents do not perceive disorder to be a problem in their neighborhoods, then this sequence will not unfold. From a purely theoretical standpoint, perceived disorder is the most appropriate measure to assess the validity of BWT. That being said, research finds considerable consistency between observed and perceived measures of *physical disorder*, the focus of this study (Perkins, Meeks, & Taylor 1992; Sampson & Raudenbush, 2004; Hinkle & Yang, 2014; Yang & Pao, 2015; Ren, Zhao, & He, 2017). Together, these findings suggest that physical disorder is more uniformly interpreted by residents as signaling neighborhood decline, resulting in higher levels of perceived disorder and fear. For this reason, SSO emerges as an appropriate approach to capture physical disorder.

In 2013, Detroit's Blight Removal Task Force - in partnership with Michigan Nonprofit Association, Data Driven Detroit, and Loveland Technologies - developed a survey to capture the physical condition of every land parcel within Detroit. The motivation behind the city collaborative – termed the Motor City Mapping (MCM) project - was to create a comprehensive, crowd-sourced database in an effort to identify problem properties and track them over time. Surveyors were recruited from within the city to facilitate this initiative. As a result of this recruitment strategy, the surveyors possessed detailed knowledge on Detroit neighborhoods and the city as a whole. Once recruited, they received extensive training to ensure a comprehensive understanding of survey items and definitions. Additionally, the surveyors received training on how to use a mobile app: "Blexting." This app was created to record property conditions and was downloaded on Nexus 7 tablets which were provided to each surveyor. Over a 10-week period, teams of surveyors were assigned to micro-hoods – 0.25 square-mile areas - to conduct physical audits of the entire city utilizing the Blexting app. For each property, the surveyors took photographs and responded to a series of survey items.

Furthermore, a mission control center was established where staff performed quality checks of the data submitted by the surveyors in real-time. Data collection was completed in the winter of 2014. The parcel-level data was later aggregated to the block-group level and made available through the city of Detroit's open data portal. Among the available data, seven survey items are particularly well-suited for the current study. A description of each item - taken from the MCM project codebook - is provided below (see Motor City Mapping, 2020). Summary statistics are provided in Table 2.

- Percent Poor Condition – Number of parcels with structures that are in poor condition divided by the number of parcels surveyed with structures. Structures that are in poor condition need major repairs. Their windows and doors may be broken or boarded. They may also have light fire damage that can be repaired. Other indicators of structures in poor condition include damaged, non-load-bearing elements like awnings, or porches collapsed, and damaged roof.

- Percent Suggested Demolition – Number of parcels with structures that are suggested for demolition divided by the number of parcels surveyed with structures. Structures that are suggested for demolition include structures that are no longer shaped like a building. They are damaged beyond practical repair or renovation, and are uninhabitable.

- Percent Structure Unoccupied – Number of parcels with structures that are perceived to be unoccupied divided by the number of parcels surveyed with residential structures. Common characteristics of unoccupied structures include neglected facades, eviction notices, empty interiors, substantial physical or structural damages, extensive security measures, uncut or tall grass, weeds, scrub trees, trash or debris accumulated over time, or accumulated flyers on the porch or door.

- Percent Structure Need Boarding – Number of parcels with structures that are in need of boarding divided by the number of parcels surveyed with structures. A structure is in need of boarding if it has missing windows, doors or is otherwise open and accessible to scrappers, squatters, or vandals.

- Percent Structure Fire Damaged – Number of parcels with structures that are fire-damaged divided by the number of parcels surveyed with structures. A structure is classified as having fire damage if it has visible indicators of fire damage in or around it, from as small as melted siding to structures that have burned down to the ground.

- Percent Total Parcels Dumping – Number of parcels, with or without structures, that have dumping divided by the number of parcels surveyed. A building or vacant lot is considered to have dumping when debris has been purposely left or placed on the property. This does not include litter or debris from a recent fire or ongoing demolition

- Percent Lots Unmaintained – Number of parcels without structures that are unmaintained divided by the number of parcels without structures surveyed. Characteristics of an unmaintained lot include tall grass, overgrown trees or bushes, weeds in the cracks of pavement, and so on.

Table 2. Physical Disorder Summary Statistics.

Variable	Mean	Std. Dev.	Min	Max
Percent Poor Condition	3.57	4.00	0	17.89
Percent Suggested Demolition	1.86	2.64	0	17.52
Percent Structure Unoccupied	19.32	12.50	0	59.33
Percent Structure Need Boarding	11.34	9.58	0	50.00
Percent Structure Fire Damage	2.79	2.99	0	20.18
Percent Total Parcels Dumping	2.38	2.67	0	24.84
Percent Lots Unmaintained	46.65	23.76	0	100.00

A principal component factor analysis revealed that all items loaded strongly onto a single factor (see Table 3). For this reason, a single composite measure representing physical disorder was generated from a regression-weighted scale of constituent characteristics and adjusted so that all values were positive ($\bar{X}$ = 1.37, SD = 1.00, Min = 0, Max = 4.76). Figure 9 visually captures this measure by block-group, with classifications based upon natural-breaks. As can be seen, low levels of physical disorder (indicated by blue tones) are predominantly concentrated in north-west Detroit, while high levels (indicated by red tones) are concentrated in several areas, most notably in west, central, and north-east Detroit. Furthermore, Figure 10 displays the distribution of physical disorder captured as a density. The data reveal a moderate right skew.

Table 3. Principal Component Factor Analysis.

Physical Disorder	Factor Loading
Percent Poor Condition	0.86
Percent Suggested Demolition	0.82
Percent Structure Unoccupied	0.91
Percent Structure Need Boarding	0.93
Percent Structure Fire Damage	0.87
Percent Total Parcels Dumping	0.73
Percent Lots Unmaintained	0.75

Note: α = 0.78. For factor analysis, N = 857. A principal component factor estimation was used with no rotation.

Figure 10. Physical Disorder.

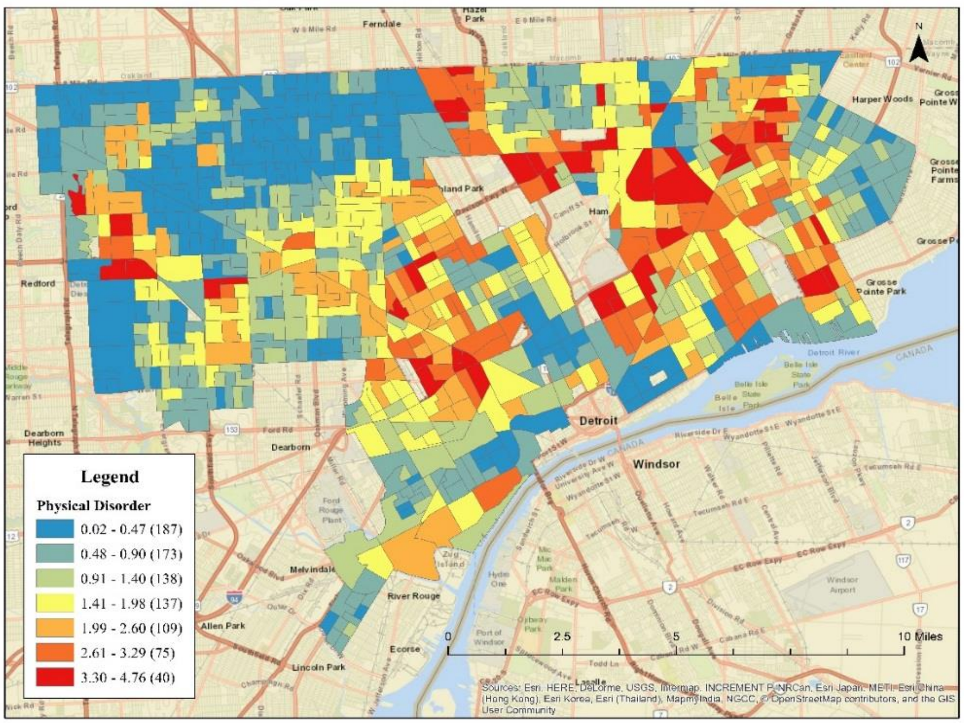
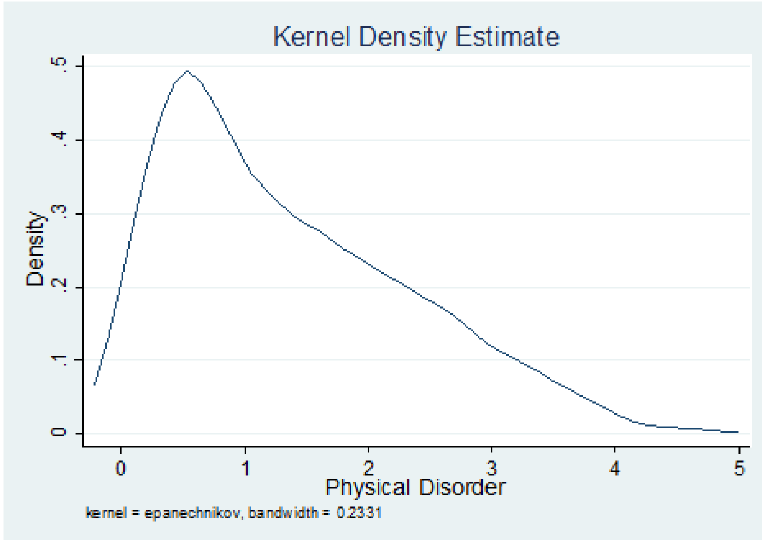


Figure 11. Physical Disorder: Kernel Density Estimate.



Violent Crime Rate. Broken windows research customarily uses either incident or CFS data as a measure of violent crime (see Braga et al., 2015). However, CFS may contain unsubstantiated crimes and are often considered as a proxy of residents' *perceptions* of crime levels (see Hinkle & Weisburd, 2008). To this point, it can be argued that BWT does not place emphasis on residents' perceptions of violence, but rather substantiated incidents thereof. Furthermore, CFS tend to be a less accurate record of crime types (Klinger & Bridges, 1997). For these reasons, this study collects violent crime incidents (homicide, rape, robbery, and aggravated assaults) that occurred in 2015 from the DPD's record management system to construct its dependent variable. Importantly, this study's focus on violent crime is not only consistent with BWT, but also helped ensure that its dependent variable is conceptually distinct from disorder (Weisburd et al., 2015). These data were geocoded in ArcMap (version 10.8), aggregated to the block-group level, and recorded as a rate (per 1,000 people) using five-year population estimates obtained from the American Community Survey (ACS) ($\bar{X}$= 22.18, SD = 19.09, Min = 1.71, Max = 277.37).

Figure 11 visually captures violent crime rate by block-group, with classifications based upon natural-breaks. As can be seen, high levels of violent crime rate (indicated by red tones) are predominantly concentrated in central and north-east Detroit, while low levels (indicated by blue tones) are concentrated in several areas, most notably in north-east and south Detroit. Furthermore, Figure 12 displays the distribution of violent crime rate captured as a density. An extreme right-skew is evident.

Figure 12. Violent Crime Rate (2015).

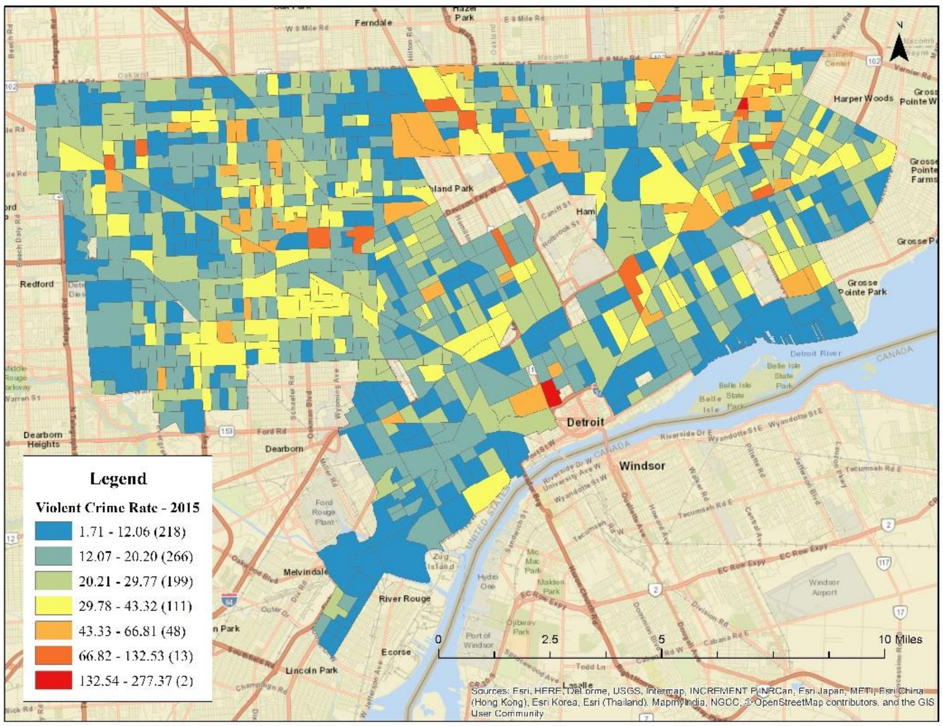
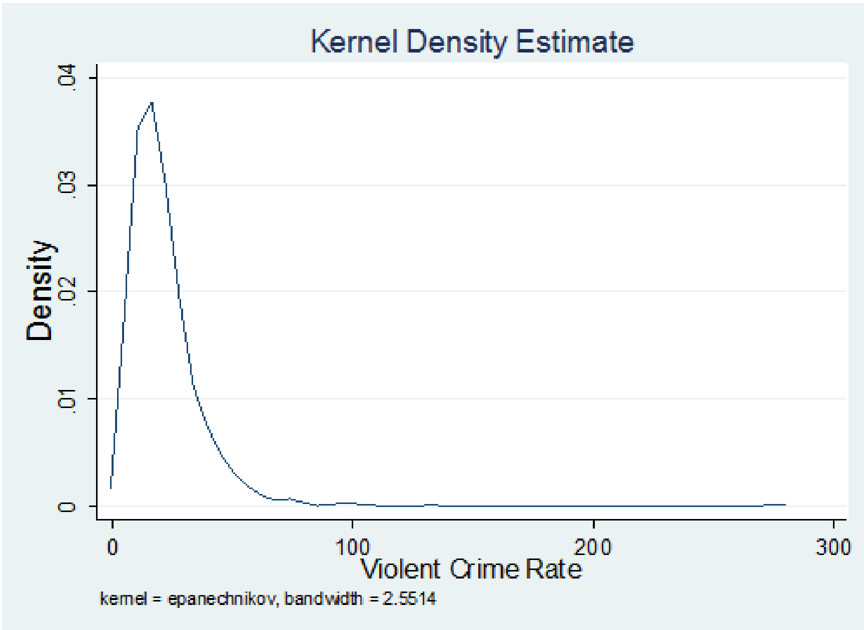


Figure 13. Violent Crime Rate (2015): Kernel Density Estimate.



Control/Matching Variables. This study constructed land use and socioeconomic variables from data sources compiled in 2014 (see Table 4). The identification of these variables was informed by the previous review of broken windows and related research. Given their presumed correlation with physical disorder, the omission of these variables from this study's analysis would result in omitted variable bias. This consequence has severe implications for the ability of the GPS approach to draw causal inferences, a point of later discussion.

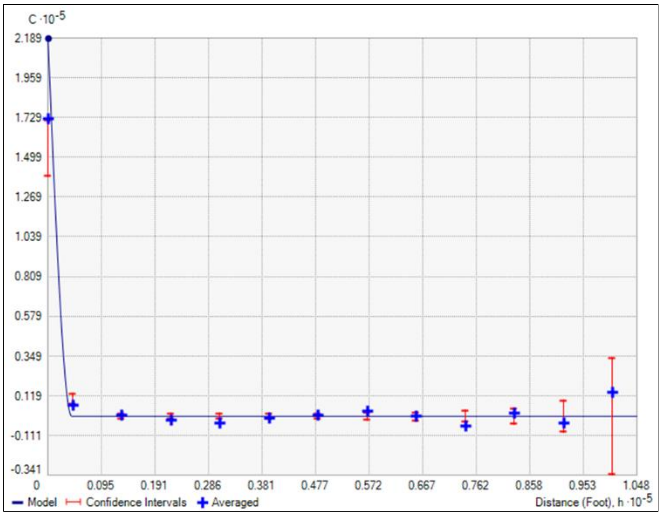
Data on land use characteristics were obtained from the MCM project and constructed as a proportion of all parcels. The measures constructed include percent garden/park, percent commercial, percent residential, percent industrial, and percent mixed. Furthermore, socioeconomic variables were constructed using five-year estimates obtained from the ACS and DPD's record management system. These variables include measures of population density (measured per square-mile), percent male between the ages of 15 and 24, percent population under the age of 18, percent unemployed, percent receiving public assistance, percent female-headed households, percent owner-occupied homes, percent same residence (for at least one year), percent African American, percent Hispanic/Latino origin, percent white, percent foreign-born, and violent crime rate.[7]

A measure of population growth was constructed using population estimates from the 2000 Census. Block-group boundaries in 2014 do not perfectly match those used over ten years ago. For this reason, areal interpolation was necessary in order to re-aggregate the data to the block-group boundaries used in 2014.[8] This was achieved using the "Areal Interpolation" tool available in ArcMap. The process first begins with the construction of a valid variograph model from which population estimates are calculated. According to the ESRI (2020) user-guide, 90% of the empirical covariances (blue cross) should fall within the (red) confidence intervals (see Figure 13). Furthermore, the root-mean-square standardized value should also be close to one (RMSE = 1.10). The constructed model met both of these criteria. The re-aggregated population estimates were then used to calculate a measure of population growth, defined as

$$\frac{Population\ \ in\ 2014 - Population\ in\ 2000}{Population\ in\ \ 2000}\ x\ 100.$$

[7] A measure of concentrated disadvantage, consisting of percent unemployed, percent receiving public assistance, percent female-headed household, percent African American, and percent population under 18, could not be constructed due to low correlation between its constituent characteristics (α = 0.39).
[8] Areal interpolation is a broad term used to include a set of methods that "can estimate an aggregate attribute of one areal unit system based on that of another, spatially incongruent, system in which the attribute data were collected" (Qiu, Zhang, & Zhou, 2012, p. 645).

Figure 14. Variograph Model.



In addition, a spatially lagged variable was created to quantify the spatial relationship among block-groups on physical disorder. This relationship is suggested in BWT and has garnered much support (see Keizer, Lindenberg, & Steg, 2008; Cerdá et al., 2009; Boggess & Maskaly, 2014; Steenbeek & Kreis, 2015; Wheeler, 2018). Indeed, this study's measure of physical disorder exhibits spatial structure, characterized as significant clustering (*Moran's I* = 0.52, *z-score* = 33.57).[9] Consequently, an inverse distance decay function was utilized to quantify this relationship. This function assumes that block-groups that are closer to the focal block-group are more influential than those that are further away. In particular, all block-groups within two miles of a focal block-group were considered to be influential and used in the calculation of the spatial lag. A distance of two miles was selected to reflect prior research which suggests that most offenders commit crime in nearby areas (Wright & Decker, 1997; Wiles & Costello, 2000; Wright, Brookman, & Bennett, 2006; Bernasco & Block, 2009).[10]

---

[9] The Moran's I test for spatial autocorrelation was conducted using a row normalized inverse distance weighted matrix and 999 permutations. This test was conducted in R using the *sp* package.

[10] Additional spatial lag variables were constructed utilizing an inverse distance decay function with a cap of five-miles, as well as a queens contiguity matrix which defines neighbors as block-groups that share either a common boundary or point with the focal block-group. The resulting lag variables were highly correlated ($r < 0.90$) with the lag variable used in this study and produced essentially identical results.

Table 4. Control/Matching Variables: Full Sample (N = 857).

Variable	Mean	Std. Dev.	Min	Max
Population Density (square-miles)	6442.03	3565.60	369.95	23569.58
Percent Population Under 18	27.04	5.88	0	45.90
Population Growth	-12.50	46.93	-92.24	393.55
Percent Male (15-24)	8.02	5.62	0	41.14
Percent Unemployed	28.36	14.52	0	88.06
Percent Receiving Public Assistance	8.32	7.71	0	69.81
Percent Female-headed Family Household	30.70	14.24	0	78.76
Percent Owner-Occupied Homes	53.62	20.50	0	100
Percent Same Residence (at least one year)	84.42	11.68	29.07	100
Percent African American	83.85	25.28	0	100
Percent Hispanic/Latino Origin	5.59	16.93	0	92.29
Percent White	10.86	18.17	0	100
Percent Foreign-born	4.08	9.36	0	61.09
Violent Crime Rate (per 1000)	23.12	21.92	0.72	350.37
Percent Garden/Park	1.90	6.00	0	100
Percent Commercial	4.91	6.16	0	71.43
Percent Residential	91.92	11.40	5.39	100
Percent Industrial	0.76	3.87	0	77.78
Percent Mixed	0.42	1.05	0	9.38

Table 4 (cont'd)

Physical Disorder Lag	1.38	0.60	0	3.04

Note: If needed, these variables will be modified to facilitate matching.

Analytical Strategy

Levels of physical disorder are not randomly assigned and therefore a robust quasi-experiential design that takes into account selection effects is needed in order to evaluate the relationship between physical disorder and violent crime rate. In the absence of an exogenous instrument needed to perform instrumental variable techniques, this study utilized the generalized dose-response propensity score (GPS) method to estimate the average treatment effect of various levels of physical disorder, measured at $\mathrm{t}_1 = 2014$, on violent crime rate, measured at $\mathrm{t}_2 = 2015$. This examination was performed using the program *gpscore2* available in STATA (version 16.0) statistical software (Guardabascio & Ventura, 2013).

Dose-response models offer several advantages over traditional approaches. In particular, they are superior to regression approaches that address selection effects through the addition of control variables because they create matched groups that must achieve balance on covariates in order for credible comparisons to be made. If balance is unable to be achieved, it means that "apples to apples" comparisons are not possible. Furthermore, unlike traditional propensity score matching which is restricted to a dichotomous causal variable, dose-response models allow covariate balancing across *levels* of a causal variable. Similar to traditional propensity score matching, the dose-response approach eliminates, where possible, bias associated with covariate imbalances (Hirano & Imbens, 2004). Another advantage of dose-response models is that they explicitly model the functional form across each level of the causal variable in such a way as to create balance among covariates.

Notwithstanding these advantages, it is important to acknowledge that both regression and propensity score matching approaches uphold the *weak unconfoundedness assumption* (or *selection on observables assumption)* (Heckman & Robb, 1985). Under this assumption, bias associated with the selection into treatment and treatment-specific outcomes is removed by controlling or conditioning on observable unit characteristics. If there are hidden biases, however, then we lose confidence in this assumption and, consequently, in our ability to draw causal inferences. For this reason, Loughran et al. (2015, p. 636) argue that "despite the many practical advantages of propensity score matching over linear regression, there is nothing magical about propensity score matching that makes it immune to hidden biases that plague regression based causal inferences."

As it pertains to the current study, the GPS method is an appropriate approach for three important reasons. First, it allows for threshold effects to be estimated. Second, it addresses selection effects through covariate balancing across matched levels of physical disorder. Third, it requires that the treatment variable occurs *before* the outcome variable. This requirement complements BWT which supports a delayed effect of disorder on violent crime (Wilson & Kelling, 1982; Kelling & Coles, 1996). The GPS method is conducted in three basic steps.

Step 1: Modeling the conditional distribution of the treatment given covariates

The *gpscore2* package allows the conditional distribution of the treatment to be estimated using general linear models, accommodating a variety of distributions with more flexible assumptions (see Guardabascio & Ventura, 2014). Using estimates from this model, the GPS – defined as the conditional distribution of the treatment given covariates - is computed for each level of the treatment. The general formula for estimating the GPS for each observation using a general linear model is provided as,

$$\widehat{R_i} = r(T,X) = c\left(T,\widehat{\emptyset}\right) exp\left\{\frac{T\hat{\theta}-a(\hat{\theta})}{\widehat{\emptyset}}\right\}, \qquad \text{(eq. 1)}$$

where the parameters $(\emptyset,\theta)$ are associated with distributions of the exponential family.

Hirano and Imbens (2004) utilize a blocking approach to assess how well adjustment for the GPS improves balance among covariates. To begin, the sample is divided into three equal-sized groups, cutting at the 33^{th} and 66^{th} percentiles of the treatment distribution. Within each treatment group, the GPS is evaluated at the median for *all* treatment units. Thus, all treatment units have three sets of GPSs. Next, the treatment units are divided into five blocks for each calculation of the GPS, creating three sets of five blocks. For each block within a set, a mean-difference for every covariate is calculated between treatment units that belong to the corresponding treatment group used to calculate the GPS and those that belong to a different treatment group but nonetheless have a GPS that falls within the block's boundaries. Finally, the resulting five mean-differences produced for every covariate are combined and calculated as a weighted average, and used to produce t-values of differences-in-means. In an ideal scenario, treatment units at each level of the treatment should not be significantly different from one another after adjustment for the GPS. A t-value that is below 1.96 indicates that the covariate means of treatment units belonging to a particular treatment group are no different than the covariate means of treatment units that belong to another treatment group but have similar GPSs. If significant differences are still detected, it should be demonstrated that the adjustment for the GPS greatly reduced mean-differences, despite not reaching statistical significance.

Prior to conducting Hirano and Imbens' (2004) blocking approach, however, it is important to assess overlap among groups in regards to unit characteristics, known as *common support*. It is well known that adjustment for covariates will perform poorly if there is not sufficient overlap in their distributions across treatment levels. In the case in which the treatment is binary, common support is traditionally assessed by evaluating the distribution of the estimated propensity scores for the treatment and comparison groups. Units are often restricted to those that fall within the region of common support (i.e., the region in which the distributions overlap). If included, treatment units that fall outside of the common support region not only may prevent balance from being achieved, but also may result in misleading predictions. In the case in which the treatment is continuous, there is an "infinite number of treatment groups and generalized propensity scores to compare," making matters less straightforward (Flores et al., 2012, p 161).

Flores et al. (2012) offer one approach which serves as a *gauge* of the degree of overlap across different levels of a treatment. Like before, the sample is divided into three equal-sized groups and the GPS is evaluated at the group median, resulting in three sets of GPSs for each treatment unit. Next, the distribution of the GPS for treatment units belonging to each treatment group is compared to the distribution of the GPS for treatment units outside of the evaluated treatment group. Finally, the sample is restricted to those treatment units that are simultaneously comparable across all three treatment groups. In other words, treatment units are dropped that have a GPS that is not among the common support region. It is worth emphasizing, however, that in regions in which the data are sparse, there is less assurance in the accuracy of predictions. To this point, Sullivan and Loughran (2014, p. 715) argue that "…a strong amount of support data at each level of the predictor is necessary to 'learn' the true functional form, or else the relationship will necessarily be based on strong and ultimately untestable functional form assumptions (in particular, where the functional form is off support of the data)." As it relates to the current study, predictions will be less reliable at high levels of disorder, signified by wider confidence intervals.

Step 2: Estimating the conditional expectation of the outcome given the treatment and GPS

The functionality of the GPS requires that we assume that after controlling for unit characteristics, any remaining differences in treatment intensity, *T*, are independent of potential outcomes *Y(t)*. Importantly, this assumption only requires that pairwise conditional independence of the treatment with potential outcomes is assumed, known as weak unconfoundedness. Previously introduced, this assumption requires that selection into a treatment level is random conditional on the observed covariates (Hirano & Imbens, 2004). Omitted variable bias poses a threat to weak unconfoundedness. For this reason, selection bias might still exist in a study's estimators if it does not account for all relevant variables. However, given this study's use of an extensive set of covariates, it is argued that any bias that remains is likely not large enough to influence its findings in a meaningful way.

Furthermore, the balancing property of the GPS can shed light on the assumption of weak unconfoundedness. It implies that treatment assignment is weakly unconfounded given the GPS (Hirano & Imbens, 2004). This quality indirectly addresses the assumption of weak unconfoundedness because treatment units that have similar GPSs also have similar covariates. To this point, Hirano and Imbens (2004) show that if treatment assignment is unconfounded given the covariates, then it is also the case that it is weakly unconfounded given the GPS. As a result, the GPS can be used to remove bias associated with differences in covariates in two steps, the first of which is estimating the conditional expectation of the outcome given the treatment and GPS.

The conditional expectation of the outcome, Y_i , is estimated given the treatment, T_i , and GPS, R_i, given as $\mathrm{E}[Y_i \mid T_i, R_i]$. This function is estimated as a flexible linear function of the covariates. A basic model includes the treatment, GPS, and an interaction of these variables.

Quadratic and cubic transformations of the treatment variable are commonly included, providing even greater flexibility. Unfortunately, *gpscore2* only supports dichotomous, ordinal, and continuous regression models. For this reason, violent crime rate is logged transformed. An example of a model containing a cubic approximation and interaction term is provided as,

$$\varphi\ \mathrm{E}\{(Y_i\ |T_i, R_i)\} = \lambda(\mathrm{T_i}, \mathrm{R_i};\ \alpha) \qquad \text{(eq. 2)}$$

$$= \alpha_0 + \alpha_1 T_i + \alpha_2 T_i^2 + \alpha_3 T_i^3 + \alpha_4 R_i + \alpha_5 T_i R_i,$$

where $\varphi\ (\cdot)$ is a link function of the predictor and $\lambda(\mathrm{T_i}, \mathrm{R_i};\ \alpha)$ relates to the conditional expectation. Importantly, the coefficients in this model *are not* directly interpretable (Hirano & Imbens, 2004). That being said, Kluve et al. (2012, p. 19) note that "whether all the estimated coefficients associated with the [GPS] terms are equal to zero can indicate whether the covariates introduce any bias." In other words, statistically significant GPS-related parameters suggest that the covariates introduce bias and that the propensity score matching approach is relevant in that it helps tease out the causal relationship between the treatment and outcome.

Step 3: Estimating the dose-response function to discern treatment effects

The second way in which the GPS can be used to remove bias associated with differences in covariates is by estimating the *dose-response function* at each level of the treatment. In particular, the parameters estimated in the previous step are used to estimate the average potential outcome associated with each treatment level over the GPS. This function is provided as,

$$\mathrm{E}(\widehat{Y(t)}) = \frac{1}{N}\sum_{i=1}^{N} \varphi^{-1}[\,\hat{\lambda}(\mathrm{t}, \hat{r}(\mathrm{t}, X_i);\ \hat{\alpha})] \qquad \text{(eq. 3)}$$

Furthermore, Rosenbaum (2002) suggests conducting a sensitivity analysis to determine the magnitude of hidden bias that would need to be present to alter study findings. For this reason, an estimate of uncertainty was conducted by bootstrapping standard errors, an option available in *gpscore2* (Rosenbaum, 2002). When selected, this option incorporates the estimation of the GPS along with the estimation of the other predicting parameters. Each replication helps provide a better understanding of the uncertainty associated with these estimates, which is captured by upper and lower confidence intervals.

Sensitivity Checks

Sullivan and Loughran (2014) identify assumptions of the GPS approach that may hamper the identification of the true functional form of the relationship between the treatment and outcome. To start, the approach requires that the treatment be a linear function of the covariates, an assumption of *parametric* regression. As previously discussed, the conditional expectation function is produced from a linear model and traditionally includes the treatment, GPS, and an interaction of these variables. Thus, the GPS approach assumes that the conditional expectation function is governed by a specific parametric form (Sullivan & Loughran, 2014). Sullivan and Loughran (2014, p. 714) state that "[t]here are no theoretical reasons apparent as to why this particular functional form is optimal, nor is it clear why this is the best means of estimating that function (particularly the interaction term)." That being said, they suggest that the "true functional form of the relationship would be robust to slight alterations in this specification" (Sullivan & Loughran, 2014, p. 714). As part of its sensitivity checks, the current study explores how slight alterations to the specification of the conditional expectation function, such as the inclusion of polynomial terms and the exclusion/inclusion of an interaction term, affect the estimation of the dose-response function.

Furthermore, the GPS approach has been adjusted to incorporate nonparametric techniques, and for good reason (see Flores et al., 2012; Kluve et al., 2012; Kreif, Grieve, Diaz, & Harrison, 2014; Fong, Hazlett, & Imai, 2018). These techniques relax many of the strong assumptions made by traditional regression and are thought to allow the functional form of the relationship between the treatment and outcome to more naturally emerge from the data. Extending the parametric GPS approach, Bia et al. (2014) developed a set of Stata programs – *drf* – to estimate the dose-response function using semiparametric estimators that draw from penalized spline techniques, and a kernel estimator developed by Flores et al. (2012). These approaches may capture nonlinear patterns that parametric regression models overlook. For example, polynomial regression – as shown in step two of the GPS approach - includes polynomial terms (x^2, x^3, etc.) for predictors in a linear regression model. The inclusion of polynomial terms provides regression models more flexibility to capture nonlinear relationships. However, polynomial terms impose a *global* structure on the relationship between the predictor and outcome. For example, the use of a cubic polynomial term means that the relationship between the predictor and outcome is cubic over the *entire* range of the predictor. It is clear that imposing a global structure is limiting. Perhaps the relationship between the predictor and the outcome is only cubic at low ranges of the predictor. If so, then polynomial regression will fail to capture the true functional form of the relationship. Splines are a nonparametric technique that offer an alternative approach to estimating relationships of unknown functional form and are commonly used in semiparametric regression models. These models allow "some of the covariates [to] enter the model in a parametric fashion, while other variables," such as splines, "enter as nonparametric terms" (Keele, 2008, p. 109).

Splines are customarily formed as the summation of locally defined polynomials – referred to as "basis functions" - which meet at knots that span across the entire range of the predictor. As the number of knots increase, so too does the flexibility of the smooth function. However, using a large number of knots runs the risk of overfitting the data. Conversely, using a small number of knots runs the risk of underfitting the data. Penalized splines attempt to strike a balance between overfitting and underfitting the data by imposing weights on each smooth function (Perperoglou et al., 2019). In particular, these weights are used to penalize overfitting the data while still offering enough flexibility to fit the data well. In this way, the approach reduces concerns regarding the appropriate number of knots. In fact, many studies have found knot specification to be of minor concern for penalized splines (e.g, Eilers & Marx, 1996; French, Kammann, & Wand, 2001; Ruppert, 2002; Ruppert, Wand, & Carroll, 2003). In particular, Ruppert et al. (2003) found the selection method K = min($\frac{n}{4}$, 35), where n is the number of unique T_i, to work well. This method is the default for *drf*.[11] In addition, penalized splines have been found to avoid "wild behavior near the extremes of the data" by imposing linearity constraints at boundary knots (Fox, 2000, p. 67).

As it relates to the GPS approach, penalized spline regression is conducted for the second stage of the estimation of the dose-response function. This approach utilizes different basis functions to accommodate the nonlinear structure of the data and perform penalized spline smoothing. The simplest penalized spline approach performs smoothing in an additive fashion for the treatment and GPS using bivariate basis functions (see Bia et al., 2014), aptly named for its consideration of two continuous variables (see Bia et al., 2014; Ruppert et al., 2003). This model is provided as,

$$\mathrm{E}[\mathrm{Y_i}\,|\mathrm{T_i}, \mathrm{R_i}] = \alpha_0 + \alpha_t \mathrm{T_i} + \alpha_r \mathrm{R_i} + \sum_{\mathbf{k=1}}^{\mathbf{K^t}} \boldsymbol{\mu}_{\mathbf{k}}^{\mathbf{t}}(\mathbf{T_i} - \mathbf{k}_{\mathbf{k}}^{\mathbf{t}}) + \sum_{\mathbf{k=1}}^{\mathbf{K^r}} \boldsymbol{\mu}_{\mathbf{k}}^{\mathbf{r}}(\mathbf{R_i} - \mathbf{k}_{\mathbf{k}}^{\mathbf{t}}), \qquad \text{(eq. 4)}$$

where K^t and K^r are knots for the treatment and GPS, respectively, and μ_k^t and μ_k^r are the related knot coefficients (see Ruppert et al., 2003; Bia et al, 2014).[12] Furthermore, the radial

[11] Ruppert (2002) provides empirical justifications for this knot specification method.
[12] Ruppert et al. (2003) demonstrate that in the case of a simple additive model the penalization of μ_k^t and μ_k^r is incurred by treating them as random effects.

basis function approach adds complexity to this basic structure by relying on distance calculations between every data point and knots to inform penalized spline smoothing (see Wand, 2003; Ruppert et al., 2003; Bia et al., 2014). This model is provided as,

$$E[Y_i \mid T_i, R_i] = \alpha_0 + \alpha_t T_i + \alpha_r R_i + \sum_{\mathbf{k=1}}^{\mathbf{k}} \boldsymbol{\mu}_{\mathbf{k}} \mathbf{C}\left(\left\| \begin{pmatrix} \mathbf{T_i} \\ \mathbf{R_i} \end{pmatrix} - \begin{matrix} \mathbf{k^t_{k'}} \\ \mathbf{k^r_{k'}} \end{matrix} \right\|\right), \qquad \text{(eq. 5)}$$

where C is the covariance function based on knots for the treatment and GPS.

For both applications, mixed models are used to represent the penalized splines using the *xtmixed* subcommand and are estimated using restricted maximum likelihood (REML). Mixed models perform smoothing by including coefficients that *are not* associated with knots as fixed effects, while coefficients that *are* associated with knots are included as random effects (Wand, 2003). Thus, the simple penalized spline approach includes two random effects parameters (indicated in bold in equation 4), while the radial basis function approach includes only one random effect parameter (indicated in bold in equation 5). Ultimately, a model which includes both fixed and random effects offers the greatest amount of flexibility in capturing the true functional form of the relationship between the treatment and outcome (see Ruppert et al., 2003). In a similar manner as before, the parameters estimated from these models – which are not directly interpretable - are subsequently used to calculate the average potential outcome at each treatment level by averaging over the GPS, the final step of the GPS approach.

Another approach available in *drf* utilizes a nonparametric kernel estimator to estimate the final stage of the GPS approach. In particular, this approach estimates the dose-response function using local polynomial regression and an inverse weighting estimator that is based on kernel methods, where the weights are constructed from the GPS and adjust for covariate differences (Flores et al., 2012). The global bandwidth of the kernel is selected using Fan and Gijbels' (1996) proposed procedure, the default in *drf*. In order to estimate the unknown parameters of the optimal global bandwidth, this procedure uses global polynomials of the GPS of order p +3, where p is the order of the fitted local polynomial (Bia et al., 2014). The inverse weighted estimator of the average dose-response function is provided as,

$$\widehat{\mu}(t)_{IW} = \frac{D_o(t)S_2(t) - D_1(t)S_1(t)}{S_o(t)S_2(t) - S_1^2(t)}, \qquad \text{(eq. 6)}$$

where $S_j(t) = \sum_{i=1}^{N} \widetilde{K}_{h,X}(T_i - t)\,(T_i - t)^j$ and $D_j(t) = \sum_{i=1}^{N} \widetilde{K}_{h,X}(T_i - t)\,(T_i - t)^j Y_i$, with the weighted kernel function indicated as $\widetilde{K}_{h,X}(T_i - t)$ (Flores et al., 2012). This local estimator is preferred due to its ability to avoid bias near data boundaries (Flores et al., 2012).

In summary, this study estimated the dose-response function utilizing both parametric and nonparametric techniques. The parametric approach includes an assessment of the sensitivity of the dose-response function to slight alterations to the specification of the conditional expectation function. The best fitting model is identified and discussed. Following this assessment, three different types of semiparametric methods are explored: penalized spline, radial spline, and inverse weighting kernel function.

CHAPTER 4: ANALYSIS & RESULTS

Parametric Method

The estimation of the dose-response function was first conducted using the parametric GPS approach and followed three key steps: 1) Modeling the conditional distribution of the treatment given covariates; 2) Estimating the conditional expectation of the outcome given the treatment and GPS; and 3) Estimating the dose-response function to discern treatment effects.

Step 1: Parametric Approach: Modeling the conditional distribution of the treatment given covariates

To begin, the conditional distribution of physical disorder given covariates was estimated. The prediction model was developed using control variables identified from broken windows and related research. As the treatment variable is continuous and right-skewed, log-normal and gamma distributions were evaluated to identify the most appropriate distributional fit. To this end, an evaluation of theoretical densities and goodness-of-fit criteria was conducted using the Anderson-Darling and Kolmogorov-Smirnov statistics, and Bayesian Information Criterion (BIC) (see Figure 14 & Table 5). In particular, the Anderson-Darling and Kolmogorov-Smirnov statistics compare the observed cumulative distribution function to the expected cumulative distribution function, which in this case is either a log-normal or gamma distribution. One key difference between these statistics is that the Anderson-Darling statistic gives more consideration to the tails of a distribution. In either case, a smaller test statistic indicates better fit. Based on the likelihood function, the BIC is a well-known criterion for model selection that includes a larger penalty term - determined by the number of parameters in the model - than the closely related Akaike Information Criterion. A smaller BIC indicates better model fit. Considering these goodness-of-fit criteria, there was considerable support in favor of a gamma distribution. For this reason, a gamma distribution was used to model physical disorder.

Figure 15. Theoretical Densities.

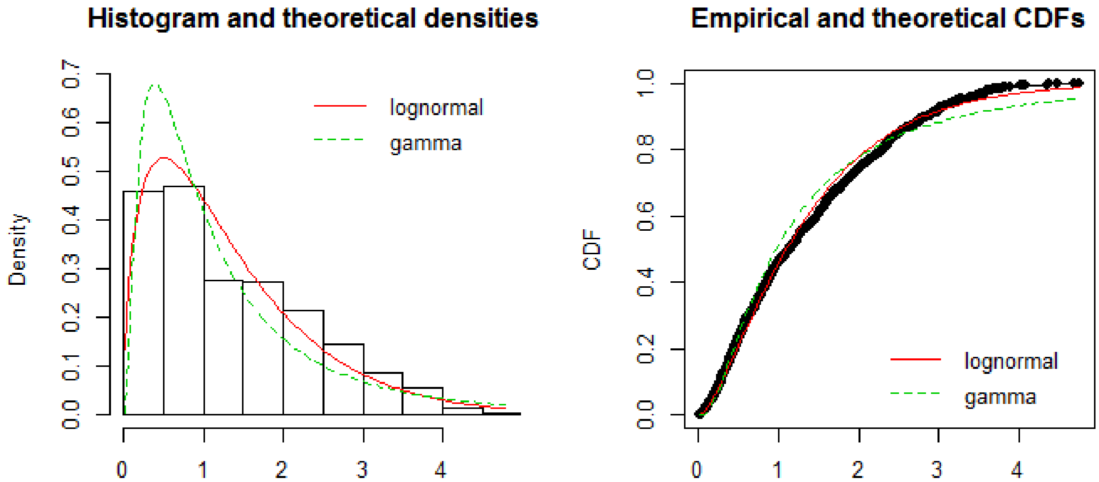


Table 5. Physical Disorder: Goodness of Fit Statistics.

Goodness of Fit Tests	Gamma	Log-normal
Anderson-Darling Statistic	2.81	9.99
Bayesian Information Criterion	2176.10	2301.61
Kolmogorov-Smirnov Statistic	0.05	0.08

Individual effects are presented in Table 6 from the estimation of the conditional distribution of physical disorder given covariates. Importantly, these findings are of interest insofar as they produce a GPS that achieves balance amongst covariates (Hirano & Imbens, 2004). Following Hirano and Imbens' (2004) suggestion, two types of transformations were utilized in an effort to achieve balance: square-root and natural log.[13] As a consequence, the interpretation of the effects presented in Table 6 is not straightforward. Proceeding with caution, the results presented in Table 6 largely reflect findings from prior studies. Although not relevant to the advancement of the GPS approach, a few of these findings are worthy of discussion.

[13] Like the natural log transformation, the square-root transformation is used to minimize right skewness, although it has a weaker effect. Unlike the natural log transformation, however, it can be applied to zero values. Herein lies the key advantage of the square-root transformation over the natural log transformation.

Residential stability is captured by percent owner-occupied homes and percent same residence. Stability within neighborhoods encourages trust, and shared values and norms amongst residents (Shaw & McKay, 1942; Coleman, 1988, 1990; Sampson, 2012; Markowitz et al. 2001; Ingoldsby & Shaw, 2002). These social processes provide fertile grounds for the development of informal social control within neighborhoods which help protect against the spread of disorder. As anticipated, percent home-owners maintains a statistically significant (p-$value \leq 0.001$), negative relationship with physical disorder, suggesting that home-ownership positively contributes to the social processes that occur within neighborhoods that help protect them against disorder. Contrary to what was expected, however, percent same residence maintains a marginally significant (p-$value = 0.09$), positive relationship with physical disorder. That being said, one year may not be long enough to capture a protective effect. To this point, remaining in the same residence for at least *five years* is a more common metric by which to capture residential stability using Census data (e.g., Warner & Rountree, 1997; Sampson et. al., 1997; Boggess & Hipp, 2010). However, this measure could not be constructed at the block-group level using five-year estimates provided by the ACS.

Furthermore, population density maintains a statistically significant (p-$value \leq 0.001$), negative relationship with physical disorder, indicating that the most disorderly neighborhoods are those that are the least densely populated. Since 2000, Detroit has experienced a massive loss in population. As Detroit's population dwindled, its number of abandoned and neglected properties increased. Consequently, neighborhoods that experienced the greatest losses in population also experienced the largest increases in physical disorder. This effect is reflected by the statistically significant (p-$value \leq 0.001$), negative relationship observed between population growth and physical disorder.

In addition, percent garden/park maintains a statistically significant (*p-value* ≤ 0.05), negative relationship with physical disorder. Community gardens and, to a lesser extent, parks signify residents' investments in their neighborhoods and promote collective efficacy by providing opportunities for residents to informally interact with one another, establishing shared norms, trust, and solidarity (Cohen, Inagami, & Finch, 2008; Teig et al., 2009; Clayton, 2007; Kearney, 2009; Alaimo et al., 2010). For these reasons, percent garden/park can be considered to be a reasonable (but imperfect) proxy of collective efficacy, helping explain its negative relationship with physical disorder. In fact, other studies have relied on similar indicators to serve as proxies in the absence of traditional measures of collective efficacy (e.g., Wheeler, 2018, 2019).

Table 6. Conditional Distribution of Physical Disorder given Covariates.

Variable	β
ln(Population Density) (square-miles)	-0.47*** (0.06)
Population Growth	-0.003*** (0.001)
sqrt(Percent Population Under 18)	0.29*** (0.05)
sqrt(Percent Male (15-24))	0.01 (0.02)
sqrt(Percent Unemployed)	0.08*** (0.02)
sqrt(Percent Receiving Public Assistance)	0.04** (0.01)
Percent Female-headed Family Household	0.001 (0.002)
Percent Owner-Occupied Homes	-0.01*** (0.001)

Table 6 (cont'd)

Percent Same Residence	0.003+ (0.002)
Percent African American	0.0004 (0.003)
sqrt(Percent Hispanic/Latino Origin)	0.02 (0.02)
Percent White	-0.005 (0.003)
sqrt(Percent Foreign-born)	0.007 (0.02)
ln(Violent Crime Rate) (per 1000)	0.04* (0.02)
sqrt(Percent Garden/Park)	-0.05* (0.03)
sqrt(Percent Commercial)	0.09** (0.03)
sqrt(Percent Residential)	0.24*** (0.07)
sqrt(Percent Industrial)	0.001 (0.01)
sqrt(Percent Mixed)	0.25*** (0.06)
Physical Disorder Lag	0.06*** (0.01)
Constant	-0.47 (0.98)

+ p-value p≤ .10; * p-value ≤ .05; ** p-value ≤.01; *** p ≤ .001.

Using estimates from this model, the GPS was computed for each level of the treatment, and common support and balance assessed. To begin, three equal-sized groups were created by dividing the distribution of physical disorder at the 33th and 66^{th} percentiles. Flores et al.'s (2011) approach was subsequently applied to identify the region of common support. Block-groups that fell outside of this region (i.e., off support block-groups) were removed, resulting in an 11.52% reduction of this study's sample (N = 760).

Table 7 displays the characteristics of the block-groups that fell *within* the common support region. For a more grounded understanding of these characteristics, they are presented without the transformations utilized to generate the GPS. A comparison of means between the common support (N = 760) and full sample (N = 857) was conducted: $H_0\colon\ \bar{X}_{CS} - \bar{X}_{FS}\ = 0;\ H_1\colon\ \bar{X}_{CS} - \bar{X}_{FS}\ \neq 0.$ A bonferroni-adjusted p-value was utilized - $\frac{0.05}{22}$ = 0.0022 - to account for the increased probability of type 1 error associated with multiple comparisons. None of the identified differences reached statistical significance as determined by this conservative standard. Relaxing this standard, however, two statistically significant mean-differences were detected. The mean levels of physical disorder (*p-value* = 0.04) and physical disorder lag (*p-value* = 0.07) are lower in the common support sample than in the full sample. While still positively skewed, the distribution of physical disorder now contains fewer block-groups with high levels of physical disorder. As a result, sparse data at these levels will reduce the reliability of this study's predictions.

For ease of comparison, Table 8 displays mean characteristics for *each group* within the common support region without transformations. Utilizing a one-way analysis of variance (ANOVA), statistically significant mean group differences were identified for 15 out of the 22 variables used in this study: $H_0\colon\ \bar{X}_{G1} = \bar{X}_{G2} = \bar{X}_{G3}$; $H_1\colon\ not\ all\ \bar{X}_{i\ (i=1,2,3)}\ are\ equal.$[14] These variables include physical disorder (*p-value* ≤ 0.001), violent crime rate (2015) (*p-value* ≤ 0.001), violent crime rate (2014) (*p-value* ≤ 0.001), population density (*p-value* ≤ 0.001), population under 18 (*p-value* ≤ 0.001), population growth (*p-value* ≤ 0.001), percent unemployed (*p-value* ≤ 0.001), percent receiving public assistance (*p-value* ≤ 0.001), percent female-headed households (*p-value* = 0.09), percent owner-occupied home (*p-value* ≤ 0.001), percent African American (*p-value* = 0.07), percent white (*p-value* = 0.03), percent foreign-born (*p-value* = 0.06), percent garden/park (*p-value* ≤ 0.001), and physical disorder lag (*p-value* ≤ 0.001). Together, these findings support the relevance of the GPS approach, as there are substantial imbalances across covariates examined by group without utilizing transformations or adjusting for the GPS. The ability of the GPS approach to create balance where it is needed will soon be presented.

Table 7. Common-support Sample (N = 760).

Variable	Mean	Std. Dev.	Min	Max
Independent Variable (2014)				
Physical Disorder	1.27	0.91	0.02	4.76
Dependent Variable (2015)				
Violent Crime Rate (per 1000)	21.25	16.38	2.41	277.37
Control Variables (2014)				
Population Density) (square-miles)	6681.73	3448.05	658.02	23569.58
Percent Population Under 18	27.22	5.789	0.00	45.90

[14] A one-way ANOVA simultaneously compares all group means. As a result, it is able to maintain the type 1 error probability at a user-designated level. Thus, this method of comparison – as a single test - is not affected by the multiple comparison problem.

Table 7 (cont'd)

Population Growth	-11.24	45.96	-92.24	393.55
Percent Male (15-24)	8.03	5.48	0.00	41.14
Percent Unemployed	27.87	13.94	0.00	87.24
Percent Receiving Public Assistance	8.07	7.15	0.00	43.33
Percent Female-headed Family Household	30.58	14.03	0.00	78.76
Percent Owner-Occupied Homes	54.28	19.92	0.00	100.00
Percent Same Residence (at least one year)	84.71	11.39	36.93	100.00
Percent African American	83.41	25.94	0.00	100.00
Percent Hispanic/Latino Origin	6.03	17.76	0.00	92.28
Percent White	11.13	18.45	0.00	93.25
Percent Foreign Born	4.35	9.76	0.00	61.08
Violent Crime Rate (per 1000)	22.17	19.46	1.20	350.36
Percent Garden/Park	1.47	3.91	0.00	40.00
Percent Commercial	4.79	6.26	0.00	71.42
Percent Residential	92.28	11.24	5.39	100.00
Percent Industrial	0.61	2.83	0.00	40.00
Percent Mixed	0.22	0.56	0.00	5.60
Physical Disorder Lag	1.33	0.56	0.00	3.04

Table 8. Group Means from Common-support Sample (N = 760).

Variable	Group 1 Mean (N = 260)	Group 2 Mean (N = 276)	Group 3 Mean (N = 224)
Independent Variable (2014)			
Physical Disorder	0.38	1.05	2.34
Dependent Variable (2015)			
Violent Crime Rate (per 1000)	17.61	22.51	23.55
Control Variables (2014)			
Population Density (square-miles)	7533.65	7332.87	5219.4
Population Growth	1.10	-9.62	-24.82
Percent Male (15-24)	7.89	7.71	8.46
Percent Unemployed	23.69	28.33	31.51
Percent Receiving Public Assistance	6.70	8.42	9.07
Percent Female-headed Family Household	29.01	31.57	31.15
Percent Owner-Occupied Homes	58.92	53.68	50.34
Percent Same House (at least one year)	84.69	83.83	85.60
Percent African American	82.38	81.37	86.41
Percent Hispanic/Latino Origin	5.71	7.75	4.69
Percent White	12.62	12.21	8.64
Percent Foreign-born	4.38	5.37	3.33

Table 8 (cont'd)

Violent Crime Rate (per 1000)	17.33	23.41	25.66
Percent Garden/Park	3.11	1.17	1.14
Percent Commercial	5.00	4.54	4.82
Percent Residential	91.68	93.17	91.99
Percent Industrial	0.37	0.72	0.75
Percent Mixed	0.24	0.20	0.22
Physical Disorder Lag	1.02	1.32	1.64

In addition, Figure 15 displays a map of the block-groups from the common support sample indicated in gray. Off-support block-groups are indicated in red. As can be seen, off support block-groups appear to be more highly concentrated in central and north-east Detroit. Characteristics of these block-groups are provided in Table 9 and shown without transformations. Out of the 22 examinations conducted, 6 statistically significant mean-differences were identified between the off support (N = 97) and common support (N = 760) samples: H_0: $\bar{X}_{OS} - \bar{X}_{CS} = 0$; H_1: $\bar{X}_{OS} - \bar{X}_{CS} \neq 0$. Statistical significance was determined utilizing the previously calculated bonferroni-adjusted p-value (*p-value* ≤ 0.0022). In particular, mean levels of physical disorder, physical disorder lag, violent crime rate (2015), violent crime rate (2014), and percent unemployed are significantly higher in the off support sample, while population density is significantly lower. Relaxing this conservative standard for statistical significance, 11 other differences emerged and are presented in Table 10. With few exceptions, these findings suggest that the excluded block-groups include those with the most severe social problems, indicated by elevated levels of physical disorder, violence, and indicators of disadvantage (*p-value* ≤ 0.05).

Figure 16. Common-support (gray) and Off-support (red) Block-groups.

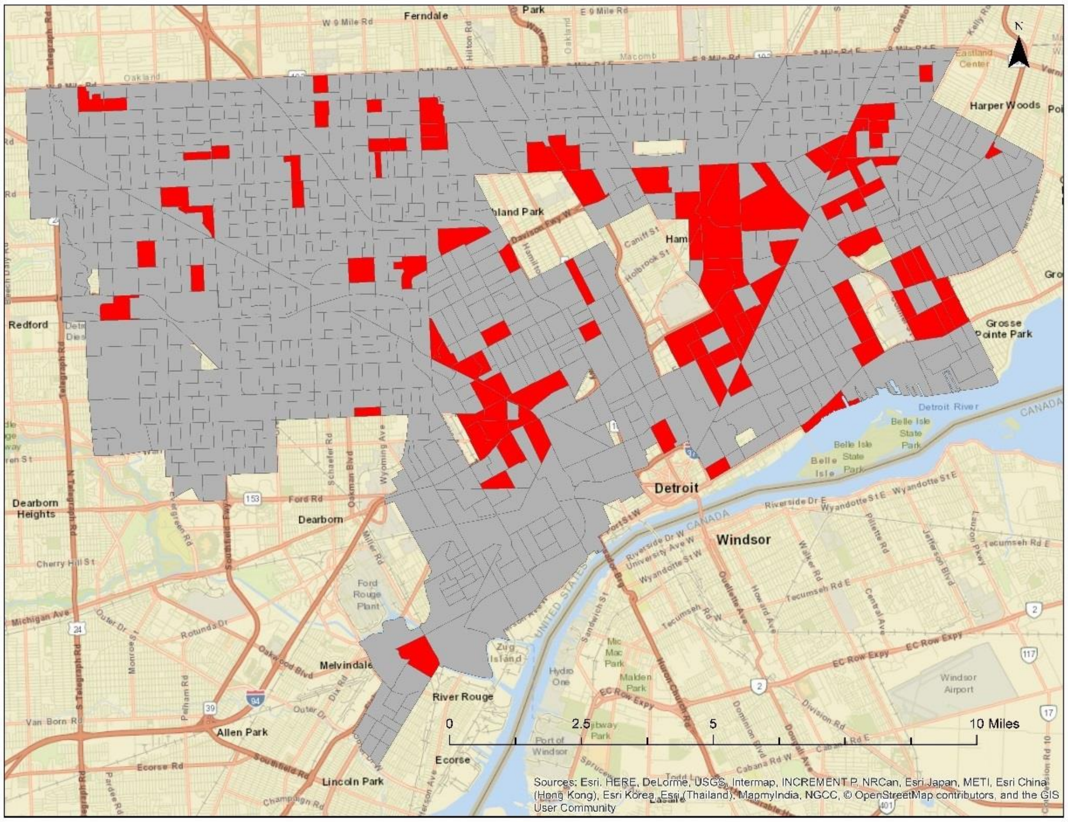


Table 9. Off-support Sample (N = 97).

Variable	Mean	Std. Dev.	Min	Max
Independent Variable (2014)				
Physical Disorder	2.15	1.27	0.02	4.48
Dependent Variable (2015)				
Violent Crime Rate (per 1000)	29.52	32.64	1.70	276.32
Control Variables (2014)				
Population Density) (square-miles)	4563.95	3916.52	369.95	17261.04
Percent Population Under 18	25.61	6.40	5.54	39.62
Population Growth	-22.38	53.16	-90.18	268.62

Table 9 (cont'd)

Percent Male (15-24)	7.96	6.64	0.00	30.26
Percent Unemployed	27.87	13.94	0.00	87.24
Percent Receiving Public Assistance	10.21	11.00	0.00	69.81
Percent Female-headed Family Household	32.13	15.46	0.00	77.35
Percent Owner-Occupied Homes	48.88	23.71	29.07	100.00
Percent Same Residence (at least one year)	82.12	13.58	29.07	100.00
Percent African American	88.71	15.80	22.51	100.00
Percent Hispanic/Latino Origin	1.99	6.86	0.00	46.32
Percent White	7.27	11.83	0.00	67.96
Percent Foreign-born	1.84	4.57	0.00	24.69
Violent Crime Rate (per 1000)	30.54	35.05	0.72	302.63
Percent Garden/Park	1.47	3.91	0.00	40.00
Percent Commercial	5.91	5.29	0.00	25.14
Percent Residential	89.88	9.40	56.00	100.00
Percent Industrial	1.05	2.95	0.00	16.67
Percent Mixed	0.26	0. 37	0.00	1.57
Physical Disorder Lag	1.79	0.74	0.31	2.86

Table 10. Off-support and Common-support Mean Differences.

Variable	Mean-difference	P-value
Percent Population Under 18	-1.61	0.01
Population Growth	-11.14	0.03
Percent Receiving Public Assistance	2.14	0.01
Percent Owner-Occupied Homes	-5.40	0.01
Percent Same Residence (at least one year)	-2.59	0.04
Percent African American	5.29	0.05
Percent Hispanic/Latino Origin	-4.04	0.03
Percent White	-3.85	0.05
Percent Foreign-born	-2.51	0.01
Percent Commercial	1.12	0.09
Percent Residential	-2.39	0.05

Following this assessment, the balancing property was evaluated on the common support sample using the previously created groups and transformations identified in Table 6. Covariate means for each group were compared against the remaining groups to assess how balance was affected by adjustment for the GPS, resulting in 60 mean group comparisons (see Tables 11, 12, & 13). Without adjusting for the GPS, 23 mean-differences were identified to be statistically significant, indicated by a t-value greater than or equal to 1.96. Thus, pre-adjustment comparisons indicate substantial imbalances across groups. To inspire confidence in the GPS approach, adjustment for the GPS should eliminate or substantially reduce these imbalances. After adjustment, the mean-difference for population density remained statistically significant for treatment group 3. That being said, the GPS adjustment resulted in a considerable improvement, reducing the mean-difference by 61.29%. Overall, the GPS adjustment reduced mean-differences by an average of 55.18%.

Table 11. Adjustment for the GPS: Group 1.

Treatment Group 1 [0.02, 0.68]					
Covariates	Pre-GPS		Post-GPS		
	Diff.	t-value	Diff.	t-value	Percent Diff.
Population Density (square-miles)	0.20	**4.70**	0.05	0.92	75.00
Percent Population Under 18	0.32	**6.99**	0.10	1.67	68.75
Population Growth	-0.18	**-5.63**	-0.02	-0.53	88.89
Percent Male (15-24)	0.02	0.18	0.01	0.15	50.00
Percent Unemployed	0.62	**5.62**	0.11	0.82	82.26
Percent Receiving Public Assistance	-0.40	**-3.59**	-0.06	-0.44	85.00
Percent Female-headed Family Household	2.33	**2.16**	0.94	0.66	59.66

Table 11 (cont'd)

Percent Owner-Occupied Homes	-6.93	**-4.57**	-0.86	-0.44	87.59
Percent Same Residence (at least one year)	0.19	0.15	0.04	0.04	78.95
Percent African American	3.71	1.43	1.54	0.77	58.49
Percent Hispanic/Latino Origin	-0.21	-0.99	-0.01	-0.10	95.24
Percent White	-2.61	-1.47	-2.22	-0.16	14.94
Percent Foreign-born	-0.30	-1.72	-0.09	-0.75	70.00
Violent Crime Rate (per 1000)	0.71	**6.33**	0.23	1.55	67.61
Percent Garden/Park	-0.28	**-3.21**	-0.12	-1.13	57.14
Percent Commercial	-0.07	-0.86	-0.06	-0.59	14.29
Percent Residential	0.08	1.54	0.07	1.19	12.50
Percent Industrial	0.36	1.66	0.19	0.69	47.22
Percent Mixed	0.10	**2.54**	0.01	0.28	90.00
Physical Disorder Lag	2.43	**8.31**	0.36	0.96	85.19

Table 12. Adjustment for the GPS: Group 2.

	Treatment Group 2 [0.69, 1.52]				
Covariates	Pre-GPS		Post-GPS		
	Diff	t-value	Diff	t-value	Percent Diff.
Population Density (square-miles)	-0.16	**-3.80**	-0.12	-1.80	25.00
Percent Population Under 18	-0.21	**-2.48**	-0.19	-1.95	9.52
Population Growth	-0.02	-0.72	-0.01	-0.27	50.00
Percent Male (15-24)	0.07	0.83	.04	0.52	42.86

Table 12 (cont'd)

Percent Unemployed	-0.13	-1.20	-0.05	-0.43	61.54
Percent Receiving Public Assistance	-0.19	-1.75	-0.12	-1.03	36.84
Percent Female-headed Family Household	-1.52	-1.32	-1.47	-1.36	3.29
Percent Owner-Occupied Homes	0.88	0.57	0.24	0.14	72.73
Percent Same House (at least one year)	1.31	1.50	0.94	1.00	28.24
Percent African American	3.05	1.52	2.56	1.22	16.07
Percent Hispanic/Latino Origin	-0.28	-1.64	-0.23	-1.36	17.86
Percent White	-1.60	-1.12	-0.61	-0.41	61.88
Percent Foreign-born	-0.22	-1.68	-0.20	-1.43	9.09
Violent Crime Rate (per 1000)	-0.63	-1.29	-0.13	-1.17	79.37
Percent Garden/Park	0.13	1.44	0.12	1.26	7.69
Percent Commercial	0.10	1.11	0.06	0.77	40.00
Percent Residential	-0.08	-1.46	-0.07	-1.44	12.50
Percent Industrial	-0.17	-0.74	-0.15	-0.72	11.76
Percent Mixed	0.04	1.13	0.03	0.91	25.00
Physical Disorder Lag	-0.39	-1.30	-0.21	-0.65	46.15

Table 13. Adjustment for the GPS: Group 3.

Treatment Group 3 [1.53, 4.76]					
Covariates	Pre-GPS		Post-GPS		
	Diff	t-value	Diff	t-value	Percent Diff.
Population Density (square-miles)	-0.31	**-4.70**	-0.12	**-2.34**	61.29
Percent Population Under 18	-0.11	**-2.35**	-0.04	-0.58	63.64
Population Growth	-0.11	**6.02**	- 0 .03	-0.75	72.73
Percent Male (15-24)	-0.07	-0.67	-.06	-0.58	14.29
Percent Unemployed	-0.48	**-4.33**	-0.19	-1.32	60.42
Percent Receiving Public Assistance	-0.20	-1.80	-0.01	-0.07	95.00
Percent Female-headed Family Household	-0.85	-0.79	0.33	0.23	138.82
Percent Owner-Occupied Homes	5.96	**3.94**	1.14	0.56	80.87
Percent Same House (at least one year)	-0.91	-1.05	-0.61	-0.51	32.97
Percent African American	-4.53	**-2.28**	-1.75	-0.64	61.37
Percent Hispanic/Latino Origin	0.29	1.74	0.17	0.77	41.38
Percent White	3.77	**2.68**	1.26	0.63	66.58
Percent Foreign-born	0.31	**2.43**	0.14	0.82	54.84
Violent Crime Rate (per 1000)	-0.57	**-5.04**	-0.04	-0.28	92.98
Percent Garden/Park	0.15	1.74	-1.2e-05	-9.2e-05	100.01
Percent Commercial	-0.14	-1.62	-0.11	-0.97	21.43

Table 13 (cont'd)

Percent Residential	-0.02	-0.27	-0.004	-0.08	80.00
Percent Industrial	0.20	0.92	0.02	0.09	90.00
Percent Mixed	0.09	1.70	0.04	1.13	55.56
Physical Disorder Lag	-2.79	**-9.76**	-0.45	-1.56	83.87

Step 2: Parametric Approach: Estimating the conditional expectation of the outcome given the treatment and GPS

The second step of the GPS approach involves estimating the conditional expectation of logged violent crime rate given physical disorder and the GPS using ordinary least squares regression.[15] In order to assess whether the functional form is robust to slight alterations in model specifications, models with and without interactions were separately conducted for base, quadratic, and cubic transformations of physical disorder. This procedure resulted in the estimation of 6 models (see Tables 14, 15, & 16):

- Model 1a – Base, interaction
- Model 1b – Base, no interaction
- Model 2a – Quadratic, interaction
- Model 2b – Quadratic, no interaction
- Model 3a – Cubic, interaction
- Model 3b – Cubic, no interaction

These models were then compared using a series of likelihood ratio tests. Model 1b served as the reduced/restricted model and was determined to have the best fit overall (see Table 17). Across all models, the GPS term was statistically significant, lending support to the relevance of the GPS approach (Hirano & Imbens, 2004).

[15] The residuals from model 1b (the best fitting model) were assessed using a Moran's I test for spatial autocorrelation available in the *sp* package in R. In particular, this test was conducted using a row normalized inverse distance weighted matrix and 999 permutations. While statistically significant clustering was detected (*p-value* = 0.02), the level of spatial autocorrelation identified is unlikely to greatly affect model results.

Table 14. Base.

	Model 1a		Model 1b (Best)	
	Beta	SE	Beta	SE
Physical Disorder	0.05	0.04	0.05	0.04
GPS	-0.28***	0.08	-0.28***	0.08
Physical Disorder x GPS	-0.03	0.37	-	-
Constant	2.93***	0.16	2.92***	0.08
Adjusted R-Squared	0.0528		0.0540	
BIC	1438.637		1432.012	

+p-value≤ .10; * p-value ≤ .05; ** p-value ≤.01; *** p ≤ .001.

Table 15. Quadratic.

	Model 2a		Model 2b	
	Beta	SE	Beta	SE
Physical Disorder	0.07	0.13	0.06	0.13
Physical Disorder^2	-0.003	0.03	-0.002	0.03
GPS	-0.28**	0.11	-0.28**	0.11
Physical Disorder x GPS	-0.04	0.39	-	-
Constant	2.92***	0.22	2.91***	0.14
Adjusted R-Squared	0.0516		0.0528	
BIC	1445.261		1438.64	

+ p-value ≤ .10; * p-value ≤ .05; ** p-value ≤.01; *** p ≤ .001.

Table 16. Cubic.

	Model 3a		Model 3b	
	Beta	SE	Beta	SE
Physical Disorder	-0.39	0.32	-0.31	0.29
Physical Disorder^2	0.23	0.16	0.19	0.14
Physical Disorder^3	-0.04	0.02	-0.02	0.02
GPS	-0.41**	0.14	-0.39**	0.14
Physical Disorder x GPS	0.23	0.42	-	-
Constant	3.10***	0. 22	3.13***	0.21
Adjusted R-Squared	0.0532		0.0541	
BIC	1449.603		1443.263	

+ p-value ≤ .10; * p-value ≤ .05; ** p-value ≤.01; *** p ≤ .001.

Table 17. Likelihood Ratio Tests.

Model Comparisons	D.F.	Chi-square Statistic	Probability
Model 1b vs. Model 2b	1	0.01	0.93
Model 1b vs. Model 3b	2	2.02	0.36
Model 1b vs. Model 1a	1	0.01	0.93
Model 1b vs. Model 2a	2	0.02	0.99
Model 1b vs. Model 3a	3	2.31	0.51

Step 3: Parametric Approach: Estimating the dose-response function to discern treatment effect

Utilizing the coefficients calculated in the previous step, the final step of the GPS approach involves estimating the dose-response function to discern treatment effects. Figures 16 and 17 display the average predicted values of logged violent crime rate across each level of physical disorder. In particular, Figure 16 displays models with an interaction term ("a" model type), while Figure 17 displays models without an interaction term ("b" model type). Both figures display the dose-response function with and without confidence intervals. For ease of comparison, Figure 18 jointly displays the dose-response function produced by each model.

Figure 17. Interaction Models: Dose-response Functions across Model Specifications.

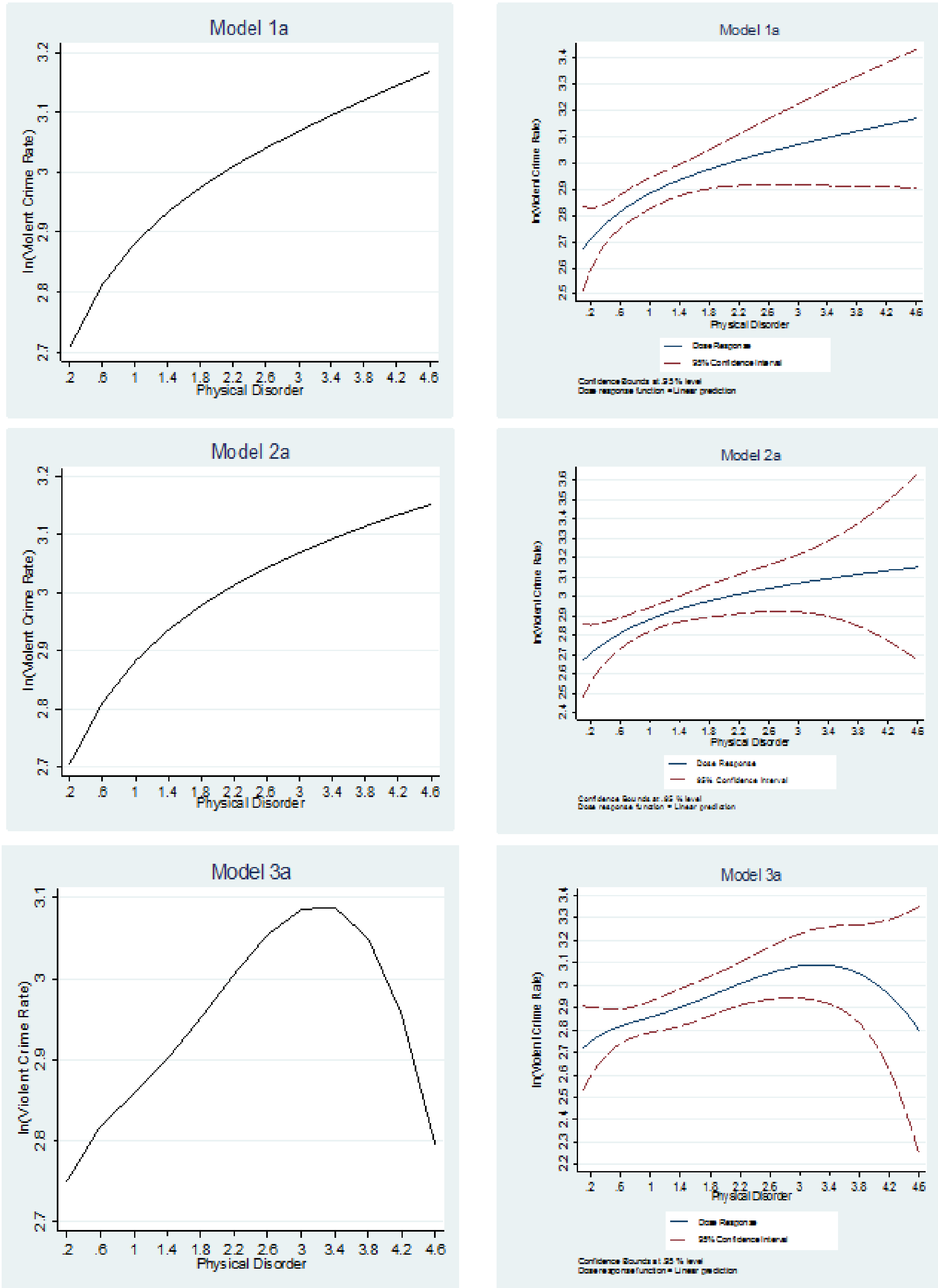


Figure 18. Noninteraction Models: Dose-response Functions across Model Specifications.

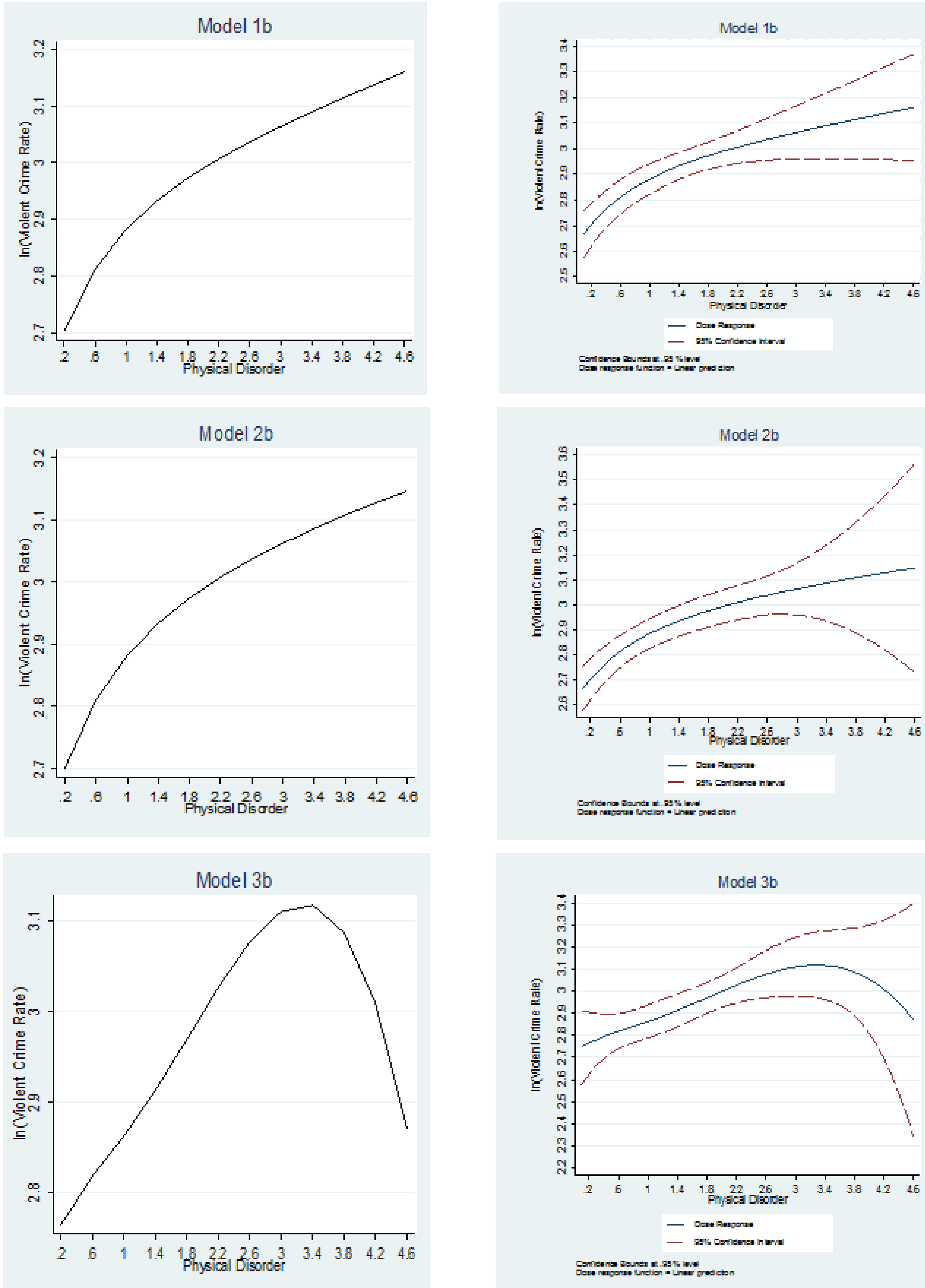
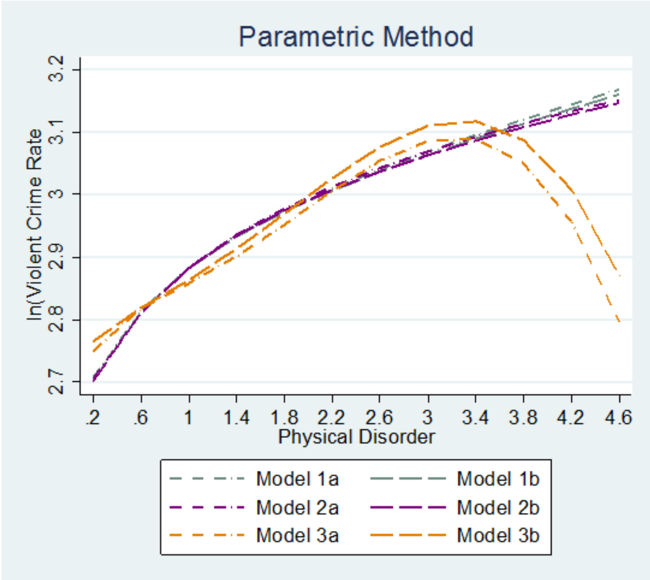


Figure 19. Combined Display of Parametric Methods.



Across all models, the relationship between physical disorder and logged violent crime rate showcases linearity. With the exception of models 3a and 3b, physical disorder maintains a positive relationship with logged violent crime rate across all levels of physical disorder. This relationship is exhibited by a steep rise in logged violent crime rate at low levels of physical disorder, followed thereafter by a steady, approximately linear increase at a treatment level of $\sim$ 1.00. Models 3a and 3b also showcase linearity at low and mid-range levels of physical disorder, as seen by a steady, positive increase in logged violent crime rate. The rate of increase, however, is slower than in models 1a, 1b, 2a, and 2b. Furthermore, the relationship between physical disorder and logged violent crime rate levels off at a treatment level of $\sim$ 3.00, and then swiftly drops. This finding is suggestive of a potential inoculation effect, whereby the severity of exposure to disorder is lessened over time as individuals adapt to their surroundings (Taylor & Shumaker, 1990; Sampson & Raudenbush, 2004).

Other commonalities across models include tight confidence intervals at low and mid-range levels of physical disorder. This is where the majority of the data lies. The widening of the confidence intervals at high levels of physical disorder exposes values with limited data. Excluding the highest levels of physical disorder from consideration, a positive, linear trend is still apparent.

Overall, the dose-response function is somewhat robust to alterations in model specifications. Within model types 1, 2 and 3, the inclusion of an interaction term did little to change the dose-response function. This finding may be due to the fact that in all cases the interaction term was insignificant. There is also substantial overlap between models 1a, 1b, 2a, and 2b (see Figure 18). Focusing on model 1b (i.e., the best fitting model), the relationship between physical disorder and logged violent crime rate does not resemble a threshold effect. Thus, no support is found for hypotheses 2a, 2b, or 3. While not entirely linear, physical disorder maintains a linear relationship with violent crime rate across a substantial portion of its distribution. For this reason, hypothesis 1, which maintains that disorder has a positive, linear relationship with logged violent crime rate, garners more support than hypothesis 4, which supports nonlinearity.

Semiparametric Method

Following the parametric GPS approach, the dose-response function was estimated using three types of semiparametric methods: penalized spline, radial spline, and inverse weighting kernel function.

Step 1: Semiparametric Approaches: Modeling the conditional distribution of the treatment given covariates

Step 1 of the GPS approach is the same for both parametric and semiparametric estimations of the dose-response function.

Step 2: Semiparametric Approaches: Estimating the conditional expectation of the outcome given the treatment and GPS

With the previously identified common support sample and estimated GPS, step 2 of the GPS approach was conducted utilizing penalized spline regressions that included additive spline bases and radial basis functions (see Bia et al., 2014). For brevity, the former regression approach is referred to as the penalized spline model (or method), while the latter is referred to as the radial spline model (or method). Estimates from the penalized spline and radial spline models are presented in Tables 18 and 19, respectively. Reviewing these tables, there are several factors to consider. To start, both physical disorder (*p-value* $\leq$ .001) and GPS (*p-value* $\leq$ .001) are statistically significant in the penalized spline model. However, only the GPS (*p-value* $\leq$ .05) is statistically significant in the radial spline model. In addition, the likelihood ratio test is statistically significant for the penalized spline model (*Prob > Chi-square* = 0.0062), but only marginally so for the radial spline model (*Prob > Chi-square* = 0.1056). In the context of mixed effects models, the likelihood ratio test compares the fit of the evaluated mixed effect model to a standard regression model (i.e., the reduced model) that does not include random effects parameters. A statistically significant chi-square statistic suggests that the mixed effect model improves model fit over the reduced model. The likelihood ratio tests conducted for the penalized spline and radial spline models use the same reduced model for comparison. Therefore, these tests can be compared to shed light on the superior approach: penalized spline or radial spline. A statistically significant chi-square statistic for the penalized spline model suggests that the inclusion of random effect parameters improves model fit. The chi-square statistic for the radial spline model is only marginally significant. Therefore, the penalized spline model emerges as superior.[16]

Table 18. Penalized Spline Model.

Fixed Effects	Beta	SE		
Physical Disorder	0.13***	0.03		
GPS	0.71**	0.27		
Constant	2.54***	0.11		
Random Effects			95% Confidence Interval	
Knot Physical Disorder	6.01e-07	1.20e06	1.20e-08	3.04e-05
Knot GPS	0.59	0.35	0.18	1.91

Likelihood Ratio Test: Chi-square (2) = 10.18, Prob > Chi-square = 0.0062
+p-value≤ .10; * p-value ≤ .05; ** p-value ≤.01; *** p ≤ .001

Table 19. Radial Spline Model.

Fixed Effects	Beta	SE		
Physical Disorder	0.09	0.11		
GPS	-0.19*	0.09		
Constant	0.09	0.17		
Random Effects	Beta	SE	95% Confidence Interval	
Knots for Physical Disorder & GPS	0.06	0.04	0.01	0.22

Likelihood Ratio Test: Chi-square (2) = 1.56, Prob > Chi-square = 0.1056
+ p-value ≤ .10; * p-value ≤ .05; ** p-value ≤.01; *** p-value ≤ .001

Step 3: Semiparametric Approaches: Estimating the dose-response function to discern treatment effects

The coefficients calculated from the penalized spline and radial spline models were then used to estimate the dose-response function. The dose-response function was also estimated utilizing an inverse weighting kernel approach, with an optimal bandwidth (*bw* = 0.32) selected from Fan and Gijbels' (1996) proposed procedure (see Flores et al., 2012). Recall Figure 19 displays the dose-response function estimated from each approach shown with and without

[16] Shedding light on this finding, the standard error associated with the *physical disorder* coefficient of the radial spline model displayed in Table 19 is almost four times higher than the corresponding standard error of the penalized spline model displayed in Table 18.

confidence intervals. For ease of comparison, Figure 20 jointly displays the dose-response function produced from each method.

Figure 20. Semiparametric Methods.

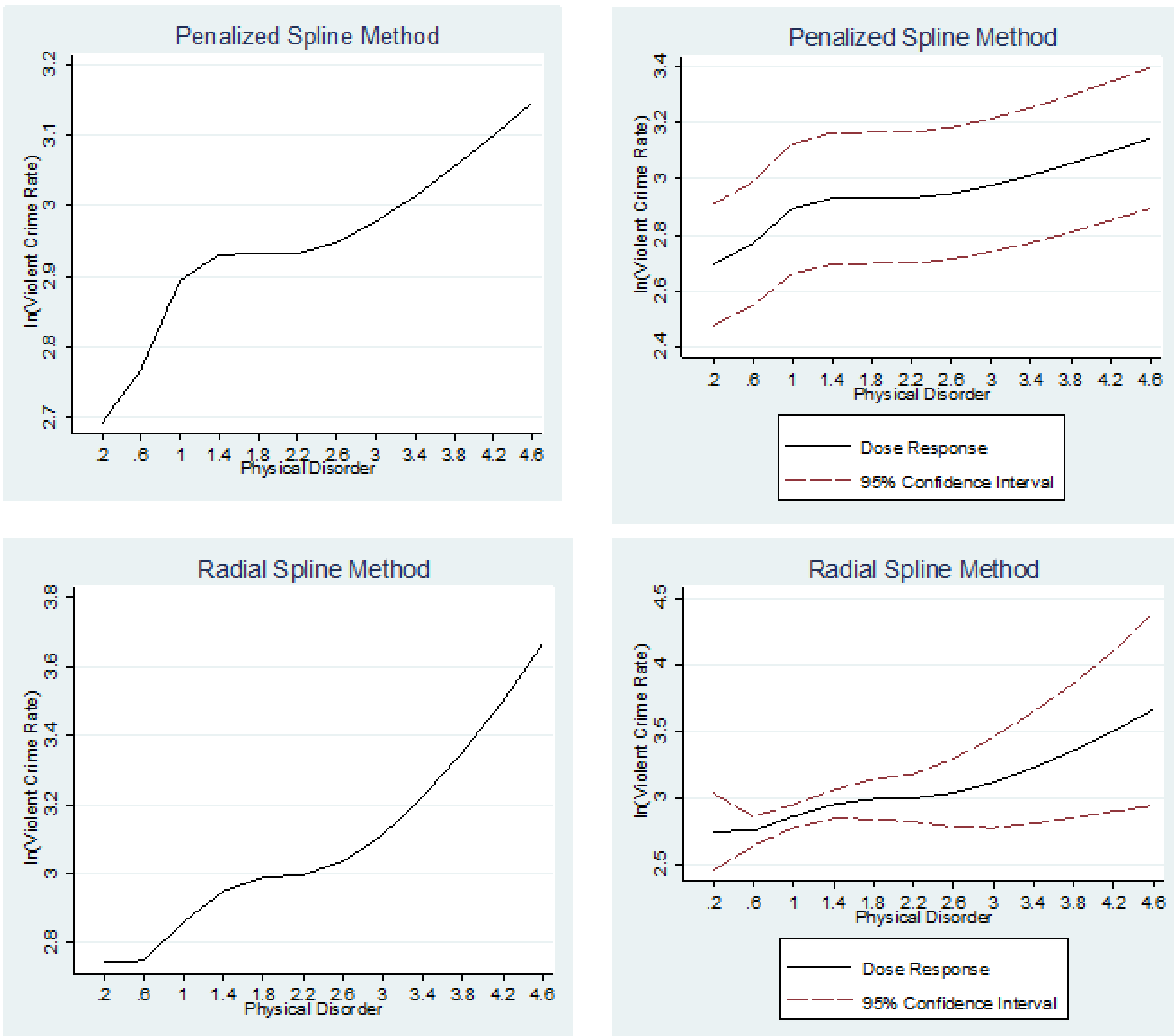


Figure 20 (cont'd)

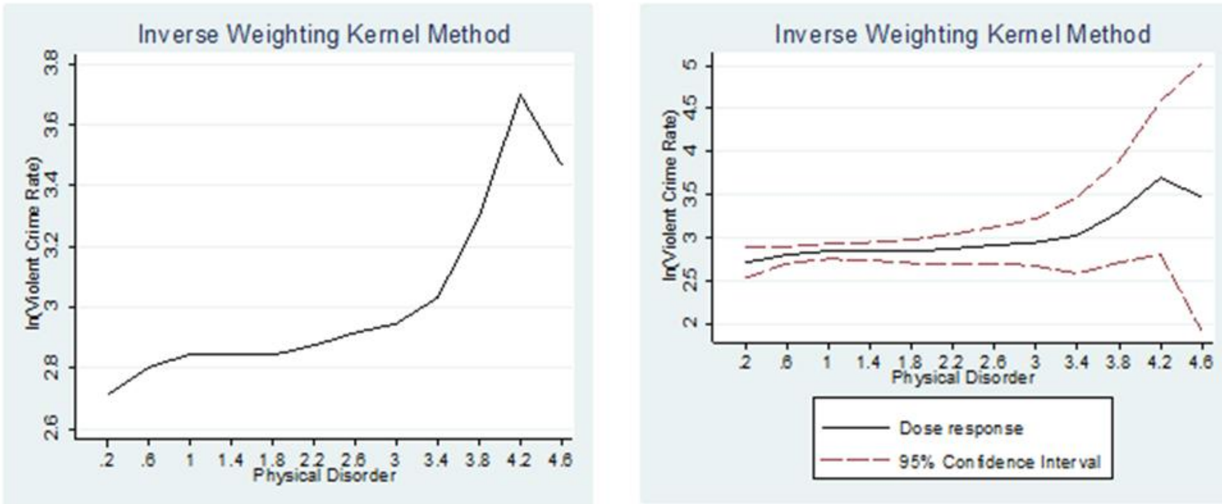
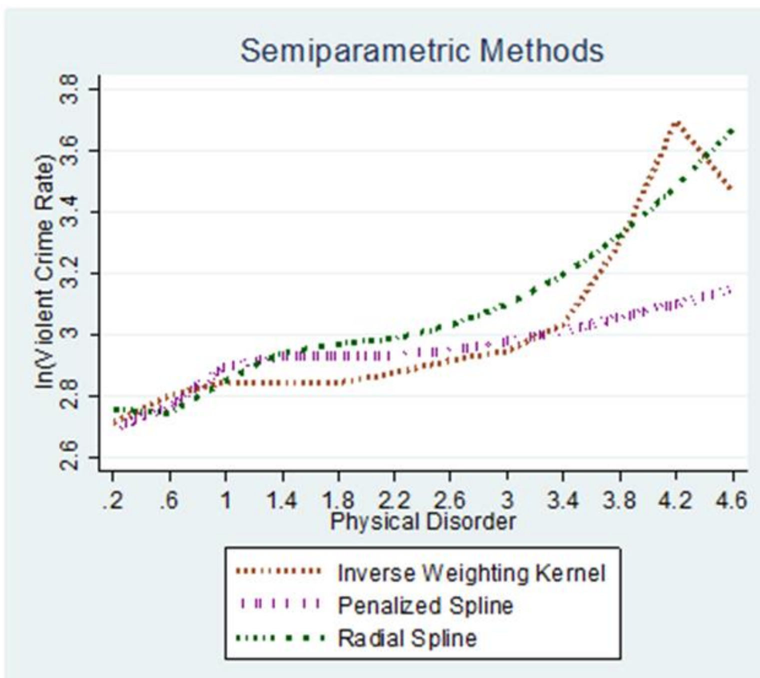


Figure 21. Combined Display of Semiparametric Methods.



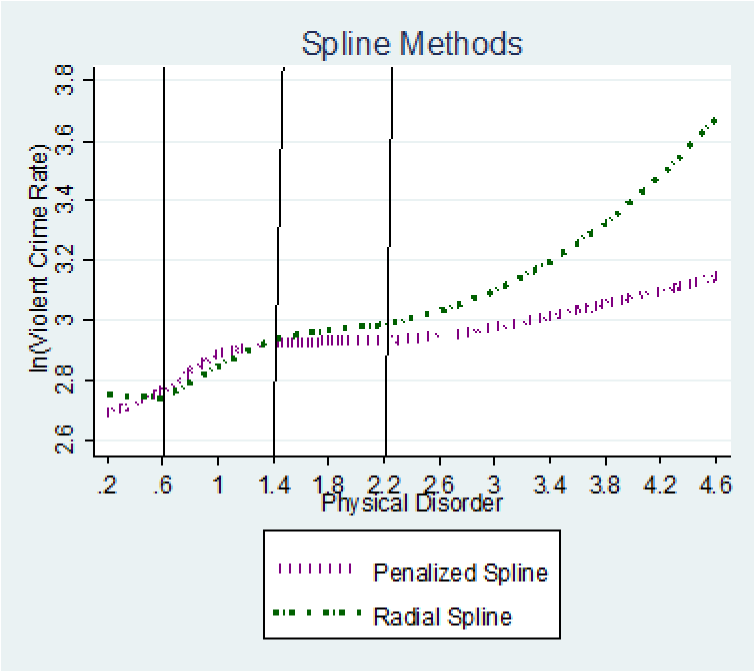
As indicated by more instances of nonlinearity, it is immediately apparent that the semiparametric methods do indeed allow for greater flexibility in the estimation of the dose-response function than the previously presented parametric method. Therefore, there is clear support for hypothesis 4, indicating nonlinearity. Another clear distinction between these methods is seen in their estimation of confidence intervals. The semiparametric methods generate much wider confidence intervals than the parametric method, especially at high levels of physical disorder. This finding is expected. The structure provided by parametric estimators allows extrapolation from regions in which data are abundant to regions in which data are scarce (Bia et al., 2014). Nonparametric methods are not afforded the same luxury. Therefore, estimates generated from limited support data are done so with a greater degree of uncertainty.

Recall the penalized spline model fit the data better than the radial spline model. However, this method to generate the dose-response function does so with a much greater degree of uncertainty, indicated by wide confidence intervals across *all* levels of physical disorder. For this reason, estimates produced from the penalized spline method must be interpreted with more caution.

Proceeding with caution, the penalized spline method identifies a positive relationship between physical disorder and logged violent crime rate at low levels of physical disorder. This relationship is shown by a modest rate of increase. In comparison, the radial spline method does not identify physical disorder to have an effect on logged violent crime rate at very low levels. A relationship does not emerge until a treatment level of ~ 0.60, indicated by an uptick in logged violent crime rate. However, it is important to take heed of the wide confidence intervals at very low levels of physical disorder, suggesting that the radial spline estimates produced at these levels must be interpreted with greater caution.

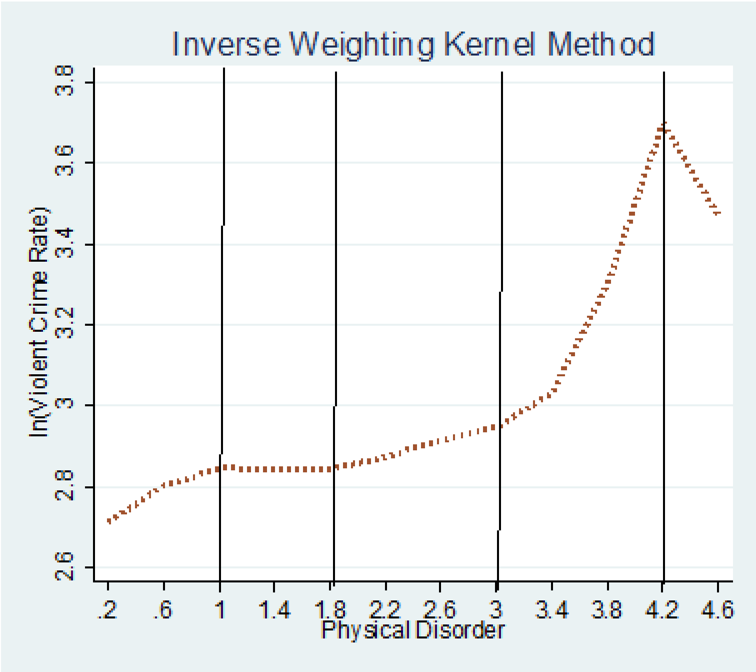
After a treatment level of ~ 1.40, the penalized spline method indicates that physical disorder has no effect on logged violent crime rate. Logged violent crime rate remains relatively constant until a treatment level of ~ 2.20. Past this level, the penalized spline method shows logged violent crime rate steadily increasing with physical disorder. In contrast, the radial spline method shows logged violent crime rate increasing with physical disorder after a treatment level of ~ 1.40, although at a much slower rate than before. Past a treatment level of ~ 2.20, the rate of increase substantially increases. Estimates produced beyond this level far exceed those produced by the penalized spline method. Figure 21 jointly displays the spline estimates and identifies the discussed treatment levels by vertical black lines.

Figure 22. Spline Methods with Relevant Treatment Levels Highlighted.



Focusing now on the inverse weighting kernel method, logged violent crime rate increases at low levels of physical disorder at a modest and steady rate. Similar to the penalized spline method, the relationship between physical disorder and logged violent crime rate is relatively constant between the treatment levels of ~ 1.00 and ~ 1.80. Past these levels, the inverse weighting kernel method shows logged violent crime rate steadily increasing, picking up speed at a treatment level of ~ 3.00. This change in rate occurs later than shown in either spline method. Reaching a precipice at a treatment level of ~ 4.20, the relationship between physical disorder and logged violent crime rate abruptly drops. This drop distinguishes the inverse weighting kernel method from its spline counterparts, and is suggestive of a potential inoculation effect. However, it is again important to remember that estimates produced at high levels of physical disorder are done so with a greater degree of uncertainty. Figure 22 jointly displays the inverse weighting kernel estimates and identifies the discussed treatment levels by vertical black lines.

Figure 23. Inverse Weighting Kernel Method with Relevant Treatment Levels Highlighted.



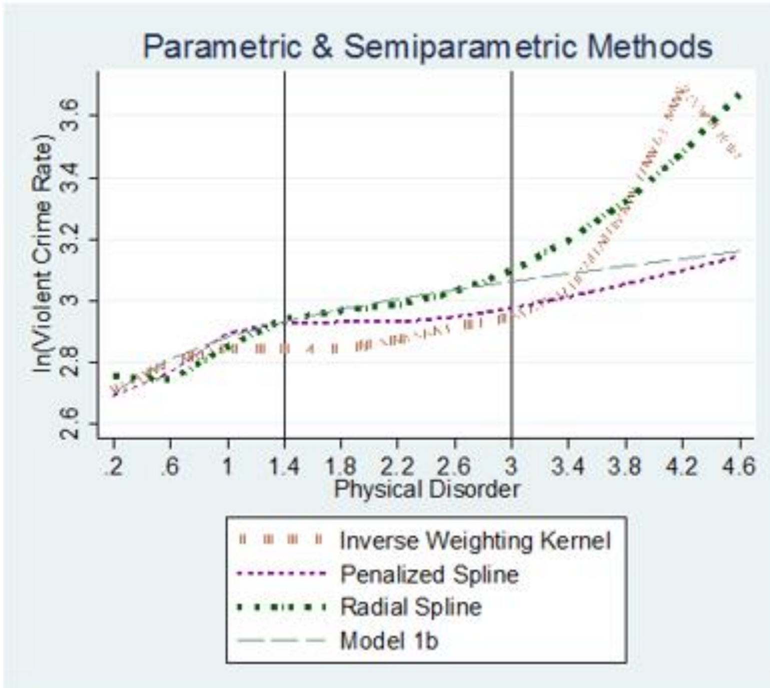
Overall, a relatively similar image of the dose-response function emerges across semiparametric approaches. To summarize, logged violent crime rate rises at a modest rate at low levels of physical disorder. Transitioning from low levels of physical disorder, the strength of the positive relationship between physical disorder and logged violent crime rate is either greatly reduced or becomes nonexistent. Past some level located in the middle of the physical disorder distribution, the rate of increase picks up. Excluding very high levels of physical disorder from consideration (i.e., those with the widest confidence intervals), this general pattern is still apparent.

This description of the relationship between physical disorder and logged violent crime rate closely parallels the broken windows tipping point, with two caveats. First, there is not a dramatic break at mid-range levels of physical disorder. Although there is an increased change in rate, the transition is smoother than originally expected. In other words, there is an *attenuated threshold effect*. Second, a reduced effect at mid-range levels of physical disorder was also not expected. In spite of these differences, the semiparametric approaches lend support in favor of hypothesis 2a: the broken windows tipping point as a threshold effect.

Summary of Findings

To review, Figure 23 displays the dose-response function produced from the best fitting parametric method (i.e., model 1b), alongside those produced from semiparametric methods. Vertical black lines demarcate low, mid-range, and high levels of physical disorder, with less consideration given to *very* high levels due to wide confidence intervals. Across methods, the relationship between physical disorder and logged violent crime rate is quite similar at low levels of physical disorder. Previously thought to be steep, the increase in logged violent crime rate observed in model 1b at low levels of physical disorder is quite modest, demonstrating the importance of scale in interpreting results. Furthermore, model 1b closely follows the radial spline method at mid-range levels of physical disorder, but diverges at high levels. To this point, clear divergences between methods can be seen at high levels of physical disorder, with the inverse weighting kernel method predicting the greatest amount of crime, followed by the radial spline method, model 1b, and the penalized spline method.

Figure 24. Parametric & Semiparametric Methods.



At each level of physical disorder, intersubjective agreement across methodological approaches boosts confidence in this study's findings. Overall, however, the dose-response function was not consistently estimated across parametric and semiparametric methods. The method and the assumptions that underlie it influenced the estimation of the dose-response function. The inconsistencies that emerged demonstrate the importance of considering the potential sensitivities of each methodological approach and their impact on estimation.

Complementing Figure 16, Table 20 re-states this study's hypotheses alongside their level of support. Strictly speaking, all of the identified relationships are nonlinear. However, labeling all of these relationships in this way masks intricacies that are revealed upon closer examination across the distribution of physical disorder. To this point, model 1b exposes a predominantly linear relationship between physical disorder and logged violent crime rate at mid-range and high levels of physical disorder, lending support to hypothesis 1. Furthermore, the semiparametric methods all identify an increased change in rate at mid-range levels of physical disorder. The largest change is shown by the inverse weighting kernel method, followed by the radial spline and penalized spline methods. However, these changes are more gradual than expected to be considered true threshold effects. At the very least, however, these findings lend partial support in favor of hypothesis 2a.

Table 20. Hypotheses & Support.

Hypotheses	Support
H1: Physical disorder maintains a positive linear effect on violent crime rate.	▪ Model 1b ○ Mid-range & high levels of physical disorder
H2a: Physical disorder maintains a threshold effect on violent crime rate such that small variations exert a modest positive effect on violent crime rate at low levels, and a dramatic positive effect past a critical level located somewhere between low and high levels.	▪ Penalized Spline Method, Radial Spline Method, & Inverse Weighting Kernel Method ○ Increased change of rate at mid-range levels of physical disorder
H2b: Physical disorder maintains a threshold effect on violent crime rate such that small variations exert a modest positive effect on violent crime rate at low and high levels, and a dramatic positive effect past a critical level located somewhere between these two extremes.	▪ No Support
H3: Physical disorder maintains a threshold effect on violent crime rate such that small variations exert a modest positive effect on violent crime rate at low and mid-range levels, and a dramatic positive effect past a critical level located somewhere at high levels.	▪ No Support
H4: Physical disorder maintains a nonlinear effect on violent crime rate.	▪ Model 1b, Penalized Spline Method, Radial Spline Method, & Inverse Weighting Kernel Method

CHAPTER 5: DISCUSSION & CONCLUSION

An Overview: The Search for The Broken Windows Tipping Point

Wilson and Kelling (1982) provide a simple instruction for the implementation of order-maintenance policing: direct limited police resources to the *broken windows tipping point*. In doing so, they imply a certain functional form of the relationship between disorder and violent crime. That is, Wilson and Kelling's (1982) description of the tipping point suggests that the disorder-crime relationship is best captured as a threshold effect: the impact of disorder on violent crime dramatically increases beyond some critical level of disorder located at mid-range levels. If this is indeed the case, then a proper test of the validity of BWT should accommodate nonlinearity. To this point, misspecification of the functional form of the disorder-crime relationship obscures tests of validity and does not advance criminological theory.

With few exceptions, broken windows research has ignored the tipping point, modeling the disorder-crime relationship as linear (e.g., Skogan, 1990; Sampson & Raudenbush, 1999; Harcourt, 2001; Eck & Maguire, 2005; Steenbeek & Kreis, 2015; Wheeler, 2018; Konkel et al., 2019). In an interesting twist, the few studies that evaluate this and similar phenomenon provide reason to doubt Wilson and Kelling's (1982) interpretation. For example, Crane's (1991) and Raleigh and Galster's (2015) findings suggest that the tipping point may be located at high levels of disorder, while Geller's (2007) finding suggests that the tipping point may not resemble a threshold effect at all. While mixed, this body of research does support a *nonlinear* relationship between disorder and violent crime, underscoring the importance of efforts to accommodate nonlinearity. Furthermore, this finding has implications for policing disorder initiatives. A nonlinear relationship between disorder and violent crime suggests that some neighborhoods may be more or less amenable to these initiatives than others. If so, police resources should be allocated in such a way as to have the most optimal effect on crime. To this point, Wilson and Kelling (1982) prioritize preventing future crime emergence by directing limited police resources to the tipping point. While this approach is contentious, their instruction suggests that a proper test of the effectiveness of policing disorder initiatives involves implementing them at the tipping point. Beyond their ability to reduce violent crime, evaluations of the effectiveness of these initiatives should include nontraditional metrics, such as their ability to improve residents' quality of life, reduce fear of crime, strengthen police-community relations, and avoid the financial and social costs associated with future violence.

This study empirically examined the functional form of the relationship between physical disorder and violent crime rate as a *first step* toward identifying the broken windows tipping point. Great strides were taken to accommodate nonlinearity by adopting a methodological approach that allows flexibility in modeling decisions, while also allowing for the identification of causal effects. In this regard, the generalized dose-response propensity score (i.e., the GPS method) was perfectly suited. This approach explicitly models the functional form of the disorder-crime relationship at each level of physical disorder while addressing selection effects through covariate balancing and consists of three key steps: 1) Modeling the conditional distribution of the treatment given covariates; 2) Estimating the conditional expectation of the outcome given the treatment and GPS; and 3) Estimating the dose-response function to discern treatment effects. To facilitate its analysis, this study utilized block-group level data on physical disorder, violent crime, and socioeconomic and land use characteristics from the DPD's record management system, MCM project, and Census. As part of its sensitivity checks, this study explored how slight alterations to the specification of the conditional expectation function and nonparametric techniques affected the estimation of the dose-response function. Despite its comprehensive analysis, the functional form of the disorder-crime relationship remains unclear. That being said, the bulk of the evidence favors a nonlinear relationship, with partial support for Wilson and Kelling's (1982) interpretation of the broken windows tipping point.

Directions for Future Research

This study found considerable support in favor of a nonlinear relationship between physical disorder and logged violent crime rate. However, this finding is far from definitive. Additional research is needed to establish whether the disorder-crime relationship is truly nonlinear. There are four important domains in which research can be developed: 1) Measures of disorder; 2) Confounding factors; 3) Neighborhood context; and 4) Longitudinal data analysis.

Measures of Disorder

Research suggests that disorder is socially constructed (Harcout, 2001; Sampson & Raudenbush, 2004; Hinkle & Yang, 2014). Turning to BWT for insight, an emphasis is placed on perceptions of disorder, rather than objective measures thereof. Residents must *perceive* disorder to be a problem within their neighborhoods and respond fearfully for the broken windows development sequence to unfold. Unlike the case of social disorder, there is considerable overlap between objective and perceived measures of physical disorder, suggesting that either is appropriate for examinations of the disorder-crime relationship (Perkins et al., 1992; Sampson & Raudenbush, 2004; Hinkle & Yang, 2014; Yang & Pao, 2015; Ren et al., 2019). Nonetheless, future research should prioritize perceived measures of disorder in examinations of the disorder-crime relationship, especially when social disorder is considered. Building upon this instruction, the allocation of police resources may be better informed by evaluations of the relationship between social disorder and violent crime. The reason being that the police often play a larger role in addressing social nuisances than the physical conditions of the environment.

That being said, variation in police responsibilities across departments is expected. In the case of Detroit, for example, the police play a substantial role in efforts to address physical disorder.

In addition, indicators used to measure disorder may differentially influence residents' perceptions that disorder is a problem. For example, Franzini et al. (2008) found that perceptions of disorder are more strongly influenced by severe, long-lasting indicators, such as abandoned properties, than by those that can be more easily rectified, such as trash and graffiti. For this reason, they argue that efforts to address disorder should focus on the former rather than the latter indicators in order to have the desired effect on crime (Franzini et al., 2008). Complementing this finding, other studies have found the disorder-crime relationship to be stronger for some classifications of disorder than for others (e.g., O'Brien & Sampson, 2015; O'Brien et al., 2015; Wheeler, 2018; Konkel et al., 2019). For these reasons, future research should give particular attention to the indicators that comprise disorder and how they shape residents' perceptions, as well as how disorder is classified and the unique contributions of these classifications to explaining crime.

Neighborhood Context

Residents' perceptions of disorder are shaped by observable cues of disorder, as well as neighborhood social structure. Generally, individuals (of all races) perceive higher levels of disorder in predominately poor, minority neighborhoods (Sampson & Raudenbush, 2004; McCord et al., 2007; Hipp, 2010; Sampson, 2012; Wickes et al., 2013). That being said, white individuals generally perceive more disorder than minorities living in the same neighborhood (Sampson & Raudenbush, 2004; Franzini et al., 2008; Hipp, 2010; Sampson, 2012). Sampson and Raudenbush (2004) offer an explanation for this finding: if minority residents have a greater past exposure to disorder, then they may have a higher threshold for perceiving disorder to be a problem. Together, these findings suggest that the disorder-crime relationship may vary across neighborhood contexts, specifically across racial/ethnic lines. In predominately minority neighborhoods, greater levels of disorder may be needed than in predominantly white neighborhoods to elicit a fear response, inciting the broken windows cycle.

Relatedly, evaluating the *average* effect of disorder on violent crime across levels of disorder may mask differences that exist across neighborhood contexts that affect the functional form of the disorder-crime relationship. To be clear, BWT makes a global statement about the relationship between disorder and violent crime; the process through which disorder influences violent crime is the same across all neighborhoods. The consideration of average effects is aligned with this framing. In light of new knowledge since the advent of BWT, however, future studies should explore the role of neighborhood context in shaping the disorder-crime relationship, perhaps creating neighborhood typologies based upon indicators correlated with disorder, such as concentrated disadvantage (Sampson & Raudenbush, 2004; Wilcox et al., 2004; Gau & Pratt, 2010). Furthermore, such evaluations may help flesh out the effect of disorder on violent crime at high levels of disorder, which this study estimated with much uncertainty. Attention should also be given to how neighborhood context affects the development of informal social control, and the barriers within neighborhoods that serve to undermine police-community partnerships needed for the implementation of order-maintenance policing.

Confounding Factors

The GPS method estimates causal relationships by controlling the effect of known confounding factors (i.e., factors that affect selection into treatment and treatment-specific outcomes). The exclusion of such factors results in omitted variable bias. Across regression approaches, omitted variable bias impairs the identification of causal effects. In the case of the

GPS method, the omission of confounding factors effects the extent to which confidence can be placed in the estimation of the dose-response function.

As it relates to BWT, collective efficacy has been shown to mitigate the relationship between disorder and violent crime across a variety of neighborhood contexts (e.g., Sampson et al., 1997; Browning et al., 2004; Reisig & Cancino, 2004; Sampson, 2004; Wells et al., 2006; Warner, 2007; Mazerolle, Wickes, & McBroom, 2010; Maxwell, Garner, & Skogan, 2011; Swatt et al., 2013). In other words, it is a known confounding factor. In the absence of formal measures of collective efficacy, the current study utilized a proxy: percent gardens/parks.

Within neighborhoods and crime research, collective efficacy has been traditionally defined as the willingness of residents to intervene to maintain order, coupled with trust and solidarity amongst residents (Sampson et al. 1997; Browning et al., 2004; Mazerolle et al., 2010; Sampson, 2013). Sampson (2013) has discussed the merits and shortcomings of various measures of collective efficacy. He ultimately concluded that collective efficacy is a theory of process "involving shared expectations about order and control, activated ties, and acts of informal control. How these concepts are measured and interrelate will vary depending on the research context" (Sampson, 2013, p. 20).

Community gardens and, to a lesser extent, parks provide opportunities for the development of collective efficacy, and signify neighborhood investments (Cohen et al., 2008; Teig et al., 2009; Clayton, 2007; Kearney, 2009; Alaimo et al., 2010). However, they do not capture residents' expectations of control and their ability to activate social ties to bring about neighborhood change, key components of collective efficacy. Future evaluations of the functional form of the disorder-crime relationship should strive to include formal measures of collective efficacy that capture all of its dimensions in order to bolster confidence in causal inferences.

Another confounding factor to consider involves the way in which the police are deployed to neighborhoods. The relationship between disorder and violent crime will be impacted if the police are deployed based upon a neighborhood's level of disorder. For example, the attenuated threshold effect observed by this study may be an artifact of deployment strategy if it so happens that the police are disproportionally deployed to neighborhoods with mid-range levels of disorder, as compared to low or high levels. In the case of Detroit, however, routine police patrol primarily focuses on neighborhoods that have high population densities and violent crime rates. For this reason, this study included a measure of population density, as well as a temporal lag of violent crime rate to address the confounding factor of police deployment.

Adding another layer of complexity, the level of police commitment to order-maintenance policing activities is a related confounding factor. Across Detroit, three to five NPOs are deployed to every SCA. As previously discussed, NPOs responsibilities are consistent with order-maintenance policing. However, the extent to which they engage in these activities is unknown. If NPOs or patrol officers, for that matter, are more likely to engage in order-maintenance policing activities in neighborhoods with mid-range levels of disorder, then the attenuated threshold effect observed by this study is suspect. Absent the eradication of the police, systematic differences in police deployment and order-maintenance policing activities across neighborhoods will pose issues for evaluations that seek to investigate the causal relationship between disorder and violent crime. For this reason, future evaluations should strive to account for these differences when possible. This effort is especially important for evaluations in which functional form is of interest.

Longitudinal Data Analysis

Longitudinal data are best equipped to study the disorder-crime relationship given the process through which Wilson and Kelling (1982) argue disorder affects violent crime. Compared to cross-sectional data, longitudinal data are better-suited to tease out causal relationships and the processes that underlie them, such as how informal social control develops and declines over time, as well as the role of neighborhood context in shaping this process. In light of this study's focus, the advantages of longitudinal data are important insofar as they affect analyses of the functional form of the disorder-crime relationship. Quite obviously, if disorder does not *cause* violent crime, then explorations of the functional form of the disorder-crime relationship are meaningless. In the absence of longitudinal data, this study attempted to mitigate the issue of causality by adopting a methodological approach that addressed selection effects, when possible, through covariate balancing across matched levels of physical disorder.

Setting aside the issue of causality, cross-sectional data require that we assume a neighborhood's stage of progression in the broken windows cycle based upon its *current* level of disorder. Following the logic of BWT, the deeper entrenched a neighborhood is in this cycle, the higher its level of disorder, and therefore the higher its level of violent crime. Operating under this assumption, the functional form of the disorder-crime relationship can be determined if cross-sectional evaluations include neighborhoods across *all* stages of decline. Compared to other levels, this study included the fewest neighborhoods with very high levels of physical disorder, presumedly those that are the deepest entrenched in the broken windows cycle. Therefore, the disorder-crime relationship was estimated with more uncertainty at these levels. Future research should strive to collect sufficient data across all levels of disorder in order to enhance confidence in the estimation of causal effects.

As previously discussed, there is reason to believe that average effects mask nuances that exist across neighborhood contexts which affect the functional form of the disorder-crime relationship. Methodological techniques that are able to capture developmental patterns as they unfold over time can help shed light on the extent to which neighborhood context shapes the disorder-crime relationship, as well as the dynamics that underlie it. One technique that shows promise is the dual GBTM. An extension of GBTM, the dual GBTM was designed to capture the relationship between two related but distinct development trajectories, such as disorder and violent crime, which evolve contemporaneously or over different time periods (Nagin & Tremblay, 2001). Several renowned scholars within the field of Criminology have raised concerns regarding the existence of distinct developmental trajectories and the extent to which units within trajectories adhere to them (see Sampson & Laub, 2005; Raudenbush, 2005). In response to this critique, other approaches, such as growth mixture modeling (GMM) and nonparametric growth mixture modeling (NP-GMM), have become popular. Unlike GBTM, GMM and NP-GMM include random effects in the estimation of trajectory models which allow for within-group variability (Nagin & Odgers, 2010). Future research should explore these (and other) modeling alternatives in an effort to establish intersubjective agreement.

Closing Remarks

Despite the need for future research, several implications for theory, practice, and policy can be tentatively drawn from this study. To start, this study suggests that broken windows research should accommodate nonlinearity in its exploration of the relationship between disorder and violent crime. Misspecification of the functional form of this relationship might cause researchers to over- or under-state effects depending on the nature of this relationship across levels of disorder. Therefore, efforts to accommodate nonlinearity not only stand to improve model fit, but also to provide a more accurate assessment of the disorder-crime relationship. Likewise, efforts to gauge the effectiveness of policing disorder initiatives should be mindful of this relationship and manage their expectations for crime control accordingly.

There are several possible options that researchers can pursue to accommodate nonlinearity. Known for their simplicity, polynomial transformations may be an entirely reasonable away to capture nonlinear effects. However, researchers must be aware of their shortcomings. Previously discussed, polynomial transformations assume that the relationship between *X*, the independent variable, and *Y*, the dependent variable, do not vary across the distribution of *X*. In other words, they force researchers to assume a global fit. Another shortcoming of this approach is that the selection of polynomial transformation is often arbitrary. While some theories may indicate a nonlinear effect, the actual power of the effect is often not clearly known and the incorrect selection may obfuscate results (see Keele, 2008).

In light of these shortcomings, researchers have turned to nonparametric techniques. These techniques do not require a priori assumptions about functional form, but rather locally estimate it from the data. Given this feature, nonparametric techniques are well-suited for theory testing. This study conducted local estimation using penalized spline, radial spline, and inverse weighting kernel methods. However, these methods are by no means the only available to capture nonlinear effects. To this point, machine learning systems, such as neural networks or tree-based models, are able to implicitly detect complex nonlinear relationships through an automated process that learns from the (data) environment and applies changes to improve predictions. However, they come at a cost. Known as “black box” approaches, the internal logic of machine learning systems is often unclear (see Rudin & Carlson, 2018). Most concerning, it is not well understood how variables contribute to the model or how to interpret model results (see

Rudin & Carlson, 2018). For these reasons, these computationally intensive methods are less aligned with theory testing which seeks to establish causality.

In addition, this study found a fair amount of support for Wilson and Kelling's (1982) interpretation of the broken windows tipping point, hypothesis 2a. Across semiparametric methods, an increased change in logged violent crime rate was identified at mid-range levels of physical disorder. However, this change is not nearly as severe as expected. Ultimately, a slower ascent provides reason to reconsider Wilson and Kelling's (1982) interpretation of the mechanisms that underlie the tipping point, as well as their instruction for the implementation of order-maintenance policing.

In neighborhoods located at the tipping point, recall Wilson and Kelling (1982) advocate the minimal use of formal mechanisms to address disorder in order to avoid its proliferation and precipitous rise of violent crime. While neighborhoods located at the tipping point may indeed have weakened levels of informal social control, their effect on violent crime may not be as significant as once thought. It may be the case that Wilson and Kelling (1982) overstated the effect of disorder on residents' fear of crime. To this point, there may be another tipping point at play that affects the disorder-crime relationship. As previously discussed, past exposure to disorder may increase the threshold for perceiving it to be a problem (Taylor & Shumaker, 1990; Sampson & Raudenbush, 2004). On average, minority residents are more likely to have been previously exposed to disorder and therefore will need to be exposed to greater levels before they perceive it to be a problem. In light of Detroit's predominantly low income, African American neighborhoods, this phenomenon may help explain the attenuated threshold effect identified by this study. It may also help explain the negative effect of disorder on violent crime at very high levels of disorder identified by the inverse kernel weighting method, and models 3a and 3b.[17] If future research consistently uncovers similar relationships between disorder and violent crime, then there is a strong basis for investigating its underlying causes.

Furthermore, Wilson and Kelling's (1982) instruction to focus police resources at neighborhoods at the tipping point was in part motivated by the practical limitations of allocating *limited* police resources to high disorder, high crime neighborhoods. In such neighborhoods, Wilson and Kelling (1982) argue that the demands to police resources would come at too great a cost. Consequently, the best (and only) option is to focus police resources at neighborhoods at the tipping point; neighborhoods that are at the brink of decline. Contrary to what was expected, however, this study's findings suggest that neighborhoods are *not* catapulted into a high disorder, high crime state past some level of disorder located in the middle of the disorder distribution. Rather, this study found an attenuated threshold effect. As a result, Wilson and Kelling's (1982) instruction loses significance. There may be more opportunities for residents to strengthen and exercise informal social control without needing the assistance of the police, as the impact of fear of crime may not be as debilitating as previously thought. There may also be more opportunities for the police to intervene *beyond* mid-range levels of disorder without great cost. Ultimately, there is less motivation to heed Wilson and Kelling's (1982) instruction without the looming threat of the tipping point.

The magnitude of the effect of disorder on violent crime is also worth consideration, especially past mid-range levels of disorder. Across parametric and semiparametric approaches, the magnitude of the effect of physical disorder on logged violent crime rate is diverse, with the inverse weighting kernel method predicting the largest effect and the penalized spline method

[17] Minority individuals are the most likely to have extensive previous exposure to disorder, as well as to live in neighborhoods with very high levels of disorder (and violent crime).

predicting the smallest. Although they do not accommodate the possibility of nonlinearity, other studies have identified a modest effect of disorder on violent crime (e.g., Sampson & Raudenbush, 1999; Taylor, 1999, 2001; Boggess & Maskaly, 2014; Wheeler, 2018; Konkel et al., 2019). Nonetheless, efforts to address disorder may still be worth pursuing. Historically, they have played a key role in policies aimed at spurring neighborhood revitalization (e.g., Newman, 1972; Brown & Perkins, 2001; Brown, Brown, & Perkins, 2004; Day et al., 2007; Dulin-Keita et al., 2015; Schuetz, Spader, & Cortes, 2016; Spader et al., 2016; Prener, Braswell, & Monit, 2020; Rupp et al., 2020). Outside of their effect on crime, efforts to address disorder may also improve residents' quality of life and reduce fear of crime (Skogan, 1990; Perkins & Taylor, 1996; Day et al., 2007; Chappell, Monk-Turner, & Payne, 2010; Dulin-Dulin-Keita et al., 2015; Johnsen, Neal, & Gasteyer, 2015; Rupp et al., 2020).

While it is premature to cast aside BWT, this study's findings bolster a competing instruction for the allocation of police resources. It has been repeatedly demonstrated that a small fraction of targets (e.g., places, victims, and offenders) account for the vast majority of crime (Wolfgang et al., 1972; Forst et al., 1978; Sherman et al., 1989; Weisburd et al., 2004; Eck et al., 2007). Termed the "power few," Sherman (2007) argues that resources should be concentrated to targets that produce the greatest amount of harm in order to have the largest crime reduction effect.[18] Consistent with this finding, this study suggests that the allocation of police resources to hot spots may be preferable to Wilson and Kelling's (1982) instruction to direct police resources to neighborhoods at the tipping point. Indeed, hot spot policing has been shown to achieve significant crime reduction gains (Braga et al., 1999; Braga & Bond, 2008; Braga et al., 2012, 2014). That being said, Sherman (2007, p. 308) warns that "[p]ower few targets," such as hot

[18] Sherman (2017, p. 13) calls for the creation of a total harm index, whereby each classification of crime "would be based on the extent of injury or ripple effects of injury."

spots, “are anything but the proverbial ‘low-hanging fruit’ that is easiest to harvest. The power few may, in fact, be the hardest nuts to crack: the cases that are most difficult to solve because they have so many simultaneous or ‘co-morbid’ problems.”

The issue of comorbidity presents a particularly complex challenge to tackle. To this point, the effect of hot spot policing on crime, however large, has yet to be shown to produce *long-term* crime reduction gains (see Telep & Weisburd, 2014). One potential explanation concerns the extenuating role of economic disadvantage and collective efficacy, co-morbid features of places. These features have been found to contribute to the developmental patterns of disorder and crime (Weisburd et al., 2012, 2013). In light of this finding, it has been argued that crime prevention strategies which focus on initiating social change within hot spots are better equipped to produce long-term crime reduction gains, as compared to policing strategies driven by opportunity theories (Weisburd et al., 2012; Weisburd, Davis, & Gill, 2015). In this regard, order-maintenance policing holds promise. Order-maintenance policing can be considered to be a future-oriented policing strategy. That is, it is geared toward *preventing* future crime emergence by targeting neighborhoods that are on the cusp of decline; neighborhoods at the tipping point. In order to achieve this goal, proponents of order-maintenance policing and, more generally, BWT have long advocated the role of the police in strengthening informal social control within neighborhoods, a key component of collective efficacy.

According to Weisburd et al. (2015, p. 272), police efforts that seek to promote collective efficacy “…will require (especially in large agencies) a shift from the myopic focus of crisis response to a bifurcated approach that allows space for community-building efforts at hot spots.” To facilitate these efforts, the police must establish relationships of trust with residents. The recent murder of George Floyd by a Minneapolis police officer has shed much needed light on the growing chasm between the police and the communities which they serve. Deeply negative attitudes toward the police - amplified by widely publicized incidents of police brutality - feed this divide. Moving forward from this horrific event, procedural justice in police-resident interactions that are supported by community-oriented policing strategies, such as order-maintenance policing, are paramount in order to restore residents' trust in the police.[19] In an effort to secure long-term crime reduction gains, Weisburd et al. (2015, p. 269) argue that these strategies should be implemented within hot spots and extended to emphasize "the direct impact of everyday police intervention on informal social control and structuring a concrete approach for building community engagement and collective efficacy." They reason that such an approach will help empower residents to take responsibility for crime within their communities and self-regulate safety, ultimately decreasing reliance on the police (Weisburd et al., 2015). This approach aligns with the recent call to support *alternative* solutions to address social problems within communities in an effort to minimize reliance on the police (see Chang & Wilson, 2020; Hawkins, Mettler, & Stein, 2020).

In addition to the challenge of comorbidity, evaluations of hot spot policing suffer from the same challenges of evaluations conducted elsewhere (see Sherman, 2007). One such challenge that is particularly relevant to this study regards determining the appropriate dosage of police response needed to ensure the identification of valid treatment effects. As previously mentioned, co-morbid features of hot spots pose significant challenges for police efforts to spur social change. Unfortunately, little is known about the dosage of police response necessary to

[19] Broadly defined, community-oriented policing "is a philosophy that promotes organizational strategies that support the systematic use of partnerships and problem-solving techniques to proactively address the immediate conditions that give rise to public safety issues such as crime, social disorder, and fear of crime" (Community Oriented Police Services, 2020, p. 1).

achieve long-term crime reduction gains in hot spots, or elsewhere, for that matter. In particular, it is unknown whether long-term crime reduction gains can be achieved in hot spots with *limited police resources*. For this reason, Wilson and Kelling's (1982) instruction to allocate limited police resources to neighborhoods at the tipping point - in an effort to avoid the future potential of increased levels of violence - remains a tenable alternative.

In light of this study's findings, there are likely more places *beyond* mid-range levels of disorder in which order-maintenance policing could be implemented without great cost to police resources. Nonetheless, efforts to elicit social change in these places will likely come much easier than similar efforts implemented in hot spots, as the challenges posed by co-morbid features will likely not be as significant. That being said, if *short-term* crime reduction gains are the priority, then the allocation of limited police resources to hot spots emerges as the superior strategy.[20] Ultimately, a comprehensive assessment of the optimal allocation of police resources is necessary. In addition to their potential to reduce crime, this analysis should compare each strategy's potential to improve residents' quality of life, reduce fear of crime, strengthen police-community relations, and avoid the financial and social costs of future violence.

[20] The logic here being that limited police resources can be spatially distributed in such a way as to produce a high spatial dosage of these resources.

REFERENCES

REFERENCES

Alaimo, K., Reischl, T. M., & Allen, J. O. (2010). Community gardening, neighborhood meetings, and social capital. *Journal of Community Psychology, 38*(4), 497-514.

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.

Andresen, M. A., & Hodgkinson, T. K. (2018). Evaluating the impact of police foot patrol at the micro-geographic level. *Policing: An International Journal, 41*(3), 314-324.

Armstrong, T. A., Katz, C. M., & Schnebly, S. M. (2015). The relationship between citizen perceptions of collective efficacy and neighborhood violent crime. *Crime & Delinquency, 61*(1), 121-142.

Baker, T. E., and Wolfer, L. (2003). The crime triangle: Alcohol, drug use, and vandalism. *Police Practice on Research: An International Journal*, *4*(1), 47–61.

Berk, R., & MacDonald, J. (2010). Policing the homeless: An evaluation of efforts to reduce homeless-related crime. *Criminology and Public Policy, 9*(4), 813-840.

Bernasco, W., & Block, R. (2009). Where offenders choose to attack: A discrete choice model of robberies in Chicago. *Criminology, 47,* 93-130

Bia, M., Flores, C. A., Flores-Lagunes, A., & Mattei, A. (2014). A Stata package for the application of semiparametric estimators of dose-response functions. *Stata Journal, 14*(3), 580-604.

Bittner, E. (1967). Police discretion in emergency apprehension of mentally ill persons. *Social Problems*, *14*(3), 278-292.

Blight Removal Task Force Plan. (2014). *Every neighborhood has a future…And it doesn't include blight*. https://assets.documentcloud.org/documents/1173946/detroit-blight-removal-task-force-plan-may-2014.pdf

Boggess, L. N., & Hipp, J. R. (2010). Violent crime, residential instability and mobility: Does the relationship differ in minority neighborhoods? *Journal of Quantitative Criminology, 26*(3), 351-370.

Boggess, L. N., & Maskaly, J. (2014). The spatial context of the disorder–crime relationship in a study of reno neighborhoods. *Social Science Research, 43,* 168-183.

Bowers, W., & Hirsch, J. (1987). The impact of foot patrol staffing on crime and disorder in Boston: An unmet promise. *American Journal of Police, 6*, 17-44.

Braga, A. A., & Bond, B. J. (2008). Policing crime and disorder hot spots: A randomized controlled trial. *Criminology, 46*(3), 577-607.

Braga, A. A., Hureau, D. M., & Papachristos, A. V. (2011). The relevance of micro places to citywide robbery trends: A longitudinal analysis of robbery incidents at street corners and block faces in Boston. *Journal of Research in Crime and Delinquency, 48*(1), 7-32.

Braga, A. A., Papachristos, A. V., & Hureau, D. (2012). The effects of hot spots policing on crime. Campbell Systematic Reviews, *8*(8), 1-96.

Braga, A. A., Papachristos, A. V., & Hureau, D. (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly, 31*(4), 633-663.

Braga, A. A., Weisburd, D. L., Waring, E. J., Mazerolle, L. G., Spelman, W., & Gajewski, F. (1999). Problem-oriented policing in violent crime places: A randomized controlled experiment. *Criminology, 37*(3), 541-580.

Braga, A. A., Welsh, B. C., & Schnell, C. (2015). Can policing disorder reduce crime? A systematic review and meta-analysis. *Journal of Research in Crime and Delinquency, 52*(4), 567-588.

Branas, C. C., South, E., Kondo, M. C., Hohl, B. C., Bourgois, P., Wiebe, D. J., & MacDonald, J. M. (2018). Citywide cluster randomized trial to restore blighted vacant land and its effects on violence, crime, and fear. *Proceedings of the National Academy of Sciences of the United States of America, 115*(12), 2946-2951.

Bratton, W., Kelling, G. (2006, February 28). There are no cracks in the broken windows. National Review. https://www.nationalreview.com/2006/02/there-are-no-cracks-broken-windows-william-bratton-george-kelling/

Brown, M. K. (1981). *Working the street: Police discretion and the dilemmas of reform*. New York, NY: Russell Sage Foundation.

Brown, B. B., & Altman, I. (1983). Territoriality, defensible space and residential burglary: An environmental analysis. *Journal of Environmental Psychology, 3*(3), 203-220.

Brown, B. B., & Perkins, D. D. (2001*). Final Project Report: Neighborhood revitalization & disorder: An intervention evaluation.* National Institute of Justice, Washington, DC.

Brown, G., Brown, B. B., & Perkins, D. D. (2004). New housing as neighborhood revitalization: Place attachment and confidence among residents. *Environment and Behavior, 36*(6), 749-775.

Browning, C. R. (2002). The span of collective efficacy: Extending social disorganization theory to partner violence. *Journal of Marriage and Family*, *64*(4), 833-850.

Browning, C. R., Calder, C. A., Soller, B., Jackson, A. L., & Dirlam, J. (2017). Ecological networks and neighborhood social organization. *AJS; American Journal of Sociology, 122*(6), 1939-1988.

Browning, C. R., Feinberg, S. L., & Dietz, R. D. (2004). The paradox of social organization: Networks, collective efficacy, and violent crime in urban neighborhoods. *Social Forces*, *83*(2), 503-534.

Bursik, R. J. (1988). social disorganization and theories of crime and delinquency: Problems and prospects. *Criminology*, *26*(4), 519-552.

Bursik, R. J. (1999). The informal control of crime through neighborhood networks. *Sociological Focus: Community, Inequality and Crime*, *32*(1), 85-97

Bursik, R., & Grasmick, H. G. (1993). *Neighborhoods and crime: The dimensions of effective community control*. New York; Toronto: Lexington Books.

Cerdá, M., Tracy, M., Messner, S. F., Vlahov, D., Tardiff, K., & Galea, S. (2009). Misdemeanor policing, physical disorder, and gun-related homicide: A spatial analytic test of "broken-windows" theory. *Epidemiology*, *20*(4), 533-541.

Chang, S., & Wilson, A. (2020). Opinion: Police should partner with social workers. *Detroit News*. https://www.detroitnews.com/story/opinion/2020/06/23/opinion-police-should-partner-social-workers/3235618001/

Chappell, A. T., Monk-Turner, E., & Payne, B. K. (2011). Broken windows or window breakers: The influence of physical and social disorder on quality of life. *Justice Quarterly*, *28*(3), 522-540.

Clarke, P., Ailshire, J., Melendez, R., Bader, M., & Morenoff, J. (2010). Using Google Earth to conduct a neighborhood audit: Reliability of a virtual audit instrument. *Health & Place*, *16*(6), 1224-1229.

Clayton, S. (2007). Domesticated nature: Motivations for gardening and perceptions of environmental impact. *Journal of Environmental Psychology*, *27*, 215-224.

Cohen, D.A., Inagami, S., & Finch, B. (2008). The built environment and collective efficacy. *Health and Place*, *14*, 198–208.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*, S95-S120.

Coleman, J. S. (1990). *Foundations of Social Theory*. Cambridge, MA: The Belknap Press of Harvard University Press.

Community Oriented Policing Services. (2020). *Community policing defined*. U.S. Department of Justice, Washington, D.C.

Cordner, G. (1986). Fear of crime and the police: An evaluation of a fear-reduction strategy. *Journal of Police Science and Administration, 14*(3), 223-233.

Covington, J. M., & Taylor, R. B. (1991). Fear of crime in urban residential neighborhoods: Implications of between- and within-neighborhood sources for current models. *Sociological Quarterly*, *32*(2), 231-249.

Cowell, B. M., & Kringen, A. L. (2016). *Engaging communities one step at a time: Policing's tradition of foot patrol as an innovative community engagement strategy*. Police Foundation, Washington, D.C.

Crane, J. (1991). The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing. *American Journal of Sociology, 96*(5), 1226-1259.

Crank, J. P., Giacomazzi, A., & Heck, C. (2003). Fear of crime in a nonurban setting. *Journal of Criminal Justice, 31*(3), 249-263.

Curman, A.S., Andresen, M. A., & Brantingham, P. J. (2015). Crime and place: A longitudinal examination of street segment patterns in Vancouver, BC. *Journal of Quantitative Criminology, 31*(1), 127-147.

Cutrona, C. E., Russell, D. W., Hessling, R. M., Brown, P. A., & Murry, V. (2000). Direct and moderating effects of community context on the psychological well-being of African American women. *Journal of Personality and Social Psychology, 79*(6), 1088-1101.

Motor City Mapping. (2020). Survey Key. https://motorcitymapping.org/survey-key

Day, K., Anderson, C., Powe, M., McMillan, T., & Winn, D. (2007). Remaking minnie street: The impacts of urban revitalization on crime and pedestrian safety. *Journal of Planning Education and Research, 26*(3), 315-331.

Detroit Public Safety Foundation. (2013). Community policing pilot program reduces crime by double digits in first twelve months. https://media4.manhattan-institute.org/pdf/detroit_one_year_press_release.pdf

Dulin-Keita, A., Clay, O., Whittaker, S., Hannon, L., Adams, I. K., Rogers, M., & Gans, K. (2015). The influence of HOPE VI neighborhood revitalization on neighborhood-based physical activity: A mixed-methods approach. *Social Science & Medicine, 139*, 90-99.

Dunham, R. G., & Alpert, G. P. (1988). Neighborhood differences in attitudes toward policing: Evidence for a mixed-strategy model of policing in a multi-ethnic setting. *The Journal of Criminal Law and Criminology. 79*(2), 504-523.

Eck, J. E., & Guerette, R. T. (2012). Place-based crime prevention: theory, evidence, and policy. In D.P. Farrington & B.C. Welsh (Eds.), *The Oxford Handbook of Crime Prevention* (pp. 354-383). New York: Oxford University Press.

Eck, J. E., & Maguire, E. R. (2005). Have changes in policing reduced violent crime? An assessment of the evidence. (pp. 207-265) Cambridge University Press.

Eck, J., Clarke, R., & Guerette, R. (2007). Risky facilities: Crime concentrations in homogeneous sets of establishments and facilities. *Crime Prevention Studies, 21*, 255–264.

Eilers, P. H., & Marx, B.D. (1996). Flexible smoothing with b-splines and penalties. *Statistical Science, 11(*2), 89-121.

Esbensen, F., & Taylor, T. (1984). Foot patrol and crime rates. *Southern Journal of Criminal Justice, 8,* 184-194.

ESRI. (2020). Using Areal Interpolation to perform polygon-to-polygon predictions. https://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/using-areal-interpolation-to-predict-to-new-polygons.htm

Fan, J., & Gijbels, I. (1996). *Local polynomial modelling and its applications* (1st ed.). London; New York: Chapman & Hall.

Flores, C. A., Flores-Lagunes, A., Gonzalez, A., & Neumann, T. C. (2012). Estimating the effects of length of exposure to instruction in a training program: The case of job corps. *The Review of Economics and Statistics, 94*(1), 153-171.

Fong, C., Hazlett, C., & Imai, K. (2018). Covariate balancing propensity score for a continuous treatment: application to the efficacy of political advertisements. *Annals of Applied Statistics, 12*(1), 156-177.

Forst, B., Lucianovic, J., & Cox, S. (1978). *What happens after arrest? A court perspective of police operations in the District of Columbia.* Washington, DC: INSLAW.

Foster-Fishman, P. G., Cantillon, D., Pierce, S. J., & Van Egeren, L. A. (2007). Building an active citizenry: The role of neighborhood problems, readiness, and capacity for change. *American Journal of Community Psychology, 39*, 91-106.

Fox, J. (2000). *Nonparametric simple regression: Smoothing scatterplots*. Thousand Oaks, CA: Sage.

Franzini, L., Caughy, M. O., Nettles, S. M., & O'Campo, P. (2008). Perceptions of disorder: Contributions of neighborhood characteristics to subjective perceptions of disorder. *Journal of Environmental Psychology, 28*(1), 83-93.

French, J., Kammann, E., & Wand, M. (2001). Comment on paper by Ke and Wang. *Journal of American Statistical Association, 96*, 1285-1288.

Galster, G. C. (2014). Nonlinear and threshold aspects of neighborhood effects. *KZfSS Kölner Zeitschrift Für Soziologie Und Sozialpsychologie, 66*(S1), 117-133.

Galster, G. C. (2018). Nonlinear and threshold effects related to neighborhood: Implications for planning and policy. *Journal of Planning Literature, 33*(4), 492-508.

Garofalo, J. (1981). The fear of crime: Causes and consequences. *The Journal of Criminal Law and Criminology, 72*(2), 839-857.

Garvin, E. C., Cannuscio, C. C., & Branas, C. C. (2013). Greening vacant lots to reduce violent crime: A randomised controlled trial. *Injury Prevention, 19*(3), 198-203.

Gau, J. M. (2014). Procedural justice and police legitimacy: A test of measurement structure. *American Journal of Criminal Justice, 39*(2), 187-205.

Gau, J. M., & Brunson, R. K. (2010). Procedural justice and order maintenance policing: A study of inner-city young men's perceptions of police legitimacy. *Justice Quarterly, 27*(2), 255-279.

Gau, J. M., & Pratt, T. C. (2008). Broken windows or window dressing? citizens'(in) ability to tell the difference between disorder and crime. *Criminology & Public Policy, 7*(2), 163-194.

Gau, J. M., & Pratt, T. C. (2010). Revisiting broken windows theory: Examining the sources of the discriminant validity of perceived disorder and crime. *Journal of Criminal Justice, 38*(4), 758-766.

Geller, A., (2007). Neighborhood disorder and crime: An analysis of broken windows in New York city. Unpublished PH.D. dissertation. New York: Columbia University.

Gelman, A., Fagan, J., & Kiss, A. (2007). An analysis of the New York City police department's "stop-and-frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association, 102*(479), 813-823.

Gill, C., Wooditch, A., & Weisburd, D. (2017). Testing the "Law of crime concentration at place" in a suburban setting: Implications for research and practice. *Journal of Quantitative Criminology, 33*(3), 519-545.

Groff, E. R., Ratcliffe, J. H., Haberman, C. P., Sorg, E. T., Joyce, N. M., & Taylor, R. B. (2015). Does what police do at hot spots matter? The Philadelphia policing tactics experiment. *Criminology, 53*(1), 23-53.

Guardabascio, B., & Ventura, M. (2014). Estimating the Dose–Response function through a generalized linear model approach. *The Stata Journal: Promoting Communications on Statistics and Stata, 14*(1), 141-158.

Hanink, P. (2013). don't trust the police: Stop question frisk, compstat, and the high cost of statistical over-reliance in the NYPD. *Journal of the Institute of Justice and International Studies,* (13),1-18.

Harcourt, B. E. (2001). *Illusion of order: The false promise of broken windows policing*. Cambridge, Mass: Harvard University Press.

Harcourt, B. E., & Ludwig, J. (2006). Broken windows: New evidence from New York City and a five-city social experiment. *The University of Chicago Law Review, 73*(1), 271-320.

Hawkins, D., Mettler, K., & Stein, P. (2020). 'Defund the police' gains traction as cities seek to respond to demands for a major law enforcement shift. *The Washington Post*. https://www.washingtonpost.com/nation/2020/06/07/protests-defund-police/

Heckman, J. J., & Robb, R. (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics, 30*(1), 239-267.

Hill, T. D., Ross, C. E., & Angel, R. J. (2005). Neighborhood disorder, psychophysiological distress, and health. *Journal of Health and Social Behavior, 46*(2), 170-186.

Hinds, L. (2007). Building Police-Youth relationships: The importance of procedural justice. *Youth Justice, 7*(3), 195-209.

Hinkle, J. C., & Weisburd, D. (2008). The irony of broken windows policing: A micro-place study of the relationship between disorder, focused police crackdowns and fear of crime. Journal *of Criminal Justice, 36*(6), 503-512.

Hinkle, J. C., & Yang, S. (2014). A new look into broken windows: What shapes individuals' perceptions of social disorder? *Journal of Criminal Justice, 42*(1), 26-35.

Hinkle, J. C., Weisburd, D., Famega, C., & Ready, J. (2013). The problem is not just sample size: The consequences of low base rates in policing experiments in smaller cities. *Evaluation Review, 37*(3-4), 213-238.

Hipp, J. R., Kim, Y., & Kane, K. (2019). The effect of the physical environment on crime rates: Capturing housing age and housing type at varying spatial scales. *Crime & Delinquency, 65*(11), 1570-1595.

Hirano, K., & Imbens, G.W. (2004). The propensity score with continuous treatments. In D. Rubin, A. Gelman and X.L. Meng (Eds.), *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives* (pp. 73 – 84). West Sussex, England: Wiley.

Ingoldsby, E. M., & Shaw, D. S. (2002). Neighborhood contextual factors and early-starting antisocial pathways. *Clinical Child and Family Psychology Review*, *5*(1), 21-55.

Innes, M. (2004). Signal crimes and signal disorders: Notes on deviance as communicative ac tion. *The British Journal of Sociology*, *55*(3), 335-355.

Jacobs, J. (1961). *The death and life of great American cities*. New York: Vintage Books.

Jay, J., Miratrix, L. W., Branas, C. C., Zimmerman, M. A., & Hemenway, D. (2019). Urban building demolitions, firearm violence and drug crime. *Journal of Behavioral Medicine*, *42*(4), 626-634.

Jeffery, C. R. (1971). *Crime prevention through environmental design*. Beverly Hills, CA: Sage.

Johansen, R., Neal, Z., & Gasteyer, S. (2015). The view from a broken window: How residents make sense of neighbourhood disorder in flint. *Urban Studies*, *52*(16), 3054-3069.

Kane, R. J., & Cronin, S. W. (2013). Associations between order maintenance policing and violent crime: Considering the mediating effects of residential context. *Crime & Delinquency*, *59*(6), 910-929.

Kasarda, J. D., & Janowitz, M. (1974). Community attachment in mass society. *American Sociological Review*, *39*, 328-39.

Katz, C. M., Webb, V. J., & Schaefer, D. R. (2001). An assessment of the impact of quality-of-life policing on crime and disorder. *Justice Quarterly*, *18*(4), 825-876.

Kearney, S. C. (2009). *The community garden as a tool for community empowerment: A study of community gardens in Hampden County* (Unpublished doctoral dissertation or master's thesis). University of Massachusetts Amherst.

Keele, L. (2008). *Semiparametric regression for the social sciences*. GB: Wiley.

Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science*, *322*(5908), 1681-1685.

Kelling, G. L. (1974). *The Kansas City preventive patrol experiment: A technical report*. Washington: Police Foundation.

Kelling, G. L. (1981). *The Newark Foot Patrol Experiment*. Washington, DC: The Police Foundation.

Kelling, G. L., & Coles, C. M. (1996). *Fixing broken windows: Restoring order and reducing crime in our communities* (First Touchstone ed.). New York: Simon & Schuster.

Kelling, G. L., & Sousa, W. H. (2001). Do police matter?: An analysis of the impact of New York City's police reforms. New York, NY: CCI Center for Civic Innovation at the Manhattan Institute.

Kleinhans, R., & Bolt, G. (2014). More than just fear: On the intricate interplay between perceived neighborhood disorder, collective efficacy, and action. *Journal of Urban Affairs, 36*(3), 420-446.

Klingler, D. A., & Bridges, G. S. (1997). Measurement error in calls-for-service as an indicator of crime. *Criminology, 35*(4), 705-726.

Kluve, J., Schneider, H., Uhlendorff, A., & Zhao, Z. (2012). Evaluating continuous training programmes by using the generalized propensity score. *Journal of the Royal Statistical Society. Series A, Statistics in Society, 175*(2), 587-617

Kitts, J. A. (1999). Not in our backyard: Solidarity, social networks, and the ecology of environmental mobilization. *Sociological Inquiry*, *69*(4), 551-574.

Kondo, M. C., Hohl, B., Han, S., & Branas, C. (2016). Effects of greening and community reuse of vacant lots on crime. *Urban Studies*, *53*(15), 3279-3295.

Kondo, M. C., Andreyeva, E., South, E. C., MacDonald, J. M., & Branas, C. C. (2018). Neighborhood interventions to reduce violence. *Annual Review of Public Health, 39*(1), 253-271.

Konkel, R. H., Ratkowski, D., & Tapp, S. N. (2019). The effects of physical, social, and housing disorder on neighborhood crime: A contemporary test of broken windows theory. *ISPRS International Journal of Geo-Information, 8*(12), 583-602.

Koper, C. S. (1995). Just enough police presence: Reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. *Justice Quarterly, 12*(4), 649-672.

Kreif, N., Grieve, R., Díaz, I., & Harrison, D. (2015). Evaluation of the effect of a continuous treatment: A machine learning approach with an application to treatment for traumatic brain injury. *Health Economics*, *24*(9), 1213-1228.

Krivo, L. J., Velez, M. B., Lyons, C. J., & Phillips, J. B. (2018). Race, crime and the changing fortunes of urban neighborhoods, 1999-2013. *Du Bois Review: Social Science Research on Race. 15*(1), 47-68.

Kubrin, C. E., & Weitzer, R. (2003). New directions in social disorganization theory. *Journal of Research in Crime and Delinquency, 40*(4), 374-402.

LaGrange, R. L., Ferraro, K. F., & Supancic, M. (1992). Perceived risk and fear of crime: Role of physical and social incivilities. *Journal of Research in Crime and Delinquency, 29*(3), 311-334.

Larson, M., Xu, Y., Ouellet, L., & Klahm, C. F. (2019). Exploring the impact of 9398 demolitions on neighborhood-level crime in Detroit, Michigan. *Journal of Criminal Justice, 60*, 57-63.

Leroux, E. J., & McShane, K. (2017). Changing youth attitudes toward the police through community policing programming. *Journal of Community Psychology, 45*(6), 810-822.

Loftin, C., & McDowall, D. (2010). The use of official records to measure crime and delinquency. *Journal of Quantitative Criminology, 26*(4), 527-532.

Long, D. A., & Perkins, D. D. (2007). Community social and place predictors of sense of community: A multilevel and longitudinal analysis. *Journal of Community Psychology, 35*(5), 563–581.

Loughran, T. A., Wilson, T., Nagin, D. S., & Piquero, A. R. (2015). Evolutionary regression? assessing the problem of hidden biases in criminal justice applications using propensity scores. *Journal of Experimental Criminology, 11*(4), 631-652.

Lowenkamp, C. T., Cullen, F. T., & Pratt, T. C. (2003). Replicating Sampson and Groves's test of social disorganization theory: Revisiting a criminological classic. *Journal of Research in Crime and Delinquency, 40*(4), 351-373.

MacDonald, J. M., & Stokes, R. J. (2020). Gentrification, land use, and crime. *Annual Review of Criminology, 3*(1), 121-138.

Markowitz, F. E., Bellair, P. E., Liska, A. E., & Lui, J. (2001). Extending social disorganization theory: Modeling the relationships between cohesion, disorder, and fear. *Criminology, 39*(2), 293-319.

Maxwell, C. D., Garner, J. H., & Skogan, W. G. (2011). *Collective efficacy and criminal behavior in Chicago, 1995-2004*. Shepherdstown, WV: Joint Center for Justice Studies, Inc.

Mazerolle, L., Wickes, R., & McBroom, J. (2010). Community variations in violence: The role of social ties and collective efficacy in comparative context. *Journal of Research in Crime and Delinquency, 47*(1), 3–30.

McGarrell, E. F., Giacomazzi, A. L., & Thurman, Q. C. (1997). Neighborhood disorder, integration, and the fear of crime. *Justice Quarterly, 14*(3), 479-500.

McGarrell, E. F., Giacomazzi, A. L., & Thurman, Q. C. (1999). Reducing disorder, fear, and crime in public housing: A case study of place-specific crime prevention. *Justice Research and Policy*, *1*(2), 61-87.

Mears, D. P., Cochran, J. C., & Beaver, K. M. (2013). Self-control theory and nonlinear effects on offending. *Journal of Quantitative Criminology, 29*(3), 447-476.

Messner, S. F., Galea, S., Tardiff, K. J., Tracy, M., Bucciarelli, A., Piper, T. M., . . . Vlahov, D. (2007). Policing, drugs, and the homicide decline in New York City in the 1990s. *Criminology, 45*(2), 385-414.

Millie, A. (2008). Anti-social behaviour, behavioural expectations and an urban aesthetic. *The British Journal of Criminology, 48*(3), 379-394.

Mooney, S. J., DiMaggio, C. J., Lovasi, G. S., Neckerman, K. M., Bader, M. D., Teitler, J. O., & Rundle, A. G. (2016). Use of Google Street View to assess environmental contributions to pedestrian injury. *American Journal of Public Health*, *106*(3), 462-469.

Morenoff, J. D., Sampson, R. J., & Raudenbush, S. W. (2001). Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology, 39*(3), 517-558.

Nagin, D. S., & Odgers, C. L. (2010). Group-based trajectory modeling in clinical research. *Annual Review of Clinical Psychology, 6*(1), 109-138.

Nagin, D. S., & Tremblay, R. E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods, 6*(1), 18-34.

Newman, O. (1972). *Defensible space: Crime prevention through urban design*. New York: Macmillan.

Newman, O. (1996). Creating defensible space. Washington, DC: U.S. Department of Housing and Urban Development.

O’Brien, D. T., & Sampson, R. J. (2015). Public and private spheres of neighborhood disorder: Assessing pathways to violence using large-scale digital records. *Journal of Research in Crime and Delinquency*, *52*(4), 486-510.

O'Brien, D. T., Farrell, C., & Welsh, B. C. (2019). Looking through broken windows: The impact of neighborhood disorder on aggression and fear of crime is an artifact of research design. *Annual Review of Criminology, 2*(1), 53-71.

O’Brien, D. T., Sampson, R. J., & Winship, C. (2015). Ecometrics in the age of big data: Measuring and assessing “Broken windows” using large-scale administrative records. *Sociological Methodology, 45*(1), 101-147.

Odgers, C. L., Caspi, A., Bates, C. J., Sampson, R. J., & Moffitt, T. E. (2012). Systematic social observation of children's neighborhoods using Google Street View: A reliable and cost-effective method. *Journal of Child Psychology and Psychiatry*, *53*(10), 1009-1017.

Pace, S. A. (2010). "Assessing the impact of police order maintenance units on crime: An application of the Broken Windows Hypothesis". Unpublished M.A. Thesis, University of Nevada.

Pate, A. M., Skogan, W., Wycoff, M.A. & Sherman, L.W. (1985). *Reducing the 'Signs of Crime': The Newark Experience*. Washington, DC: Police Foundation.

Perkins, D. D., & Taylor, R. B. (1996). Ecological assessments of community disorder: Their relationship to fear of crime and theoretical implications. *American Journal of Community Psychology*, *24*(1), 63-107.

Perkins, D. D., Meeks, J. W., & Taylor, R. B. (1992). The physical environment of street blocks and resident perceptions of crime and disorder: Implications for theory and measurement. *Journal of Environmental Psychology*, *12*(1), 21-34.

Perkins, D. D., Wandersman, A., Rich, R., & Taylor, R. (1993). The physical environment of street crime: Defensible space, territoriality and incivilities. *Journal of Environmental Psychology, 13,* 29-49.

Perperoglou, A., Sauerbrei, W., Abrahamowicz, M., & Schmid, M. (2019). A review of spline function procedures in R. *BMC Medical Research Methodology, 19*(1), 46-16.

Piza, E. L., & O'Hara, B. A. (2014). Saturation foot-patrol in a high-violence area: A quasi-experimental evaluation. *Justice Quarterly, 31*(4), 693-718.

Prener, C. G., Braswell, T. H., & Monti, D. J. (2020). St. louis's "urban prairie": Vacant land and the potential for revitalization. *Journal of Urban Affairs: Promoting Social Justice and Equity in Shrinking Cities, 42*(3), 371-389.

Qiu, F., Zhang, C., & Zhou, Y. (2012). The development of an areal interpolation ArcGIS extension and a comparative study. *GIScience & Remote Sensing, 49*(5), 644-663.

Raleigh, E., & Galster, G. (2015). Neighborhood disinvestment, abandonment, and crime dynamics. *Journal of Urban Affairs, 37*(4), 367-396.

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

Ratcliffe, J. H., Groff, E. R., Sorg, E. T., & Haberman, C. P. (2015). Citizens' reactions to hot spots policing: Impacts on perceptions of crime, disorder, safety and police. *Journal of Experimental Criminology, 11*(3), 393-417.

Ratcliffe, J. H., Taniguchi, T., Groff, E. R., & Wood, J. D. (2011). The Philadelphia foot patrol experiment: A randomized controlled trial of police patrol effectiveness in violent crime hotspots: Philadelphia foot patrol experiment. *Criminology, 49*(3), 795-831.

Raudenbush, S. W. (2005). How do we study "What happens next"? *The Annals of the American Academy of Political and Social Science, 602*(1), 131-144.

Reisig, M. D. (2010). Community and Problem-Oriented policing. *Crime and Justice, 39*(1), 1-53.

Reisig, M. D., & Cancino, J. M. (2004). Incivilities in nonmetropolitan communities: The effects of structural constraints, social conditions, and crime. *Journal of Criminal Justice, 32*(1), 15-29.

Ren, L., Zhao, J., & He, N. (2019). Broken windows theory and citizen engagement in crime prevention. *Justice Quarterly, 36*(1), 1-30.

Rengifo, A. F., & Fowler, K. (2016). Stop, question, and complain: Citizen grievances against the NYPD and the opacity of police stops across New York city precincts, 2007-2013. *Journal of Urban Health: Bulletin of the New York Academy of Medicine, 93 Suppl 1*(S1), 32-41.

Rengifo, A. F., & Fratello, J. (2015). Perceptions of the police by immigrant youth: Looking at stop-and-frisk and beyond using a New York city sample. *Youth Violence and Juvenile Justice, 13*(4), 409-427.

Ridgeway, G. (2007). Analysis of racial disparities in the New York police department's stop, question, and frisk practices. Santa Montica, CA: Rand Corporation.

Robinson, J. B., Lawton, B. A., Taylor, R. B., & Perkins, D. D. (2003). Multilevel longitudinal Impacts of incivilities: Fear of crime, expected safety, and block satisfaction. *Journal of Quantitative Criminology, 19*(3), 237-274.

Rohe, W. M., & Stegman, M. A. (1994). The impact of home ownership on the social and political involvement of low-income people. *Urban Affairs Quarterly, 30*(1), 152-172.

Rosenbaum, P. R. (2002). *Observational Studies,* 2nd Edition. New York: Springer Verlag.

Rosenbaum, D. P. (2006). The limits of hot spots policing. In D. Weisburd & A. A. Braga (Eds.), *Police innovation: Contrasting perspectives*. Cambridge, MA: Cambridge University Press.

Rosenfeld, R., Fornango, R., & Rengifo, A. F. (2007). The impact of order-maintenance policing on New York City homicide and robbery rates: 1988-2001. *Criminology, 45*(2), 355-384.

Ross, C. E., & Jang, S. J. (2000). Neighborhood disorder, fear, and mistrust: The buffering role of social ties with neighbors. *American Journal of Community Psychology, 28*(4), 401-420.

Ross, C. E., & Mirowsky, J. (2001). Neighborhood disadvantage, disorder, and health. *Journal of Health and Social Behavior*, *42*(3), 258-276.

Rudin, C., & Carlson, D. (2018). The secrets of machine learning: Ten things you wish you had known earlier to be more effective at data analysis. *Informs*, 1-29.

Rundle, A. G., Bader, M. D., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using Google Street View to audit neighborhood environments. *American Journal of Preventive Medicine*, *40*(1), 94-100.

Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics, 11*(4), 735-757.

Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Cambridge; New York: Cambridge University Press.

Sabol, W. J., Coulton, C. J., & Korbin, J. E. (2004). Building community capacity for violence prevention. *Journal of Interpersonal Violence, 19*(3), 322-340.

Sampson, R. J. (1988). Local friendship ties and community attachment in mass society: A multilevel systemic model. *American Sociological Review, 53*(5), 766-779.

Sampson, R. J. (2004). Networks and neighbourhoods: The implications of connectivity for thinking about crime in the modern city. In H. McCarthy, P. Miller and P. Skidmore (Eds.), *Network Logic: Who Governs in an Interconnected World?* (pp 157-166). London: Demos.

Sampson, R. J. (2012). *Great American city: Chicago and the enduring neighborhood effect.* Chicago: University of Chicago Press.

Sampson, R. J. (2013). the place of context: A theory and strategy for criminology's hard problems. *Criminology, 51*(1), 1-31.

Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology*, *94*(4), 774-802.

Sampson, R. J., & Laub, J. H. (2005). A life-course view of the development of crime. *The Annals of the American Academy of Political and Social Science, 602*(1), 12-45.

Sampson, R. J., & Raudenbush, S. W. (1999). Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology, 105*(3), 603-651.

Sampson, R. J., & Raudenbush, S. W. (2004). Seeing disorder: Neighborhood stigma and the social construction of "broken windows". *Social Psychology Quarterly, 67*(4), 319-342.

Santos, R. G., & Santos, R. B. (2015). Practice-based research: Ex post facto evaluation of evidence-based police practices implemented in residential burglary micro-time hot spots. *Evaluation Review, 39*(5), 451-479.

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science, 277*(5328), 918-924.

Scaglion, R., & Condon, R. G. (1980). Determinants of attitudes toward city police. *Criminology, 17*(4), 485-494.

Schuetz, J., Spader, J., & Cortes, A. (2016). Have distressed neighborhoods recovered? evidence from the neighborhood stabilization program. *Journal of Housing Economics, 34*, 30-48.

Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas: A study of rates of delinquents in relation to differential characteristics of local communities in American cities*. Chicago: The University of Chicago Press.

Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology, 27*(1), 27-56.

Sherman, L. W. (2007). The power few: Experimental criminology and the reduction of harm: The 2006 Joan McCord prize lecture. *Journal of Experimental Criminology, 3*(4), 299-321

Silver, E., & Miller, L. L. (2004). Sources of informal social control in Chicago neighborhoods. *Criminology, 42*(3), 551-584.

Simpson, R. (2017). The police officer perception project (POPP): An experimental evaluation of factors that impact perceptions of the police. *Journal of Experimental Criminology, 13*(3), 393–415.

Skogan, W. G. (1990). *Disorder and decline: Crime and the spiral of decay in American neighborhood*s. New York; Toronto: Free Press.

Skogan W. G. (2006). The promise of community policing. In D. Weisburd & A. A. Braga (Eds.), *Policing innovation: Contrasting perspectives* (pp. 27-43). Cambridge, UK: Cambridge University Press.

Skogan, W. G. (2012) Disorder and crime. In B.C. Welsh & D. P. Farrington (eds), *The Oxford Handbook of Crime Prevention* (pp. 173-188). New York: Oxford University Press.

Skogan, W. G. (2015). Disorder and decline: The state of research. *Journal of Research in Crime and Delinquency, 52*(4), 464-485.

Skogan, W. G., & Frydl, K. (2004). Fairness and effectiveness in policing: The evidence. In Committee to review research on police policy and practices. Committee on law and justice, division of behavioral and social sciences and education. Washington, DC: The National Academies Press.

Small, M. L. (2009). *Unanticipated gains: Origins of network inequality in everyday life*. New York; Oxford: Oxford University Press.

Smith, M. J., & Mazerolle, L. (2013). Using civil actions against property to control crime problems. *Community Oriented Policing Services, 11*.

Spader, J., Schuetz, J., & Cortes, A. (2016). Fewer vacants, fewer crimes? impacts of neighborhood revitalization policies on crime. *Regional Science and Urban Economics, 60*, 73-84.

Spelman, W. (2004). Optimal targeting of incivility-reduction strategies. *Journal of Quantitative Criminology, 20*(1), 63-88.

Steenbeek, W., & Hipp, J. R. (2011). A longitudinal test of social disorganization theory: Feedback effects among cohesion, social control, and disorder. *Criminology, 49*(3), 833-871.

Steenbeek, W., & Kreis, C. (2015). Where broken windows should be fixed: Toward identification of areas at the tipping point. *Journal of Research in Crime and Delinquency, 52*(4), 511-533.

Sullivan, C. J., & Loughran, T. (2014). Investigating the functional form of the self-control-delinquency relationship in a sample of serious young offenders. *Journal of Quantitative Criminology, 30*(4), 709-730.

Swatt, M. L., Varano, S. P., Uchida, C. D., & Solomon, S. E. (2013). Fear of crime, incivilities, and collective efficacy in four Miami neighborhoods. *Journal of Criminal Justice, 41*, 1-11.

Taylor, R. B. (1999). Responses to disorder: Relative impacts of neighborhood structure, crime, and physical deterioration on residents and business personnel. National Institute of Justice, Washington, DC.

Taylor, R. B. (2001). *Breaking away from broken windows: Baltimore neighborhoods and the nationwide fight against crime, grime, fear, and decline*. Boulder, CO: Westview Press.

Telep, C. W., & Weisburd, D. (2012). What is known about the effectiveness of police practices in reducing crime and disorder? *Police Quarterly, 15*(4), 331-357.

Tyler, T. R., & Fagan, J. (2008). Legitimacy and cooperation. Why do people help the police fight crime in their communities? *Ohio State Journal of Criminal Law, 6*, 231–275.

Taylor, R. B., & Shumaker, S. A. (1990). Local crime as a natural hazard: Implications for understanding the relationship between disorder and fear of crime. *American Journal of Community Psychology, 18*(5), 619.

Taylor, R. B., Gottfredson, S. D., & Brower, S. (1984). Block crime and fear: Defensible space, local social ties, and territorial functioning. *Journal of Research in Crime and Delinquency*, *21*(4), 303-331.

Taylor, R. B., Koper, C. S., & Woods, D. J. (2011). A randomized controlled trial of different policing strategies at hot spots of violent crime. *Journal of Experimental Criminology*, *7*(2), 149-181.

Teig, E., Amulya, J., Bardwell, L., Buchenau, M., Marshall, J. A., & Litt, J. S. (2009). Collective efficacy in Denver, Colorado: Strengthening neighborhoods and health through community gardens. *Health and Place, 15*(4), 1115-1122.

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.

Todak, N., & James, L. (2018). A systematic social observation study of police de-escalation tactics. *Police Quarterly, 21*(4), 509-543.

Trojanowicz, R. C. (1986). Evaluating a neighborhood foot patrol program: The Flint, Michigan, Project. In Dennis P. Rosenbaum (ed.), *Community Crime Prevention: Does it Work?* Newbury Park, CA: Sage.

Trojanowicz, R. C., & Baldwin, R. (1982). An evaluation of the neighborhood foot patrol program in Flint, Michigan, (pp. 1-184). East Lansing: Michigan State University.

Trojanowicz, R. C., & Banas, D. W. (1985). *Perceptions of safety: A comparison of foot patrol versus motor patrol officers*. East Lansing, Michigan: National Neighborhood Foot Patrol Center, School of Criminal Justice, Michigan State University.

Warner, B. D., & Rountree, P. W. (1997). Local social ties in a community and crime model: Questioning the systemic nature of informal social control. *Social Problems, 44*(4), 520-536.

Weisburd, D., & Green, L. (1995). Policing drug hot spots: The Jersey City drug market analysis experiment. *Justice Quarterly, 12*(4), 711-735.

Weisburd, D., Bushway, S., Lum, C., & Yang, S. (2004). Trajectories of crime at places: a longitudinal study of street segments in the city of Seattle. *Criminology, 42*(2), 283-322.

Weisburd, D., Davis, M., & Gill, C. (2015). Increasing collective efficacy and social capital at crime hot spots: New crime control tools for police. *Policing: A Journal of Policy and Practice, 9*(3), 265-274.

Weisburd, D., Groff, E. R., & Yang, S. (2012). *The criminology of place: Street segments and our understanding of the crime problem*. US: Oxford University Press.

Weisburd, D., Hinkle, J. C., Braga, A. A., & Wooditch, A. (2015). Understanding the mechanisms underlying broken windows policing: The need for evaluation evidence. *Journal of Research in Crime and Delinquency, 52*(4), 589-608.

Weisburd, D., Hinkle, J. C., Famega, C., & Ready, J. (2011). The possible "backfire" effects of hot spots policing: An experimental assessment of impacts on legitimacy, fear and collective efficacy. *Journal of Experimental Criminology, 7*(4), 297-320.

Weisburd, D., Hinkle, J.C., Famega, C., and Ready, J. (2012). *Legitimacy, Fear and Collective Efficacy in Crime Hot Spots.* Washington, DC: National Institute of Justice.

Wells, W., Schafer, J. A., Varano, S. P., & Bynum, T. S. (2006). Neighborhood residents' production of order: the effects of collective efficacy on responses to neighborhood problems. *Crime and Delinquency, 52*(4), 523–550.

Wheeler, A. P. (2018). The effect of 311 calls for service on crime in D.C. at microplaces. *Crime & Delinquency, 64*(14), 1882-1903.

Wheeler, A. P. (2019). Quantifying the local and spatial effects of alcohol outlets on crime. *Crime & Delinquency, 65*(6), 845-871.

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.

White, A., & Trump, K. (2018). The promises and pitfalls of 311 data. *Urban Affairs Review, 54*(4), 794-823.

Wickes, R., & Hipp, J. R. (2018). The spatial and temporal dynamics of neighborhood informal social control and crime. *Social Forces, 97*(1), 277-307.

Wickes, R., Broidy, L., & Hipp, J. R. (2018). Responding to neighborhood problems: Is the division of community labor gendered? *Crime & Delinquency, 64*(9), 1215-1241.

Wickes, R., Hipp, J. R., Zahnow, R., & Mazerolle, L. (2013). “Seeing” minorities and perceptions of disorder: Explicating the mediating and moderating mechanisms of social cohesion. *Criminology*, *51*(3), 519-560.

Wiles, P. & Costello, A. (2000). *The road to nowhere: The evidence for travelling criminals*, Home Office Research Study No. 207. London: Home Office.

Williams, C. (2012, February 13). Detroit trying ‘broken windows’. *Detroit Free Press.*

Wilson, W. J. (1996). *When work disappears: the world of the new urban poor*. Random House, New York.

Wilson, W. J., & Kelling, G. (1982). The police and neighborhood safety: Broken windows. *The Atlantic*, *127*, 29-38.

Wilson, W. J., & Kelling, G. (1989). Making neighborhoods safe. *The Atlantic*, *263,* 46-52.

Wolfgang, M., Figlio, R., & Sellin, T. (1972). *Delinquency in a birth cohort*. Chicago: University of Chicago Press.

Worrall, J. L., & Wheeler, A. P. (2019). Evaluating community prosecution code enforcement in Dallas, Texas. *Justice Quarterly, 36*(5), 870-899.

Wright, R., & Decker, S. (1997). *Armed robbers in action: stickups and street culture.* Boston: Northeastern University Press.

Wright, R., Brookman, F., & Bennett, T. (2006). The foreground dynamics of street robbery in Britain. *The British Journal of Criminology, 46*(1), 1-15.

Xiao, C., & McCright, A. M. (2014). A test of the biographical availability argument for gender differences in environmental behaviors. *Environment and Behavior, 46*(2), 241-263.

Xu, Y., Fiedler, M. L., & Flaming, K. H. (2005). Discovering the impact of community policing: The broken windows thesis, collective efficacy, and citizens’ judgment. *Journal of Research in Crime and Delinquency, 42*(2), 147-186.

Yang, S. (2010). Assessing the Spatial—Temporal relationship between disorder and violence. *Journal of Quantitative Criminology, 26*(1), 139-163.

Yang, S., & Pao, C. (2015). Do we “See” the same thing? an experimental look into the black box of disorder perception. *Journal of Research in Crime and Delinquency, 52*(4), 534-566.