



How and Why is Crime More Concentrated in Some Neighborhoods than Others?: A New Dimension to Community Crime

Daniel T. O'Brien^{1,2,3,4} · Alexandra Ciomek^{3,4,5} · Riley Tucker^{2,3,4}

Accepted: 14 January 2021 / Published online: 8 February 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

Abstract

Objectives Much recent work has focused on how crime concentrates on particular streets within communities. This is the first study to examine how such concentrations vary across the neighborhoods of a city. The analysis evaluates the extent to which neighborhoods have characteristic levels of crime concentration and then tests two hypotheses for explaining these variations: the compositional hypothesis, which posits that neighborhoods whose streets vary in land usage or demographics have corresponding disparities in levels of crime; and the social control hypothesis, which posits that neighborhoods with higher levels of collective efficacy limit crime to fewer streets.

Methods We used 911 dispatches from Boston, MA, to map violent crimes across the streets of the city. For each census tract we calculated the concentration of crime across the streets therein using the generalized Gini coefficient and cross-time stability in the locations of hotspots.

Results Neighborhoods did have characteristic levels of concentration that were best explained by the compositional hypothesis in the form of demographic and land use diversity. In addition, ethnic heterogeneity predicted higher concentrations of crime over and above what would be expected given the characteristics of the individual streets, suggesting it exacerbated disparities in crime.

Conclusions The extent to which crime concentrates represents an underexamined aspect of how crime manifests in each community. It is driven in part by the diversity of places in the neighborhood, but also can be influenced by neighborhood-level processes. Future work should continue to probe the sources and consequences of these variations.

Keywords Hotspots · Urban criminology · Communities and crime · Quantitative methods

✉ Daniel T. O'Brien
d.obrien@neu.edu

¹ School of Public Policy and Urban Affairs, Northeastern University, 310 Renaissance Park, 1135 Tremont St, Boston, MA 02120, USA

² School of Criminology and Criminal Justice, Northeastern University, Boston, USA

³ Boston Area Research Initiative, Northeastern University, Cambridge, USA

⁴ Boston Area Research Initiative, Harvard University, Cambridge, USA

⁵ Department of Sociology, Harvard University, Cambridge, USA

Introduction

A rapidly growing and influential literature has highlighted how certain localized places—street segments and addresses—are responsible for disproportionate amounts of crime and disorder in a city. The foundational theory for this work is *the law of concentration of crime*, which stipulates that for a given microgeographic unit there is a narrow bandwidth of percentages for a defined cumulative proportion of crime events in a city (Weisburd 2015). It is unclear, however, what the implications of this law are for individual neighborhoods within a city. Do they all feature the same intensity of concentration of crime, or do they vary in these regards? A few recent studies, however, have found preliminary evidence that communities vary substantially in how much crime is concentrated at particular streets and addresses (Hipp and Kim 2017; O'Brien 2019). Such results raise questions about how and why crime might be highly geographically concentrated in some communities and more evenly distributed in others.

Considering how neighborhoods vary in their concentration of crime opens a new dimension in how we might describe crime in communities. Traditionally, criminologists have focused on the relative amount of crime in a community—that is, which neighborhoods in a city have high, moderate, and low levels of crime—and how these differences arise. A focus on concentrations of crime reflects a distinct component of how crime manifests in a community, analogous to how income inequality complements median income when understanding a community's socioeconomic composition. To illustrate, picture two communities with the same level of crime, but in one community events are concentrated at one or two hotspots and in the other they are distributed more evenly. Just as income inequality holds its own consequences for residents, the differences in concentration of crime might hold implications for the social dynamics that generate crime as well as the way local residents experience crime. Such differences are an underexplored aspect of cross-community variation in the nature of crime.

The current study is an initial attempt to understand how and why concentrations of crime vary across neighborhoods. We analyze violent crimes reported through 911 in Boston, MA, in order to examine two main questions. First, it is necessary to establish whether variation in the concentration of crime is in fact meaningful or whether it is merely a statistical artifact. Put more tangibly, is crime reliably more or less concentrated in some neighborhoods relative to others? We test this by comparing multiple measures of concentration of crime and assessing their stability over time. Higher correlations and stability would provide greater evidence that concentration of crime is a reliable characteristic of a neighborhood worth studying in its own right. Second, if each neighborhood does have a characteristic concentration of crime, what physical, demographic, or social factors best explain these variations? That is, what are the features and processes that either lead to the isolation of crime on a small number of street segments or, alternatively, cause it to be distributed throughout a neighborhood? Before proceeding to presentation of the data and results, the remainder of this section reviews existing evidence for variation in the concentration of crime; lays out two hypotheses for why crime might concentrate at different levels across communities; and describes how the current study will examine the questions at hand.

Concentrations of Crime in Communities

Interest in crime at places (or “criminology of place”) has seen rapidly growing interest since two seminal studies in the late 1980’s that found that ~3% of addresses account for over 50% of crimes in each of two cities (Sherman, Gartin, and Buerger 1989; Pierce, Spaar, and Briggs 1988). Since then, numerous studies have demonstrated similar patterns of concentration for street segments. Work in a wide range of cities has consistently found that 4–6% of streets generate at least 50% of crime events (Braga, Papachristos, and Hureau 2010; Andresen and Malleson 2011; Weisburd 2015). The literature has been robust in a number of ways. While the focus has largely been on streets, additional studies have replicated and extended the early results for addresses (e.g., Farrell and Pease 2001; Johnson et al. 2007; Trickett et al. 1992; O’Brien and Winship 2017), including the concentration of crime at a small number of major institutions and facilities (Eck, Clarke, and Guerette 2007). Other work has illustrated that hotspots persist across years, with the same set of streets accounting for the majority of a given city’s crime over time (Groff, Weisburd, and Yang 2010; Braga, Papachristos, and Hureau 2010; Braga, Hureau, and Papachristos 2011; Curman, Andresen, and Brantingham 2015). Further, a recent systematic review of 44 studies found that the proportion of places (streets or addresses) that were responsible for a given proportion of crimes in a city was highly consistent (Lee et al. 2017).

The extended evidence that a small number of places accounts for the vast majority of crime events led Weisburd (2015) to propose *the law of concentration of crime*: that for a given microgeographic unit there is a narrow bandwidth of percentages for a defined cumulative proportion of crime events in a city. The law leaves open the possibility, however, that the extent to which crime concentrates might vary across communities within a city. Some preliminary evidence supports this notion. First, early work on the law of concentration of crime by Weisburd, Groff, and Yang (2012) included many maps to illustrate the concentration of crime on hotspot streets in Seattle, WA. To the naked eye, it is apparent that the hotspots were not evenly distributed across the city, but were more strongly differentiated from the surrounding environment in some neighborhoods than others. More recently, O’Brien (2019) used the Gini coefficient, a standardized measure of inequality, to uncover variation in concentrations across the census tracts of Boston, MA. Last, in a paper critiquing the law of concentration of crime Hipp and Kim (2017) found that 42 cities in Southern California had substantially different levels of stability in crime hotspots, which was used as an indicator of a broader tendency for crime to concentrate. This suggests that similar variations might exist more locally as well.

The apparent variation in concentrations of crime across communities captures an underexplored dimension of how crime manifests differently across communities. Nearly all urban criminology has emphasized the amount of crime in a neighborhood (or city) relative to others. This was the fundamental inspiration of Shaw and McKay’s (1942/1969) seminal studies in Chicago, for example, and continues to be a major focus of the field. Whereas studies on the amount or frequency of crime focus on variation *between* neighborhoods, concentrations of crime quantify variation *within* a given neighborhood. This is an important point because sometimes the statement “a neighborhood has a heavy concentration of crime” is colloquially intended to communicate a high-crime neighborhood in the sense that crime is concentrated there relative to other parts of the city. Instead, the two phrases mean different things, and a neighborhood’s relative level on each need not be the same. Take, for example, the two Boston neighborhoods (census tracts) presented in Fig. 1.

Fig. 1 An illustration of the variation in the concentration of crime across neighborhoods. **a** The distribution of composite Gini coefficients for all census tracts, with two neighborhoods with the same amount of crime but different levels of concentration highlighted. Below is a direct comparison of the distribution of violent events across the streets of the two neighborhoods **b, d** and corresponding maps **c, e**. Note that scales for numbers of crimes (x -axis and coloration) were kept consistent between the two tracts to enable easier comparison

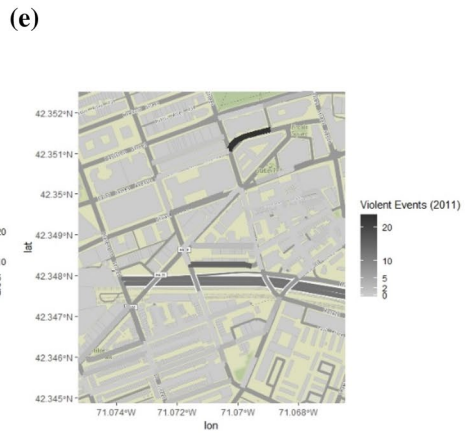
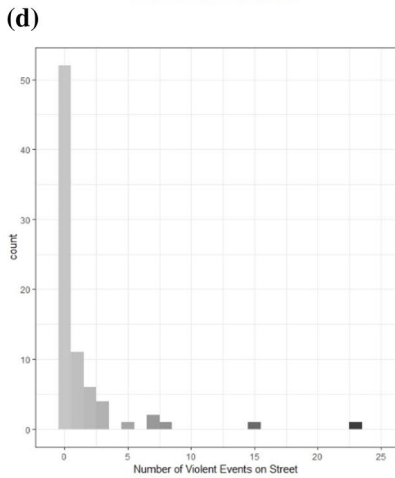
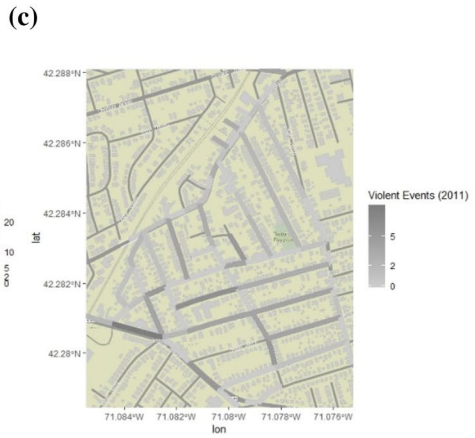
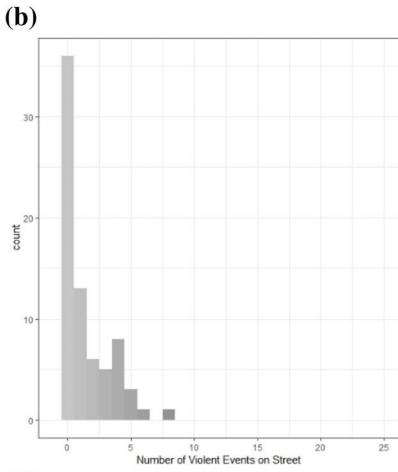
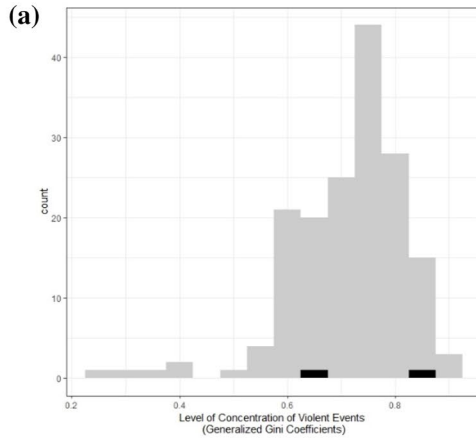
They have approximately the same amount of crime, but very different concentrations of crime. One (Fig. 1b and c) has a relatively low level of concentration, with crime distributed across multiple streets and no truly outstanding hotspots. The other neighborhood (Fig. 1d and e) has a high level of concentration, with most crimes occurring on two major hotspots, one in the center of the neighborhood and the other to the north. Despite the same amount of crime, these two neighborhoods have very different levels of concentration.

The implications of criminology of place have often been summarized in the statement that most streets in high-crime neighborhoods are safe and even low-crime neighborhoods can have hotspots (e.g., Groff, Weisburd, and Yang 2010; Steenbeek and Weisburd 2016). This phrasing, however, could be misinterpreted as implying that concentration is a given rather than a variable. The evidence and illustration presented here, however, suggest that there may be greater nuance to the matter. It could be, for instance, that most streets in some high-crime neighborhoods are safe, but that crime might occur on the majority of streets in some others. If so, there are likely differences in the dynamics pertaining to crime in those neighborhoods, despite having similar amounts of crime. In turn, different levels of concentration might have implications for the experiences of the people who live, work, and recreate there. For example, in a neighborhood with a highly uneven distribution of crime, the location of an individual's home and daily routines might have marked impacts on fear of crime (Glas, Engbersen, and Snel 2019). These sorts of questions, however, remain largely unexamined.

Why Might Concentrations of Crime Vary Across Communities?

If the level of concentration of crime does vary across communities, it begs the further question of why such variations occur and whether particular characteristics of communities lead to greater or lesser concentrations of crime. Put in simple terms, what are the features and processes that either lead to the isolation of crime on a small number of street segments or, alternatively, cause it to be distributed throughout a neighborhood? Here we focus particularly on neighborhoods, presenting two hypotheses for how this variation arises. The compositional hypothesis posits that streets in an institutionally or demographically diverse neighborhood will have greater variation in their vulnerability to crime. In contrast, the social control hypothesis argues that community members guide crime to concentrate on some streets rather than others. It is important to note that these two hypotheses are not mutually exclusive and that each might be true.

The compositional hypothesis stipulates that inequality in the incidence of crime across the streets of a community is attributable to the characteristics of those individual streets. To illustrate, some land uses generate more crime by virtue of the routine activities that occur there (i.e., the volume and types of visitors and the sorts of activities in which they participate while there; Cohen and Felson 1979; Brantingham and Brantingham 1993). It might be that a mixed-use neighborhood will experience a more uneven distribution of crime because streets with high levels of commercial and pedestrian activity see lots of crime whereas quieter residential streets see less. Likewise, in a demographically diverse



neighborhood, the population on some streets might be substantially more vulnerable to crime (or to the reporting of crime) than other streets (e.g., streets with subsidized housing within a middle- or upper-class neighborhood).

In the case that institutional or demographic diversity is associated with a greater concentration of crime within a neighborhood, it is necessary to adjudicate between two competing sub-hypotheses. The simpler interpretation is that the uneven distribution of crime arises because the land uses and sub-populations that compose the neighborhood have different vulnerabilities to crime or the tendency for it to be reported. Arithmetically, a higher concentration of crime is the result of comparing crime across these distinct places and people. Alternatively, it is possible that an elevated concentration of crime is an emergent property of the diverse context. That is to say, the heterogeneity itself drives one or more behavioral or social processes that instigate a more uneven distribution of crime across streets. For example, Legewie and Schaeffer (2016) found that ethnically heterogeneous neighborhoods in New York City produced more complaints about minor disturbances, apparently as a way for neighbors with different norms to indirectly manage each other's behaviors. The complaints were especially elevated on the street blocks where different ethnic groups interfaced. This suggests that differences between streets in the number of complaints were more than a function of multiple ethnicities with different propensities for crime living in close proximity. Instead, it would seem that the interactions between these ethnicities led to a greater number of complaints on specific streets.

The social control hypothesis has been partially articulated by Weisburd et al. (2012). Drawing off of Durkheim's (1895 / 1964) argument that crime serves a purpose in society and therefore tends toward a "natural" level, they suggested that there might also be a natural level of concentration (though dispensing with the now outmoded claim that crime serves a needed societal function). They argue that, if crime is a fact of society, then communities likely prefer that it predictably occur in certain select places and not in others. Achieving this goal would then potentially be the province of informal social control or collective efficacy (i.e., the ability to achieve shared goals, a component of which is informal social control). Much theory and research has focused on how informal social control can be directed to discourage and prevent crime in general (Sampson, Raudenbush, and Earls 1997; Bursik and Grasmick 1993), but those studies have repeatedly demonstrated that informal social control is insufficient to completely eliminate crime. Thus, an alternative role of informal social control could be to isolate crime geographically to a small set of locations, especially those whose land usage leads to limits the amount of monitoring that can occur there, and otherwise keeping it to a minimum wherever possible. This would create high concentrations of crime within the neighborhood. Where the capacity for social control is lower, however, crime and disorder are more able to permeate throughout the neighborhood.

To summarize, we present two main hypotheses for how variance in the concentration of crime might emerge. First, institutional or demographic diversity across streets will generate different levels of crime across the neighborhood *or* could instigate interactions that lead crimes (or their reporting) to be exaggerated in particular places. Second, neighborhoods with stronger social control isolate crime to certain streets. As this study compares these two hypotheses, we must also consider a third possibility. Much of the variation in the concentration of crime across the streets of a community could be a function of the idiosyncrasies of the specific set of streets, and thus not easily explained by standard correlates. This seems possible based on previous work. For example, Eck, Clarke, and Guerette (2007) found that even for institutions that are particularly vulnerable to crime, like liquor stores, only a small proportion in fact experience large number of crimes. Further, O'Brien

(2019) found that adjusting Gini coefficients according to differences in land usage across neighborhoods did little to alter the concentrations of crime in each. Meanwhile, experts in repeat victimization have argued that each crime event can lead to greater concentration of crime in a place by either “flagging” that place as suitable for crime or “boosting” its attractiveness for crime (Johnson 2008). Thus, particularities in either current routine activities or historical circumstance might be as responsible for concentrations of crime as any of a neighborhood’s social or compositional characteristics.

The Current Study

The current study examines the variation in the concentration of crime across the neighborhoods of Boston, MA. Studies of concentrations of crime often focus on particular types of crimes (e.g., robbery, Braga, Hureau, and Papachristos 2011; shootings, Loeffler and Flaxman 2018). In this first analysis of variations in concentrations of crime across neighborhoods, we choose to focus on violent crime, as it is arguably the form of crime whose presence in a neighborhood is most salient for the well-being of residents. The study uses a three-year archive of 911 dispatches. Previous work with these data has developed groupings of case types that reflect particular dimensions of crime and social disorder (O’Brien and Sampson 2015a, b), and here we analyze a single measure that combines the two dimensions that reflect violent crime: public violence and prevalence of guns.

We link each record to the street segment where it occurred and nest streets within neighborhoods, approximated using census tracts. We then tabulate the number of events on each street and examine the level of concentration across the streets of each census tract. We recognize that census geographies are at best an approximation of actual neighborhoods, and that their boundaries might be drawn in ways that can introduce error to estimates of concentrations. That said, we utilize them here for a few reasons. First, they are traditionally used to estimate neighborhoods across disciplines, permitting comparability. Second, many data sets use census tracts as a unit of measurement, allowing us to incorporate various correlates into the analysis. Our last and most substantive reason for using tracts relates to the question of geographic scale. Some have sought to mitigate measurement error by analyzing multi-tract “neighborhood clusters” that are demographically homogeneous (e.g., Schnell et al. 2017). These are typically much larger than the spaces that individuals identify with as neighborhoods, however. In fact, cognitive maps and movement patterns both regularly show that residents define their neighborhoods as being approximately the size of a census tract (about a radius of a half-mile; Guest and Lee 1984; Coulton et al. 2001; Colabianchi et al. 2007). Applying this to concentrations of crime, if a neighborhood cluster is indeed a conglomeration of communities, it would be questionable whether crime events occurring in one part of a neighborhood cluster could have feasibly occurred in another part that is far away. This geographic reachability is necessary for processes or conditions leading to the concentration of crime to be meaningful. Thus, we believe that census tracts are the more appropriate geographic scale for the question at hand.

The study pursues two main questions. First, we seek to determine whether variation in the concentration of violent crime is meaningful or a statistical artifact. While Hipp and Kim (2017) and O’Brien (2019) each uncovered this variation, they did not do the necessary tests to establish whether each community has a characteristic concentration of crime that is stable relative to other communities across time and across measures. Second, we test the compositional and social control hypotheses for why variations in concentrations

of crime might emerge across neighborhoods. Provided there is evidence for the compositional hypothesis, we will further explore whether such relationships are explained entirely by the characteristics of the individual streets or if the heterogeneity of the neighborhood has an emergent effect on concentrations of crime.

In order to examine the two main questions, we must first select a technique for measuring the level of concentration of crime in each census tract. The most common approach is the use of cumulative frequency distributions to identify what proportion of streets account for a given proportion of crimes (e.g., 50%). This approach, however, has two weaknesses: it is vulnerable to generating artifactually small values owing to the small number of actual crimes and the consequently high frequency of places with zero crimes; it relies on an arbitrary threshold for the cumulative frequency—should it be 50% of crimes? 75% of crimes? 90% of crimes? Hipp and Kim (2017) proposed techniques that attend to the first of these weaknesses to leverage multiple years of data. Specifically, they quantified whether the same streets were responsible for the preponderance of crime across years, which they referred to as “temporally adjusted crime concentration.” This approach still requires the stipulation of a proportion of streets (e.g., 5%) or crimes (e.g., 50%) that is believed to be most informative, but its use of longitudinal information makes it more robust to stochastic spikes and drops in crime.

Other authors have recently proposed more standardized methodologies that measure the concentration of crime across communities without requiring a threshold. Bernasco and Steenbeek (2017) argued that concentration of crime is essentially a case of inequality, lending it well to the Gini coefficient used by economists. The Gini coefficient offers a standardized measure of inequality based on the Lorenz curve, which plots the proportion of a given quantity on the y-axis (often wealth, though in this case crime events) held by the x% of the population with the lowest amount of that quantity (i.e., the fewest crime events). The Gini coefficient then quantifies inequality as the total distance between the points on the Lorenz curve and the line of perfect equality ($y=x$).¹ Curiel, Delmar, and Bishop (2018) took a similar though more complex approach with the rare event concentration coefficient (RECC), which decomposes the frequency distribution into multiple subpopulations with different probabilities of experiencing one or more crimes. They then used the relative size and risk of crime for these subpopulations to calculate inequality in the distribution of crime. Mohler et al. (2019) also extended the Gini coefficient strategy, applying it to an inferred Poisson-Gamma distribution, which better handles the distribution of rare events.

Given the strengths and weaknesses of the techniques for quantifying the concentration of crime in a community, we have chosen to use two in the proceeding analysis. First, the generalized Gini coefficient is a well-established, highly interpretable metric that has already been used to demonstrate differences in concentration across communities. It is also straightforward to implement, whereas the inferential procedures of its more complex cousins the RECC (Curiel et al. 2018) and Mohler et al.’s (2019) inferred Poisson-Gamma distribution require too large of a sample within each group to be appropriate for the geographic scale of neighborhoods.

¹ Formally calculated as the area between the line of equality and the Lorenz curve divided by the total area under the line of equality. This is also equal to twice the area between the line of equality and the Lorenz curve as the line of equality delineates a right triangle on a set of axes with scale 0–1, thus having a total area of .5.

Second, we use a variant of Hipp and Kim's (2017) measure of temporally adjusted crime concentration. They sorted the streets of a city in descending order of number of crimes in one year and then calculated what proportion of crimes occurred on the first 5% of streets in the following year. Our goal here is slightly different in that we want to use the same logic to determine whether crime concentrates on the same streets from one year to the next. Thus, instead of defining the proportion of streets in advance, which can limit us to very few streets in a census tract, we identified the streets that accounted for 50% of crimes in one year and calculated the proportion of overlap with streets responsible for the "first" 50% of crimes in the following year. This adds a temporal dimension that is absent from cross-sectional Gini coefficients, permitting a fuller assessment of how crime is or is not limited to certain parts of a community. Because the intent is less to adjust for stochastic influences on crime counts and more about whether the crime concentrates in the same places year to year, we refer to this as *stability in crime concentrations* from hereon.

The two measures are the vehicle for testing each of our research questions. First, we assess the extent to which concentration of crime is characteristic for a neighborhood by testing the correlation between the generalized Gini coefficient and the stability in crime concentrations. Additionally, the three-year span of the data permits us to test the cross-time stability of each measure. In the second stage of the analysis, we test the compositional and social control hypotheses using each measure of crime concentration as an outcome variable.

Methods

Data Sources

The study utilizes the archive of dispatches made by the City of Boston's 911 system from 2011 to 2013. Over this time, 1,925,516 dispatches were made. Each dispatch records the location where services were required, not necessarily the location from which the request was made, meaning it documents the emergency itself, as well as date and time the request was received. A record also includes a case type drawn from a standardized list that captures the nature of the issue and the services required. These case types were used to identify 67,792 records that referenced public violence (see Measures for more), 64,140 of which could be linked to a street segment within the city (an effective geocoding rate of 95%).²

The dispatch records were prepared for analysis using the Geographical Infrastructure for the City of Boston (GI; O'Brien et al. 2018), which organizes the city at 17 nested geographical scales. The basis of the GI is the City of Boston's Street and Address Management system and the Tax Assessments database, which track all properties (i.e., the smallest ownable unit) and land parcels (i.e., geographically-bounded lots that contain one or

² All dispatches are immediately geocoded to the City's street and address management (SAM) system at the time of receipt to specify unique latitude and longitude. The Boston Area Research Initiative then uses these latitude–longitude pairs to spatially join to the containing or nearest parcel (i.e., address), using parcel polygons provided by the City and included in BARI's Geographical Infrastructure for the City of Boston (O'Brien et al. 2018). 56,654 violent events were contained within the footprint of an address (84% of events). These were then linked directly to the containing street segment from the Geographical Infrastructure. An additional 7,486 were spatially joined to the nearest street segment (11% of events). All records were within 30 m of the nearest street, indicating strong fidelity in this process. The 5% of events not geocoded all were outside the city or lacked latitude and longitude in the original record.

more properties), respectively. The GI then maps them to U.S. Census TIGER line street segments, defined as the undivided length of street between two intersections or an intersection and a dead end, which are nested within the City's 178 census tracts.³

Street segments are the fundamental unit of measurement in the study. The GI recognizes 24,981 street segments that fall within the city, but the analysis focuses on the 12,524 that have at least one parcel.⁴ We do this because the overwhelming majority of zero-parcel streets experienced zero violent events (95% vs. 40% for streets with one or more parcels), for two reasons: substantively, they are typically short in length or meaningless in terms of land-use (e.g., on-ramp to a highway), meaning violent events are unlikely to occur there; methodologically, the 911 dispatch process seeks to link events to land parcels, meaning that only the small handful of cases lacking land parcel information can be joined to a street with zero parcels. The inclusion of no-parcel streets would thus create a false comparison with other types of streets and an excess of streets with zero violent events, thereby exaggerating concentrations of crime, especially in neighborhoods with an abundance of no-parcel streets. Consequently, the final analysis includes 60,890 violent events (95% of those geocoded) on those 12,524 streets.

The GI makes it possible to incorporate information about streets and tracts from other sources in order to describe census tracts. This study utilizes three such sources. First, the GI itself provides basic land use information. Second, demographic information is drawn from the US Census' American Community Survey 2010–2014 estimates (O'Brien and Ciomek 2017). Third, the Boston Neighborhood Survey (Injury Control Research Center and Boston Area Research Initiative 2019) provides measures of social process. The BNS was a telephone survey that recruited participants by random-digit dial. Its content and methodology was modeled after the community surveys designed by the Project on Human Development in Chicago Neighborhoods (Sampson 2012). Here we use the 2010 wave, which had 1718 adult participants drawn from a list-assisted sampling frame, with separate random probability samples for each of Boston's 16 planning districts (large regions with historical and social significance; response rate = 11%), proportional to population size.

³ Given that some street segments form the border between two tracts, rather than lying clearly inside one or another, the Geographical Infrastructure links each street to the single tract containing its centroid. This process assigns them randomly to one of the census tracts between which they form the border, limiting any systematic bias in the subsequent analyses. One consequence of this method is that 11% of streets have parcels that fall in a census tract other than the one to which they were attributed (i.e., on the other side of the census tract border from the street's centroid; 13% of all parcels). Counts of violent crimes for each street included events occurring at all addresses on the street, regardless of whether they fell in the same tract as the street's centroid. This was deemed as the most appropriate way to deal with disagreements between an addresses' street and tract, as the fundamental unit of analysis was the street segment. Further, it altered the presumed census tract of only 1,097 of violent events (2% of those geocoded).

⁴ Readers will note that these numbers differ somewhat from O'Brien & Winship (2017), which uses the same basic database and analytic structure. The 2018 update of the Boston Area Research Initiative's Geographical Infrastructure of Boston geocoded properties and parcels in a more precise manner that altered the street attribution of a modest number of parcels.

Table 1 Case types and dispatch codes from 911 dispatches composing the indices of violent crime and their frequencies on streets with land parcels, 2011–2013

Public Violence		Prevalence of Guns	
Case Type (Dispatch Code)	Count	Case Type (Dispatch Code)	Count
Assault and Battery in Progress (ABIP)	7942	Assault and Battery with Deadly Weapon (ABDWIP)	311
Assault and Battery Report (ABRPT)	4710	Person with a Gun (PERGUN)	2230
Armed Robbery (ARMROB)	1194	Shots (SHOTS)	2890
Emotionally Disturbed Person: Violent or Injured (EDP2)	19,534	Person Shot (SHOT and PSHOT)	1598
Fight (FIGHT)	17,909		
Person with Knife (PKNIFE)	2572		

Note: Occurring on 8,218 of the 12,524 streets with one or more land parcels

Measures

Concentration of Crime

Previous work with Boston's 911 dispatches used confirmatory factor analysis to develop groupings of case types that act as indices of disorder and crime (O'Brien and Sampson 2015a, b).⁵ Here we combine two such indices that reflect violent crime: *public violence* that did not involve a gun (e.g., fight); and *prevalence of guns*, as indicated by shootings or other incidents involving guns. This is justifiable because the two measures are highly correlated ($r \approx 0.6$ at the street level in all years analyzed; $r \approx 0.8$ when limiting to residential neighborhoods). Table 1 reports constituent case types for both indices and their frequencies for 2011. For each street we tabulate the number of violent events in each year. To aid interpretation and limit the number of spurious outliers, we calculate the following measures only for tracts with at least 5 streets and at least 5 violent crimes in each of the relevant years.

After tabulating the number of violent events on each street in each year, two sets of measures were calculated for each tract: cross-sectional Gini coefficients and stability in crime concentrations. The Gini coefficients followed the classical equation:

$$G = \frac{1}{n} \left(n + 1 - 2 \frac{\sum_{i=1}^n (n + 1 - i) y_i}{\sum_{i=1}^n y_i} \right)$$

where each street i in a census tract with n street segment has a quantity of violent crimes y_i . Bernasco and Steenbeek (2017) noted, however, that the Gini coefficient is ill-suited to situations in which there are fewer instances of the quantity being distributed than there are units to distribute them across (in this case, fewer crimes than places), and proposed a generalized Gini coefficient that handles this issue mathematically. Based on this, when

⁵ The confirmatory factor analyses were based on counts of events of case types for census block groups, maximizing the extent to which case types included in a single category of events (e.g., public violence) were co-incident at this level of geography. We did not re-run the confirmatory factor analysis at streets because the low frequency of events relative to streets (< 1 per street for almost all case types) would make such an analysis hard to interpret as stochasticity can obscure the correlations between types of events.

there are fewer events than units (i.e., $\sum_{i=1}^n y_i < n$), our analysis uses this modified Gini coefficient:

$$G' = \frac{n}{\sum_{i=1}^n y_i} (G - 1) + 1$$

These calculations were conducted in R using the *reldist* package's `gini` command (Handcock 2016) in conjunction with custom functions. G or G' was calculated for each tract for each year.

Stability in crime concentrations were calculated by first sorting the streets of a tract in descending order of number of violent crimes in a year. We then identified all streets accounting for the first 50% of violent crimes in the tract. We included every street through the one that reached or surpassed 50% and then added any additional streets with the same number of violent crimes as the final street because there is no clear way to select one street over another in the case of a tie (i.e., a strict cutoff would create an arbitrary division between inclusion and exclusion). The same process was then completed for the following year, and the final measure was the proportion of streets that contributed to the first 50% of violent crimes in either year that did so in both years.

Other Tract-Level Measures

We incorporate three types of information to generate descriptors of census tracts, for which all descriptive statistics are reported in Table 2. First, the GI provides information on urban form at the street level that we use to measure the institutional or land use diversity implicated by the compositional hypothesis, including identification as a “main” street (provided by MassGIS) and nature of land usage, which is a seven-group typology based on a cluster analysis of the representation of each land use (e.g., the most common street type is dominated by single-family houses, whereas streets with a predominance of parcels that are exclusively commercial are the most common type not dominated by residential; see Table 2 for all categories and their frequency and O'Brien et al. 2018 for method). These are then the basis of aggregate measures for the proportion of main streets and commercial streets (streets in either the Commercial or Mixed-Use Commercial categories). The GI also includes a categorization of census tracts as Residential, Downtown, Institutional (dominated by industry or a college or hospital campus) or Park (dominated by a recreational area), which indicates the overarching land use of the community and can determine the types of frequency of crime occurring there.

Second, we draw population descriptors from the U.S. Census' American Community Survey. First, ethnic composition (e.g., proportion Black, proportion Hispanic, etc.) was used to calculate ethnic heterogeneity, based on the Herfindahl index, $1 - \sum s_i^2$, where s_i is the proportion of residents belonging to ethnicity i . The index represents the likelihood that two randomly selected residents would be of different races. Higher values indicate a more diverse neighborhood, and as such is an additional variable for testing the compositional hypothesis. We also incorporated population density as an additional covariate reflecting land use patterns, and median household income, proportion Black, and proportion Hispanic (log-transformed before inclusion in any regressions to account for a skew distribution) as indicators of current and historical disadvantage.

Third, the BNS provided a measure of collective efficacy per resident perceptions, a critical predictor for testing the social control hypothesis. This measure was composed of

Table 2 Characteristics of streets and tracts included in the analysis

	Mean (SD) or Count (%)	Range	Mean (SD) or Count (%)	Range
<i>Street Segments (n = 12,524)</i>				
<i>Length</i>	96.52 m (67.88 m)	3.32 m – 802.31 m	7.66 (7.73)	1 – 77
<i>Main Street</i>	3,357 (27%)	–	0.74 (2.78)	0 – 92.33
<i>Predominant Zoning</i>				
Three-Family Residential with Assorted Other Uses	1,801 (14%)	–	1,159 (9%)	–
Mix of Two-Family and Single-Family Residential	2,549 (20%)	–	1,482 (12%)	–
Commercial	2,092 (16%)	–	498 (4%)	–
Single-Family Residential Only	3,228 (25%)	–	–	–
Census Tracts (n = 166) ^a				
<i>Number of Streets</i>	74.93 (46.92)	8 – 211	121.74(90.78)	3.33 – 600.33
<i>Population Density (1000 s / mile²)</i>	24.26 (16.79)	0.14 – 110.11	0.23(0.18)	0.00 – 0.82
<i>% Hispanic</i>	0.18 (0.16)	0.00 – 0.77	0.22(0.25)	0.00 – 0.90
<i>Median Household Income</i>	\$58,620 (\$29,583)	\$12,813–\$155,417	0.50(0.18)	0.04 – 0.85
<i>Income inequality (Gini)</i>	0.48 (0.07)	0.30 – 0.70	3.51(0.14)	3.12 – 3.85
<i>Neighborhood Type</i>				
Downtown	11 (7%)	–	8 (5%)	–
Industrial/ Institutional	27 (16%)	–	120 (72%)	–
<i>Concentration of Crime (G⁺; 2011)</i>	0.71 (0.11)	0.24 – 0.90	0.72 (0.11)	0.28 – 0.93
<i>Concentration of Crime (G⁺; 2013)</i>	0.72 (0.10)	0.28 – 0.91	–	–
<i>Stability in Concentrations of Crime (2011–2012)</i>	0.39 (0.22)	0.00 – 1.00	0.38 (0.24)	0.00 – 1.00

^a – Limited to those census tracts included in the regression analysis: those with at least five streets and at least five violent events in at least one year, permitting the calculation of at least one Gini coefficient for distribution of violent crime

two subscales of five items each: social cohesion (e.g., “People in this neighborhood know and like each other.”) and social control (e.g., “How likely is it that your neighborhood would organize together to do something if a child was spray-painting graffiti on a local building?”). The measure was calculated as an aggregate of residents’ responses, controlling for individual-level demographic characteristics (gender, age, ethnicity, and parental status). This methodology was identical to that originally developed by Sampson et al. (1997).

Analysis Plan

The analysis is organized to answer the two main questions posed at the outset. First, we assessed the extent to which concentration of crime is characteristic for a neighborhood. This was done by testing the correlation between the generalized Gini coefficient and the stability in crime concentrations. We also used the three-year span of the data to test the cross-time stability of each measure. In the second stage of the analysis, we use regression models to test the compositional and social control hypotheses with each measure of crime concentration as an outcome variable. These regressions were run in two stages. First, indicators of land use composition, demographic composition, and collective efficacy were entered into the model. Because many of these variables are correlated, we trimmed the models to only significant predictors to ensure that significant findings were not artifactual. All models also controlled for number of crimes and number of streets, as these elements have direct impact on the likely distribution of the Gini coefficient. We treat the model with these two variables alone as the null model when discussing variance explained.

As noted in the Introduction, it is possible for the diversity of features across streets to explain concentrations of crime in either of two ways: the direct relationships between those features and a street’s vulnerability creates variation in the aggregate; the diversity itself exacerbates variations across streets, independent of these features at the street level. To test this, we follow the example of O’Brien (2019), which ran multilevel models to control for street-level characteristics and then recalculated tract-level Gini coefficients based on the number of crimes above expected (i.e., the street-level residuals of the models). This quantifies inequality in the distribution of crime not accounted for by differences in the street-level characteristics included in the model. These characteristics included: land use, based on the composition of parcels on the street; the number of properties and length of the street; land value per sq. ft. of parcels on the street, as a proxy for wealth as census-based income is estimated only for census block groups and higher scales of geography⁶; ethnic composition of the street, as imputed from census data.⁷ Importantly, these models

⁶ This was incorporated as an interaction with land use type of the street, as it has a different interpretation vis-à-vis routine activities and crime for residential streets compared to other types of zonings (e.g., commercial). Indeed, streets dominated by single-family residential parcels with higher value per square foot had fewer crimes than other streets in the neighborhood, whereas the opposite was true for commercial streets.

⁷ This information is not readily accessible from the US Census as demographics are tabulated for census blocks and street blocks form the borders between census blocks. In order to impute information from census blocks to streets blocks, we first determined how many parcels on a street were contained in each bordering census block (per the GI). This was then the basis for a weighted average of the form $p_i = \sum_b \frac{p_{ib} * parcels_b}{\sum_b parcels_b}$, where p_i is the estimated proportion of residents of ethnicity i for the street block, p_{ib} is the same proportion for block b , and $parcels_b$ is the number of parcels on the street in block b . This was done for proportion Asian, Black, Latino, and White.

accounted for the extent to which each of these variables explained differences in the frequency of crime between streets within neighborhoods, rather than across the city as a whole (all results reported in Appendix A). It was not possible to do this for the stability in the location of crime concentrations, however, because a regression-based control would not make sense for a dichotomous measure drawn from a relative ranking of streets.

Results

Are Concentrations of Crime Characteristic of a Neighborhood?

There was substantial variation across tracts in both the generalized Gini coefficients and cross-time concentration stability. In 2011, the Gini coefficient ranged from 0.24 to 0.90, though most values were clustered between 0.60 and 0.85 ($mean=0.71$, $s.d.=0.11$), a range generally considered indicative of a high amount of inequality. Between 2011 and 2012, the average tract had 39% consistency in the streets responsible for the first 50% of violent crimes, but tracts ranged from 0 to 100% on this measure ($s.d.=0.22$). That is to say, 4 tracts (2%) had none of the same streets accounting for the first 50% of violent crimes, and 8 tracts (5%) saw no change in which streets accounted for the first 50% of violent crimes. These statistics were consistent across years for both measures (see Table 2 for all details).

We illustrate the implications for the Gini coefficients in Fig. 1. As can be seen, nearly all census tracts had a rather high level of concentration of crime across streets, but there is substantial variation therein (Fig. 1a). We compare two tracts that have nearly the same exact number of crimes and streets but fall on different sides of that distribution. One, which might be described as having moderately high concentrations of crime, featured a series of minor hotspot streets that collectively account for most of the crime in the neighborhood (Fig. 1b and c). The other had a more exaggerated level of concentration of crime, dominated by a single outlier street that accounted for a massive amount of the neighborhood's violent crime (Fig. 1d and e).

The Gini coefficients correlated substantially across the three years (cross-year $r_s=0.59$, 0.67 , 0.72 , all p -values <0.001), justifying their reduction into a single, multi-item measure (Cronbach's $\alpha=0.85$). Likewise, the two measures of cross-time concentration stability were substantially correlated ($r=0.59$, $p<0.001$). Because the Gini coefficient and concentration stability measures each have an absolute meaning whose interpretation should be consistent across years (e.g., a Gini coefficient of 0.8 has the same meaning in any year), we calculated two multi-year measures as the mean of the annual measures. The multi-year measures of the Gini coefficient and cross-time concentration stability were also correlated, albeit to a lesser extent ($r=0.36$, $p<0.001$), indicating at least partial overlap between tracts with higher concentrations of 911 dispatches for violent crime on certain streets and cross-time stability in which streets those were. As the two measures were not especially strongly correlated, however, we analyze them as separate indicators of concentration of crime moving forward.

Table 3 Parameter Estimates from Regressions using physical and demographic characteristics of census tracts to predict measures of concentration of crime across a tract's streets

	Gini			High-Crime Stability			Adjusted Gini ^a		
	(1)	(2)	Stand. Beta	(1)	(2)	Stand. Beta	(1)	(2)	Stand. Beta
	Beta	Beta	Beta	Beta	Beta	Beta	Beta	Beta	Beta
<i>Nbhd Type^b</i>									
Downtown	0.02 (0.03)	0.07	-	-0.02 (0.05)	-0.04	-	-0.06 (0.04)	-0.12	-
Institutional	0.03 (0.02)	0.12	-	0.00 (0.03)	0.01	-	-0.04 (0.03)	-0.11	-
Park	0.05 (0.03)	0.12	-	0.11* (0.05)	0.15	0.11* (0.05)	0.02 (0.04)	0.03	-
Pop. Dens	-0.001 (0.0004)	-0.14	-	-0.001* (0.0007)	-0.17	-	0.0005 (0.0006)	0.06	-
% Comm. Streets	-0.24** (0.08)	-0.48	-0.17* (0.07)	0.04 (0.11)	0.05	-	-0.09 (0.10)	-0.12	-
Street Hetero ^c	-0.30*** (0.08)	-0.46	-0.27** (0.08)	-0.06 (0.12)	-0.06	-	-0.24* (0.11)	-0.27	-
% Hispanic ^{c,d}	-0.04 (0.07)	-0.05	-	0.16 (0.10)	0.14	-	0.04 (0.09)	0.04	-
% Black ^d	-0.10 (0.05)	-0.21	-	-0.04 (0.08)	-0.05	-	0.02 (0.07)	0.02	-
Med. HH Income	0.00 (0.00)	0.06	-	0.00 (0.00)	.14	-	0.00 (0.00)	0.08	-
Ethnic Hetero	0.17*** (0.05)	0.33	0.16*** (0.04)	0.16* (0.07)	0.20	0.19*** (0.06)	0.13* (0.06)	0.18	0.12* (0.05)
Income Inequality (G)	0.26* (0.13)	0.20	0.30* (0.12)	0.17 (0.18)	0.09	-	0.18 (0.16)	0.10	-

Table 3 (continued)

	Gini				High-Crime Stability				Adjusted Gini ^a			
	(1)		(2)		(1)		(2)		(1)		(2)	
	Beta	Stand. Beta	Beta	Stand. Beta	Beta	Stand. Beta	Beta	Stand. Beta	Beta	Stand. Beta	Beta	Stand. Beta
<i>Collective Efficacy</i>	0.11 ⁺ (0.06)	0.16	0.13* (0.06)	0.19	-0.01 (0.09)	-0.01	-	-	-	-0.00 (0.08)	-0.00	-
R²	0.24		0.16		0.16		0.10		0.11		0.03	

Note: N = 166 census tracts with at least 5 streets and at least 5 violent events in at least one year, meaning a Gini coefficient could be calculated for at least one year. This was further limited to 161 census tracts for the adjusted Gini coefficients, removing any tract that had no street that generated at least one crime above expected

^a-Calculated from the number of crimes above expected on each street, controlling for land use classification, length, number of parcels, parcel value per square ft., and inferred ethnic composition, compared to other streets in the same tract. See Methods for more on measures and Appendix A for parameters from multilevel models on which the adjustments were based

^b- A series of dichotomous variables reflecting a tract's predominant usage, with Residential acting as the reference group

^c- Calculated as $p^*/(1-p)$, where p is the proportion of streets classified as commercial or mixed-use commercial (see Methods for more)

^d- Log-transformed to account of skew distribution

⁺ $-p < .1$, * $-p < .05$, ** $-p < .01$, *** $-p < .001$

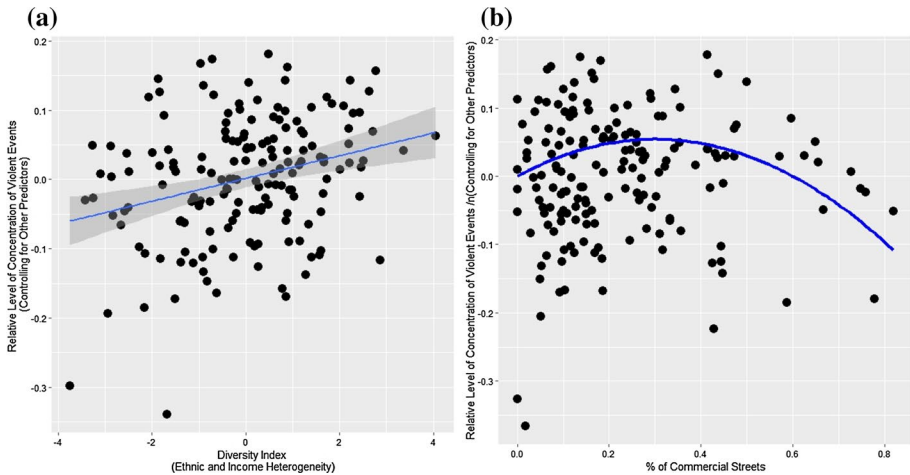


Fig. 2 Scatter plots and fit lines depicting the relationship between the concentration of crime in a census tract and **a** the level of diversity, calculated as a standardized combination of ethnic heterogeneity and income inequality, and **b** the proportion of commercial streets, each controlling for other variables included in the final model

Why Does Crime Concentrate More in Some Neighborhoods?

We used regressions to test the compositional and social control hypotheses for why violent crime, as measured by 911 dispatches, might concentrate more strongly in some neighborhoods than others. These models used land use composition, demographic composition, and collective efficacy to predict variation in each of the measures of concentration of crime (see Analysis Plan for more detail; see Table 3 for all results).

We first analyze the composite Gini coefficients—that is, the average level of unevenness in the distribution of crimes across the streets of a census tracts over the three years analyzed. The full model found that physical composition and demographic composition were significant predictors of the level of concentration of crime. First, the proportion of commercial streets and the mixture of commercial and non-commercial streets were each predictive of lower heterogeneity (% of commercial streets: $B = -0.48$, $p < 0.01$; commercial-non-commercial heterogeneity: $B = -0.46$, $p < 0.001$). Given the specific parameter estimates, these two relationships indicate a parabolic relationship (illustrated in Fig. 2), with the distribution of crime being most uneven in neighborhoods with 29% commercial streets and decreasing as a neighborhood had either more or fewer commercial streets.

In terms of demographic composition, ethnic heterogeneity had the strongest effect in the model, predicting more unevenly distributed crime ($B = 0.33$, $p < 0.001$). Income inequality was also associated with more unevenly distributed crime ($B = 0.20$, $p < 0.05$). This combined effect of demographic diversity is illustrated in Fig. 2. In addition, collective efficacy was marginally associated with higher levels of concentration ($B = 0.16$, $p < 0.10$). These effects remained consistent in the trimmed model, and collective efficacy strengthened to a conventional level of significance ($B = 0.19$, $p < 0.05$).

The analysis of stability in crime concentrations found fewer associations. As with the Gini coefficients, ethnic heterogeneity was associated with higher stability in crime

concentrations in the full model ($B=0.20$, $p<0.05$), and remained significant in the trimmed model. The locations of crime concentrations were also more stable in neighborhoods dominated by large parks ($B=0.15$, $p<0.05$). This is likely an artifact of the fact that such neighborhoods have highly segregated routine activities, often with a handful of residential streets flanking the edges of a public recreational space, meaning the differences between them in crime would likely be stable over time. Population density was associated with less uneven distributions of crime in the initial model, but this fell to a non-significant level when covariates were trimmed.

The Role of Composition: Street-Level Characteristics or a Diverse Context?

The initial tests found that both land use and demographic heterogeneity were predictive of higher concentrations of crime. There are two ways that these relationships might arise, however. First, it could be that concentrations of crime arise from aggregating a set of streets whose basic features lead them to have different levels of vulnerability to crime. Alternatively, it could be that the diversity itself exacerbates variations across streets, over and beyond these street-level characteristics, making concentrations of crime an emergent property of a diverse context. To adjudicate between these interpretations, we calculated adjusted Gini coefficients that accounted for street-level variations in land use, land value (as a proxy for wealth), ethnic composition of residents, and the number of properties and length of the street (see Analysis Plan for more detail). These form the basis for the remainder of the analysis.

We replicated the analyses above for the adjusted Gini coefficients. Again, correlations across years were high enough to justify combination into a single cross-time Gini coefficient ($r_s=0.69$, 0.70 , 0.74 , all p -values <0.001). We were forced, though, to exclude five census tracts with no streets that generated at least one crime above expected, as this would mean that the total number of crimes above expected was zero, generating a nonsensical $G^*=0$. The regressions predicting the newly calculated composite measure saw substantial changes. Ethnic heterogeneity remained predictive of more uneven distributions of crime across a neighborhood's streets, though the effect size was diminished by nearly half ($B=0.18$, $p<0.05$). The mix of commercial and non-commercial land use also had a diminished effect that turned out to be non-significant after the regression was trimmed. Meanwhile, income inequality and collective efficacy were no longer significant predictors in either the full or trimmed models. No new significant predictors emerged in this latter analysis.

It was unsurprising that the effects of some of these measures were diminished in this analysis as they are themselves measures of variation that might be accounted for by the street-level controls. The measure of collective efficacy used here, however, is a neighborhood-level social phenomenon and should not in theory be explained away in the same way by street-level characteristics. To test which of the control variables led it to lose significance, we ran the multilevel models working up from the simplest set of controls: number of parcels and street length. Controlling for only these two variables, recalculating the Gini coefficients, and then re-running the tract-level regressions, we found that collective efficacy had a non-significant relationship with the distribution of crime, $1/4^{\text{th}}$ the magnitude of that observed in the original analysis ($B=0.04$, $p=ns$). Indeed, variation in length of street within a tract was correlated with collective efficacy ($r=0.20$, $p<0.01$), suggesting that the original result might have been an artifact of this relationship.

Discussion

The analysis confirmed that each neighborhood in Boston had a characteristic level of concentration of violent crime across its streets. These relative differences between neighborhoods were stable across time and were also correlated across two different measures of concentration. To be sure, concentration was rather strong across neighborhoods, but existed on a spectrum, ranging from an exaggerated concentration on a few hotspots to instances in which violent crime was more moderately distributed across a handful of streets. This points to concentrations of crime as another dimension of how crime manifests across communities, complementing the more traditional focus on levels of crime.

Variations in the concentration of violent crime, as measured here through 911 dispatches, were best explained by a neighborhood's physical and demographic composition. Those with a greater diversity of land uses or with higher ethnic or socioeconomic heterogeneity tended to have higher concentrations of crime. Of particular interest, ethnic heterogeneity appeared to further exacerbate the uneven distribution of crime beyond what would be expected based on the demographic composition of the neighborhood's streets alone. Meanwhile, a neighborhood's collective efficacy, and thus informal social control, had a moderate relationship with the distribution of crime that further examination suggested was artifactual, though we return to this and its interpretation in the next section. It is worth noting that these multivariate relationships explained only 16% of the variance in concentrations of reported violent crimes, leaving over 80% of that variance unexplained. As we interpret these results, it is important to keep in mind that 911 records are predominantly generated by calls from community members, which means they intermingle objective events with subjective perceptions and actions. This could have implications for how we understand these correlations.

The compositional hypothesis was largely borne out by the results, especially considering that all three indicators of diversity, one pertaining to land use and two pertaining to different aspects of resident demography, were predictive of greater concentrations of reported violent crimes. Ethnic heterogeneity was also associated with higher stability in high-crime streets from year to year. As striking was how these different indicators of heterogeneity differed in their effect. Land use diversity predicted higher concentrations of reported violent crime, reaching a maximum in tracts with 29% commercial streets. But this effect was fully accounted for when recalculating concentrations while factoring out the expected number of events based on land use. Thus, the effect appears entirely owed to the differences between the land use of the streets themselves. In contrast, controlling for the demographic characteristics of streets accounted for only about one-half of the effect of ethnic heterogeneity on concentrations of crime. This suggests that the ethnic diversity in a neighborhood is exacerbating differences in crime frequency across streets.

Why ethnic heterogeneity would further concentrate violent crime (or resident reports thereof) is not immediately clear from this analysis, but we might speculate as to why this would be true. We propose three different hypotheses and use the simple thought experiment of two ethnicities to illustrate. The first is that the mere proximity of the two groups increases the vulnerability of the one that is more prone to crime. Some have argued that demographic diversity can lead to elevated crime among the population with less access to resources, either monetary or institutional. Not only can this manifest in property crime as they seek alternative ways to “keep up” with their neighbors (Kling, Jens, and Katz 2005), but also in generalized stress that leads to violent crime (Wilson and Daly 1997). A second explanation derives from theories of interethnic conflict (Banton 1983; Blalock 1967), though it might take either of two forms. In one interpretation, patterns of crime reporting

are responsible for the relationship, as opposed to the occurrence of crime itself. Legewie and Schaeffer (2016) found that neighborhoods with higher ethnic heterogeneity generated more complaints about public disorder (e.g., noise complaints). These complaints were especially high along “border” streets where residential enclaves of two or more groups interfaced, creating the same sorts of concentrations being captured here. It is also possible that one population is more likely to find the actions of another group more threatening, reporting them as violent events, especially if there is an imbalance in the material or institutional capital of the two groups. In this case the concentrations of crime would not necessarily occur on border streets but on the streets inhabited by the less resourced population. The alternative explanation is that border streets arise from concentrations in actual crime, as violent conflict could be occurring directly between the members of different ethnic groups (e.g., via gangs). Either of these explanations could underlie the findings in this study as our measures of interest were drawn from 911 reports, which ostensibly describe objective events but are filtered through the decision by local residents and passers-by to report them.

Third, the geographic arrangement of the two ethnic groups—reflecting their level of microsegregation—might combine with spillover effects to explain the greater unevenness in crime across streets without requiring any additional social or behavioral processes. We describe differences in microsegregation with two neighborhoods in Boston that feature a mix of Latinx and White residents; it is important to note, though, that we use these only for that illustrative purpose and do not reference their actual crime concentrations as measured in the study in order to avoid providing anecdotal evidence in one direction or the other. Also, we assume the positive correlation between a street’s non-White population and the level of crime or crime reporting seen in our street-level controls (see Appendix). In the most extreme case of microsegregation, the two sets of streets will sit entirely apart in two contiguous clusters. For instance, in the Jackson Square section of the Jamaica Plain neighborhood the Latinx population and the White population are largely residentially separated, with the former living primarily to the north and the latter living to the south of a major thoroughfare. This creates a minimal number of “border streets.” Let us then suppose that there is natural spillover in crime between adjacent streets. In this case, the crime (or crime reporting) on streets occupied by the population that is more vulnerable to crime will spillover almost exclusively to streets dominated by the same ethnicity, creating reinforced concentrations of crime. In contrast, the neighborhood of East Boston features a geographically integrated mix of Latinx and White residents. In this case, spillover would at times occur between streets that are ethnically dissimilar, evening out the overall distribution.

The current analysis does not adjudicate between these three hypotheses. It is also possible that more than one of these processes could be in force in a given neighborhood, and that different ones are relevant for land use and ethnic heterogeneity. This should be a focus of future research.

Lessons and Limitations

As this was the first formal examination of variations in concentrations of crime across neighborhoods, we consider how work on this subject might proceed, with an eye toward open questions, methodology, and limitations. First, we turn from what was explained in the regression models to the large amount of variation that went unexplained. Clearly, there are explanatory variables missing. There are many mechanisms that contribute to the level of crime in a community (Pratt and Cullen 2005), and the same is likely to be true for

concentrations of crime. The two hypotheses we presented and tested here are just an initial step and will necessarily be followed by other examinations. One that we want to suggest is an alternative formulation of the social control hypothesis. We took the traditional tack of treating collective efficacy as a neighborhood-level process, and then reasoning how it might be leveraged by residents to influence concentrations of crime. There is increasing evidence, however, that the level or activation of collective efficacy varies across streets and even institutions within a neighborhood (Browning et al. 2017; Weisburd et al. 2017, 2013, 2020). This could mean that streets with lower collective efficacy are more vulnerable to crime than others in the neighborhood, creating concentrations of crime in a manner analogous to varied land use composition.

There is also value in considering whether some of the variance in concentrations of crime might not be systematically explicable via street and neighborhood characteristics. Could it be that idiosyncrasies of individual streets are largely responsible for the level of concentration of crime in a neighborhood? There is reason to believe that this could be true. It is well established that certain land uses and institutions are more vulnerable to crime than others. There is substantial evidence, though, that even within “high-crime” land uses, only a handful of places account for the vast majority of crimes (Eck et al. 2007). St. Jean (2007) has illustrated this phenomenon ethnographically through interviews with drug dealers in Chicago. His interviewees repeatedly pointed out small details that made a specific street corner a good place to “work”: design features that lowered the chance of capture by police, access to quality consumers, avoidance of nosy neighbors. Chief among these was a location’s historical success as a lucrative place to do business (see also Olaghere and Lum 2018). This last point resonates strongly with work on repeat victimization, which posits that a crime at a location might either “flag” its suitability for crime, or “boost” its vulnerability to future crimes (Johnson 2008). Whether prior crime acts as a flag or boost, however, the critical point here is that the place in question is more likely to reexperience crime. Depending on the historic ordering of events, it would seem possible that these sorts of recurrences could be responsible for differing levels of concentration of crime across neighborhoods. Further, just as the neighborhood-level concentration appears to be largely subject to the features of the individual streets therein, each street’s own level and concentration of crime could be largely dependent on the individual residences and institutions located on it.

From a methodological perspective, we examined two different metrics used by previous research to study variations in concentrations of crime: the generalized Gini coefficient for quantifying unequal distributions of crime, and cross-time stability in the locations of crime concentrations (what Hipp and Kim (2017) referred to as a temporally-adjusted measure of crime concentration). We would argue that the generalized Gini coefficient performed better as a metric for the purposes here, owing to its simplicity, interpretability, and greater association with theoretically meaningful predictors. Though the stability of the locations of crime concentrations is a clever way to factor out natural randomness in the placement of crime, it was more difficult to interpret. In fact, one might debate whether its relationship to the topic of interest is merely a function of its inherent correlation with the level of crime concentration. Places with higher crime concentrations likely have particularly active hotspot streets, and it is well-established that such hotspots maintain their level of crime over time (Weisburd et al. 2004). This would in turn drive up the stability in the locations of concentrations across years. This might also be why that measure was associated only with ethnic heterogeneity out of an array of potentially related covariates, which was also the measure most consistently associated with the generalized Gini coefficient. For these reasons, we suggest that future research concentrate on the Gini coefficient.

Turning specifically to limitations of the current study, this analysis was intended as a first step in examining how concentrations of crime vary across neighborhoods. It speaks only to one type of crime in one city. The results will require replication for other crime types and in other locales, especially of different sizes or in other countries. Each of these variables will introduce greater nuance to the interpretations here. Second, we used 911 dispatches as a proxy for violent crime, but it is well established that neighborhoods can have different propensities for reporting issues (Klinger and Bridges 1997; O'Brien et al. 2015), and it would be best to replicate this work with other measures of crime that are not as subject to such tendencies, like crime reports or victimization surveys. Third, we described at the outset our choice of census tracts as the best way to approximate neighborhoods for the purposes of this study. That said, they still have their weaknesses, especially that some of the boundaries fail to reflect socially salient divisions between communities. It would be useful for future work to assess more closely how different logics for defining neighborhoods might alter the estimations of concentrations of crime. Last, as we stated at the beginning of this section, there is clearly a need for future studies that either (a) expand the set of variables that might be theoretically associated with levels of concentration of crime, or (b) demonstrate how concentrations of crime might emerge over time without consideration for compositional or social characteristics of the neighborhood.

Conclusion

Looking forward, there are two observations worth making. First is that the concentration of crime is a meaningful aspect of how crime manifests in a community, complementing the more traditional focus on levels of crime in a manner analogous to the paired measures of socioeconomic status and income inequality. Here we see initial evidence of the factors that underlie variations in these concentrations, which we anticipate will be expanded upon, but this raises additional questions about the consequences that higher and lower concentrations of crime hold for residents. For instance, are residents more able to avoid crime in neighborhoods with higher concentrations of crime than in those where the crime is more evenly distributed? What impacts might this have for fear of crime, stress, and related sequelae? This will be an important topic for further inquiry.

Second, we might return to the broader inspiration of criminology of place and its tension with traditional, neighborhood-centric perspectives. At times this has given rise to an either-or debate of “places vs. communities” in determining the proper geographic scale for studying and responding to crime. Here, though, we see that the factors explaining concentrations of crime operated at both scales. Although the features of a neighborhood’s streets were largely responsible for variations in concentration, a neighborhood’s ethnic heterogeneity appeared to exacerbate concentrations of crime across streets, above and beyond what would be expected from the characteristics of the streets alone. These insights on how concentrations of crime vary across communities align well with an ongoing theme in recent research: the primary dynamics surrounding crime are locally situated, but higher geographic scales remain relevant.

Appendix

See Table 4.

Table 4 Complete parameter estimates from multilevel models predicting counts of reports of violent crime on a street relative to other streets in the same neighborhood

	2011		2012		2013	
	Beta (Std. Error)	Odds Ratio	Beta (Std. Error)	Odds Ratio	Beta (Std. Error)	Odds Ratio
	<i>Length^a</i>	0.55*** (0.01)	1.73	0.57*** (0.01)	1.77	0.53*** (0.01)
<i># Properties</i>	0.002*** (0.0002)	1.00	0.003*** (0.0002)	1.00	0.002*** (0.0001)	1.00
<i>Main Street^b</i>	0.58*** (0.02)	1.79	0.46*** (0.02)	1.58	0.45*** (0.02)	1.57
<i>Land Usage^c</i>						
3-Family Mixed	0.96*** (0.04)	2.61	0.90*** (0.04)	2.46	0.87*** (0.04)	2.39
2-Family w/ single-family	0.66*** (0.04)	1.93	0.62*** (0.04)	1.86	0.58*** (0.04)	1.79
Pure Commercial	1.28*** (0.04)	3.60	1.22*** (0.04)	3.39	1.23*** (0.04)	3.42
Exempt	1.13*** (0.04)	3.10	1.09*** (0.05)	2.97	1.02*** (0.04)	2.77
Condominiums	0.58*** (0.05)	1.79	0.61*** (0.05)	1.84	0.60*** (0.04)	1.82
Mixed Commercial	1.17*** (0.05)	3.22	1.00*** (0.05)	2.97	1.06*** (0.04)	2.89
<i>No Population^b</i>	0.47*** (0.05)	1.60	0.67*** (0.05)	1.95	0.34*** (0.05)	1.40
<i>% Asian^d</i>	0.02 ⁺ (0.01)	1.02	0.03** (0.01)	1.03	0.05*** (0.01)	1.05
<i>% Black^d</i>	0.18*** (0.01)	1.20	0.18*** (0.01)	1.20	0.16*** (0.01)	1.17
<i>% Latino^d</i>	0.13*** (0.01)	1.14	0.20*** (0.01)	1.22	0.13*** (0.01)	1.14
<i>Value per ft.^{2e}</i>	-0.39*** (0.08)	0.68	-0.39*** (0.08)	0.68	-0.33*** (0.07)	0.72
<i>Value per ft.² × Land Usage^e</i>						
3-Family Mixed	0.22* (0.10)	1.25	0.17 ⁺ (0.09)	1.19	-0.05 (0.08)	0.95
2-Family w/ single-family	0.70*** (0.10)	2.01	0.59*** (0.09)	1.80	0.39*** (0.08)	1.48
Pure Commercial	0.60*** (0.09)	1.82	0.67*** (0.08)	1.95	0.55*** (0.07)	1.73
Exempt	0.97*** (0.09)	2.64	1.05*** (0.08)	2.86	0.99*** (0.07)	2.69
Condominiums	0.20* (0.10)	1.22	0.22* (0.10)	1.25	0.09 (0.09)	1.09
Mixed Commercial	0.41*** (0.09)	1.51	0.67*** (0.09)	1.95	0.47*** (0.08)	1.60

Table 4 (continued)

$N = 11,899$ street segments with at least one property nested in 178 census tracts. The models excluded 634 streets with no property for which average value could be calculated and 1 street whose single property falls outside of the census bounds of the city and thus has no census tract to which it can be attributed

^a100 s of meters

^bA dichotomous variable with '1' equal to variable name

^cA series of dichotomous variables reflecting a street's predominant land usage, based on a cluster analysis of land use types. Single-Family Residential is the reference group

^dLog-transformed to account for skew. Ethnic composition variables predict expected crime frequency relative to a street of 100% White residents. See footnote 8 in main text for description of calculation at the street segment level

^eCalculated from property assessments with values standardized by land use type as different land uses have inherently different valuations. Direct effect is the predicted impact on streets dominated by Single-Family Residential

* $-p < .05$, ** $-p < .01$, *** $-p < .001$

Funding Funding was provided by Division of Social and Economic Sciences (Grant Numbers SES-1637124, SES-1921281).

References

- Andresen MA, Malleon N (2011) Testing the stability of crime patterns: Implications for theory and policy. *J Res Crime Delinq* 48(1):58–82
- Banton M (1983) Racial and ethnic competition. Cambridge University Press, Cambridge
- Bernasco W, Steenbeek W (2017) More places than crimes: implications for evaluating the law of crime concentration at place. *J Quant Criminol* 33:451–467
- Blalock HM (1967) Toward a theory of minority-group relations. Capricorn Books, New York
- Braga AA, Hureau DM, Papachristos AV (2011) The relevance of micro places to citywide robbery trends: A longitudinal analysis of robbery incidents at street corners and block faces in Boston. *J Res Crime Delinq* 48(1):7–32
- Braga AA, Papachristos AV, Hureau DM (2010) The Concentration and Stability of Gun Violence at Micro Places in Boston, 1980–2008. *J Quant Criminol* 26:33–53
- Brantingham PL, Brantingham PJ (1993) Environment, routine, and situation: toward a pattern theory of crime. In: Clarke RV, Felson M (eds) *Routine Activity and Rational Choice*. Transaction Publications, New Brunswick, NJ
- Browning CR, Calder CA, Soller B, Jackson AL, Dirham J (2017) Ecological networks and neighborhood social organization. *Am J Sociol* 122(6):1939–1988
- Bursik RJ, Grasmick H (1993) *Neighborhoods and Crime: The Dimensions of Effective Community Control*. Lexington Books, New York
- Center, Injury Control Research, and Boston Area Research Initiative (2019) Boston neighborhood survey boston area research initiative harvard dataverse
- Cohen LE, Felson M (1979) Social change and crime rate trends: A routine activity approach. *Am Sociol Rev* 44(4):588–608
- Colabianchi N, Dowda M, Pfeiffer KA, Porter DE, Almeida MJC, Pate RR (2007) Towards an understanding of salient neighborhood boundaries: Adolescent reports of an easy walking distance and convenient driving distance. *Int J Behav Nutr Phys Activity* 4(1):66
- Coulton CJ, Korbin J, Chan T, Marilyn Su (2001) Mapping residents' perceptions of neighborhood boundaries: A methodological note. *Am J Community Psychol* 29(2):371–383
- Curiel RP, Delmar SC, Steven Richard Bishop. (2018) Measuring the distribution of crime and its concentration. *J Quant Criminol* 34:775–803
- Curman ASN, Andresen MA, Brantingham PJ (2015) Crime and place: A longitudinal examination of street segment patterns in Vancouver, BC. *J Quant Criminol* 31:127–147
- Durkheim E (1895/1964) *The Rules of Sociological Method*. Translated by S. A. Solovay and J. H. Mueller. New York: Free Press.
- Eck JE, Clarke RV, Guerette RT (2007) Risky facilities: Crime concentration in homogeneous sets of establishments and facilities. *Crime Prev Stud* 21:225–264
- Farrell G, Pease K (2001) *Repeat Victimization*. Criminal Justice Press, Monsey, NY
- Glas I, Engbersen G, Snel E (2019) Going spatial: applying egohoods to fear of crime research. *Br J Criminol* 59(6):1411–1431
- Groff E, Weisburd D, Yang S-M (2010) Is it important to examine crime trends at a local “micro” level: A longitudinal analysis of street to street variability in crime trajectories. *J Quant Criminol* 26:7–32
- Guest AM, Lee BA (1984) How urbanites define their neighborhoods. *Popul Environ* 7(1):32–56
- Handcock MS (2016) reldist: relative distribution methods (Version 1.6-6). CRAN. Retrieved from <http://www.stat.ucla.edu/~handcock/RelDist>
- Hipp JR, Kim Y-A (2017) Measuring crime concentration across cities of varying sizes: Complications based on the spatial and temporal scale employed. *J Quant Criminol* 33(3):595–632
- Johnson SD (2008) Repeat victimisation: A tale of two theories. *J Exp Criminol* 23:201–219
- Johnson SD, Bernasco W, Bowers KJ, Elffers H, Ratcliffe J, Rengert G, Townsley M (2007) Space-time patterns of risk: A cross national assessment of residential burglary victimization. *J Quant Criminol* 23:201–219
- Kling JR, Jens L, Katz LF (2005) Neighborhood effects on crime for female and male youth: evidence from a randomized housing voucher experiment. *Q J Econ* 120(1):87–130
- Klinger DA, Bridges GS (1997) Measurement error in calls-for-service as an indicator of crime. *Criminology* 35(4):705–726
- Lee YJ, Eck JE, Soohyun O, Martinez NN (2017) How concentrated is crime at places? A systematic review from 1970 to 2015. *Crime Science* 6. <https://doi.org/10.1186/s40163-017-0069-x>

- Legewie J, Schaeffer M (2016) Contested boundaries: Explaining where ethnoraical diversity provokes neighborhood conflict. *Am J Sociol* 122(1):125–161
- Loeffler C, Flaxman S (2018) Is gun violence contagious? A spatiotemporal test. *J Quant Criminol* 34(4):999–1017
- Mohler G, Jeffrey Brantingham P, Carter J, Short MB (2019) Reducing bias in estimates for the law of crime concentration. *J Quant Criminol* 35:747–765
- O'Brien D, Ciomek A (2017) Massachusetts Census Indicators, edited by B. A. R. Initiative
- O'Brien D, Sampson RJ (2015a) Public and Private Spheres of Neighborhood Disorder: Assessing Pathways to Violence Using Large-Scale Digital Records. *J Res Crime Delinq* 52:486–510
- O'Brien DT (2019) The action is everywhere, but greater at more localized spatial scales: Comparing concentrations of crime across addresses, streets, and neighborhoods. *J Res Crime Delinq* 56(3):339–377
- O'Brien DT, Phillips N, De Benedictis-Kessner J, Shields M, Sheini S (2018) 2018 Geographical Infrastructure for the City of Boston, edited by B. A. R. Initiative
- O'Brien DT, Sampson RJ (2015b) Public and private spheres of neighborhood disorder: Assessing pathways to violence using large-scale digital records. *J. Res. Crime Delinq.* 52(4):486–510
- O'Brien DT, Sampson RJ, Winship C (2015) Econometrics in the age of big data: Measuring and assessing “broken windows” using administrative records. *SociolMethodol* 45:101–147
- O'Brien DT, Winship C (2017) The gains of greater granularity: The presence and persistence of problem properties in urban neighborhoods. *J Quant Criminol* 33:649–674
- Olaghere A, Lum C (2018) Classifying “micro” routine activities of street-level drug transactions. *J Res Crime Delinq* 55(4):466–492
- Pierce GL, Spaar S, Briggs LR (1988) The character of police work: Strategic and tactical implications. Center for Applied Social Research, Northeastern University, Boston, MA
- Pratt TC, Cullen FT (2005) Assessing macro-level predictors and theories of crime: a meta-analysis. *Crime and Justice* 32:373–450
- Sampson RJ (2012) *Great American City: Chicago and the Enduring Neighborhood Effect*. University of Chicago Press, Chicago
- Sampson RJ, Raudenbush S, Earls F (1997) Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science* 277:918–924
- Schnell C, Braga AA, Piza EL (2017) The influence of community areas, neighborhood clusters, and street segments on the spatial variability of violent crime in Chicago. *J Quant Criminol* 33:469–496. <https://doi.org/10.1007/s10940-016-9313-x>
- Shaw C, McKay H (1942/1969). *Juvenile Delinquency and Urban Areas*. University of Chicago Press, Chicago
- Sherman LW, Gartin PR, Buerger ME (1989) Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology* 27(1):27–55
- St. Jean, Peter K.B. (2007) *Pockets of Crime: Broken Windows, Collective Efficacy, and the Criminal Point of View*. University of Chicago Press, Chicago, IL
- Steenbeek W, Weisburd D (2016) Where the action is in crime? An examination of variability of crime across different spatial units in The Hague, 2001–2009. *J Quant Criminol* 32(3):449–469
- Trickett A, Osborn DR, Seymour J, Pease K (1992) What is different about high crime areas? *Br J Criminol* 32(1):81–89
- Weisburd D (2015) The law of crime concentration and the criminology of place. *Criminology* 53(2):133–157
- Weisburd D, Bushway S, Lum C, Yang S-M (2004) Trajectories of crime at place: A longitudinal study of street segments in the city of Seattle. *Criminology* 42:283–322
- Weisburd D, Groff ER, Yang S-M (2012) *The Criminology of Place: Street Segments and Our Understanding of the Crime Problem*. Oxford University Press, New York
- Weisburd D, Groff E, Yang S-M (2013) Understanding and controlling hot spots of crime: the importance of formal and informal social controls. *PrevSci* 15(1):31–43
- Weisburd D, Shay M, Amram S, Zamir R (2017) The relationship between social disorganization and crime at the micro geographic level: Findings from Te Aviv-Yafo using Israeli census data. *AdvCriminol Theory* 22:97–120
- Weisburd D, White C, Wooditch A (2020) Does collective efficacy matter at the micro geographic level?: findings from a study of street segments. *Br J Criminol* 60(4):873–891
- Wilson M, Daly M (1997) Life expectancy, economic inequality, homicide, and reproductive timing in Chicago neighbourhoods. *BMJ* 314:1271–1274