



Multiscale spatiotemporal patterns of crime: a Bayesian cross-classified multilevel modelling approach

Matthew Quick¹

Received: 19 October 2018 / Accepted: 9 July 2019 / Published online: 2 August 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Characteristics of the urban environment influence where and when crime events occur; however, past studies often analyse cross-sectional data for one spatial scale and do not account for the processes and place-based policies that influence crime across multiple scales. This research applies a Bayesian cross-classified multilevel modelling approach to examine the spatiotemporal patterning of violent crime at the small-area, neighbourhood, electoral ward, and police patrol zone scales. Violent crime is measured at the small-area scale (lower-level units) and small areas are nested in neighbourhoods, electoral wards, and patrol zones (higher-level units). The cross-classified multilevel model accommodates multiple higher-level units that are non-hierarchical and have overlapping geographical boundaries. Results show that violent crime is positively associated with population size, residential instability, the central business district, and commercial, government-institutional, and recreational land uses within small areas and negatively associated with civic engagement within electoral wards. Combined, the three higher-level units explain approximately fifteen per cent of the total spatiotemporal variation of violent crime. Neighbourhoods are the most important source of variation among the higher-level units. This study advances understanding of the multiscale processes influencing spatiotemporal crime patterns and provides area-specific information within the geographical frameworks used by policymakers in urban planning, local government, and law enforcement.

Keywords Spatiotemporal · Crime pattern · Multilevel model · Neighbourhood · Cross-classified data

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10109-019-00305-2>) contains supplementary material, which is available to authorized users.

✉ Matthew Quick
matthewquick@asu.edu

¹ School of Geographical Sciences and Urban Planning, Arizona State University, 975 South Myrtle Avenue, Tempe, AZ 85287-5302, USA

JEL Classification C11 · C33 · R23 · R58

1 Introduction

Spatiotemporal crime patterns are influenced by characteristics of the urban environment at multiple spatial scales (Ouimet 2000; Wooldredge 2002; Boessen and Hipp 2015). Studies that explore local crime patterns, however, often analyse cross-sectional data for a single set of geographical units. Focusing on one spatial scale overlooks the complex spatial structure of urban areas and does not account for the relationships between crime and sociodemographic, political, and built environment characteristics across multiple spatial scales (Sampson 2013). From a theoretical perspective, analysing local crime patterns at two or more spatial scales provides insight into the crime-generating processes that arise over different geographical contexts and helps to distinguish which spatial scale is most important for understanding where and when crime events occur (Taylor 2015; Steenbeek and Weisburd 2016). From a policy perspective, incorporating the multiple geographical frameworks used by local government and law enforcement into quantitative analyses enables policy-relevant information to be estimated and the most suitable spatial scales for crime prevention interventions to be assessed.

This research applies a Bayesian cross-classified multilevel modelling approach to examine the spatiotemporal patterning of violent crime over five years at the small area, neighbourhood, electoral ward, and police patrol zone scales. Crime data are measured at the small-area scale (lower-level units) and small areas are nested in neighbourhoods, electoral wards, and patrol zones (higher-level units). Neighbourhoods, electoral wards, and patrol zones are non-hierarchical such that the set of small areas nested in one neighbourhood may also be nested in two or more electoral wards and two or more patrol zones (Goldstein 1994; Browne et al. 2001). For the spatiotemporal analysis of crime, cross-classified multilevel models provide a framework for integrating two or more higher-level contexts with overlapping geographical boundaries, for estimating the effects of observed and latent covariates at both lower and higher levels, and for quantifying the degree to which the spatiotemporal variation of crime is explained by each set of geographical units.

This study illustrates the first application of a multilevel cross-classified model to analyse the spatiotemporal patterning of crime. In this study, violent crime was found to be positively associated with sociodemographic, built environment, and civic engagement covariates at multiple scales, and neighbourhoods, electoral wards, and patrol zones were found to account for approximately fifteen per cent of the total spatiotemporal variation of violent crime. This advances past research that characterizes the distribution of crime at one spatial scale by showing that local crime patterns are simultaneously influenced by characteristics of the urban environment at multiple scales (Ouimet 2000; Wooldredge 2002). Also, this study extends past multilevel analyses of crime patterns by estimating the area-specific effects for multiple overlapping higher-level units that are each relevant for theoretical inference and for policy development in urban planning (neighbourhoods), local government (electoral wards), and law enforcement (patrol zones) (Steenbeek and Weisburd 2016;

Schnell et al. 2017). In the following sections of this paper, the theories and methods used to explain and analyse multiscale crime patterns in past research are reviewed, the data and the Bayesian multilevel modelling approach are detailed, the results of this study are shown, and the contributions of this study for theory and crime prevention policy are discussed.

2 Theoretical review

Local spatial and spatiotemporal patterns of violent crime are commonly explained by social disorganization theory, collective efficacy theory, and routine activity theory (Miethe et al. 1991; Braga and Clarke 2014). Social disorganization theory hypothesizes that structural characteristics influence the development and maintenance of resident-based informal social control, which, in turn, shapes the degree to which community members mobilize to control criminal behaviour (Shaw and McKay 1942). Informal social control is defined as the capacity to develop and maintain a common set of values and norms and is operationalized by variables measuring socioeconomic disadvantage, residential mobility, and ethnic heterogeneity, as high levels of these characteristics are thought to challenge the formation of social ties between residents and limit the degree to which community members can establish shared values and norms (Sampson and Groves 1989; Veysey and Messner 1999; Kubrin and Weitzer 2003). While social disorganization theory was originally proposed to describe the residential locations of juvenile delinquents in Chicago at the neighbourhood scale (Shaw and McKay 1942), past research has applied social disorganization theory to explain the geographical distribution of violent crime offences across a variety of spatial scales, including municipally defined neighbourhoods, neighbourhood clusters (aggregations of multiple census tracts), census tracts, and smaller census area units (Ouimet 2000; Weisburd et al. 2012; Sutherland et al. 2013; Law et al. 2015).

Elaborating on the ways in which informal social control is established and enforced within and between communities, the systemic model of social disorganization contends that social control functions at private, parochial, and public levels. Private social control manifests through the intimate relationships between friends and family, parochial social control results from the non-intimate relationships between community residents, and public social control is established through the relationships between communities and extra-local organizations (Bursik Jr. and Grasmick 1993). Geographically, the three levels of social control are hierarchical; private social control is exercised at the micro-scale within households or friendship networks, parochial social control operates at the meso-scale within small-area units, and public social control functions within larger geographical units, such as municipally defined neighbourhoods or community areas (Taylor 1997; Wooldredge 2002). Distinguishing between the meso- and macro-scales of social control, in particular, past studies have suggested that parochial social control is most appropriately inferred via sociodemographic structural characteristics for small areas and that public social control can be operationalized by variables that capture community-level civic engagement and/or actions that work to secure political and economic

resources from local government and law enforcement (Velez 2001; Kubrin and Weitzer 2003; van Wilsem et al. 2006).

Collective efficacy theory hypothesizes that crime patterns are explained by both informal social control and the willingness of residents to intervene on behalf of the common good (Sampson et al. 1997). Collective efficacy theory extends social disorganization theory by recognizing that local criminal behaviour is shaped by the common values shared among residents as well as the degree to which community members will take task-specific actions to achieve collective goals, such as living in a safe environment (Sampson et al. 1997; Morenoff et al. 2001). Predominately operationalized for groups of census tracts and municipally defined neighbourhoods, collective efficacy research often analyses the structural characteristics highlighted by social disorganization theory as well as survey data that asks about social cohesion and perceptions that community members will intervene in suspicious, disorderly, or criminal behaviour (Sampson et al. 1997; Sutherland et al. 2013). When representative survey data for all geographical units within an urban area are unavailable, however, researchers have inferred collective efficacy via variables that capture local civic engagement, such as the per cent of active voters, because this is an indicator of the degree to which residents engage in public affairs and take action to achieve shared goals (Weisburd et al. 2012). Related, civic engagement has also been highlighted as a dimension of social capital, or the cooperative relationships between people that facilitate action towards collective goals, with past studies showing that the per cent of active voters is negatively associated with crime after accounting for social disorganization covariates (Rosenfeld et al. 2001; Coleman 2002).

The third theoretical perspective used to explain local crime patterns is routine activity theory. Routine activity theory contends that crime offences occur when motivated offenders, suitable targets, and a lack of capable guardianship converge in space and time (Cohen and Felson 1979). Compared to social disorganization and collective efficacy theories, which focus on the social dynamics within neighbourhoods, routine activity theory centres on how the behavioural activities of potential offenders and potential victims interact with characteristics of the physical environment. Situating routine activity theory at multiple spatial scales, Brantingham and Brantingham (1993) propose that crime patterns are simultaneously influenced by activity nodes, activity paths, and the environmental backcloth. Activity nodes are specific locations where large populations come together for daily activities—such as non-residential areas used for employment, school, or shopping—activity paths are the travel routes between activity nodes—such as public transit stations and major roads—and the environmental backcloth is composed of the broader social, political, and physical contexts in which activity nodes and paths are located (Groff et al. 2010; Deryol et al. 2016). Broadly, past research has found that areas with high traffic activity nodes and/or paths have relatively higher crime rates than areas without nodes and/or paths or areas with low traffic nodes and/or paths (Wilcox and Eck 2011).

Combined, social disorganization, collective efficacy, and routine activity theories provide a theoretical background for understanding the multiscale structure of local crime patterns. Consider, for example, a group of adjacent small areas nested in larger zones used for urban planning and law enforcement purposes. The larger

zones have overlapping geographical boundaries such that the group of small areas nested in one planning zone are simultaneously nested in two different law enforcement zones (i.e. the larger zones are non-hierarchical). In each small area, violent crime may be influenced by the presence of a high traffic activity node and corresponding convergences between offenders and targets (routine activity theory), as well as structural characteristics and informal social control (social disorganization and collective efficacy theories). In addition to the small-area processes, however, there may be additional high (or low) clustering of crime common to the small areas nested in the urban planning zone due to planning policy (e.g. similarities in land use composition, housing, or the presence of an activity path) or public social control (e.g. place-based resources attained from local government). Furthermore, there may be distinct clustering of crime among the small areas nested in each law enforcement zone that is attributable to differences in law enforcement practices (e.g. frequent and proactive police patrols) or the relationships between police and community members (public social control).

3 Methods for analysing multiscale crime patterns

Existing studies that examine spatial and spatiotemporal crime patterns across multiple spatial scales have adopted four methodological approaches: the spatial point pattern test, single-level cluster detection methods, single-level regression models, and multilevel models of purely hierarchical data (i.e. lower-level spatial units nested in one higher-level unit). The spatial point pattern test quantifies the similarity of two geographically referenced point datasets at the area scale by iteratively sampling a subset of points from one dataset (i.e. crime for one year at one scale), establishing area-specific confidence intervals based on the sampled data, and calculating the per cent of areas for which the second dataset (i.e. crime for a different year at the same scale) fall within the confidence intervals from the sampled dataset. For multiscale analysis, the spatial point pattern test has been used to examine if the similarity of between-year crime patterns for large areas is different from the similarity of between-year crime patterns for the nested smaller areas (Andresen and Malleson 2011). While the spatial point pattern test helps to assess if there is variation in crime change at different scales, it does not explain the spatiotemporal patterning of crime through observed or latent covariates across any of the scales (Steenbeek and Weisburd 2016).

Past studies that explore multiscale crime patterns via single-level cluster detection methods typically compare the locations, sizes, and shapes of clusters identified separately for two or more scales. For example, Andresen (2011) used local Moran's I to compare violent crime clusters for two areal scales using ambient and residential populations as crime rate denominators, finding that, while the cluster locations were similar for both scales, the smaller-scale clusters were relatively more sensitive to the crime rate denominator. Similarly, studies that use single-level regression models generally compare model results and diagnostics from separate models fit to data aggregated at different scales. Ouimet (2000), for example, applied separate regression models to explore juvenile violent crime rates for census tracts and larger

municipally defined neighbourhoods and found that the neighbourhood-level model estimated larger regression coefficients and had greater explanatory power. Comparing and contrasting the results of single-level analyses provides insight regarding the scale at which risk factors are associated with crime, however, these approaches do not account for the simultaneous effects of risk factors operating across multiple spatial scales.

Multilevel modelling approaches provide a framework for analysing hierarchically structured data where lower-level units are nested in higher-level units. To date, few studies have applied multilevel modelling approaches to explore the multi-scale spatiotemporal patterns of crime for a comprehensive set of geographical units in an urban area, and instead, the most common use of multilevel models has been to examine the interactions between individual or household characteristics and neighbourhood contexts (van Wilsem et al. 2006; Taylor 2015). Focusing on a comprehensive set of units in a city, Steenbeek and Weisburd (2016) and Schnell et al. (2017) both used a three-level linear mixed model to examine total crime for street segments, neighbourhoods, and districts/community areas and found that the lower-level units (street segments) explained the largest proportions of variation. Johnson and Bowers (2010) and Davies and Johnson (2015) applied three-level Poisson models with street segments nested in small areas nested in larger areas and observed that street attributes, such as permeability and potential usage, were positively associated with burglary.

The multilevel models used in past studies have analysed purely hierarchical data where lower-level units are nested in one higher-level unit [e.g. street segments are nested only in neighbourhoods (Steenbeek and Weisburd 2016)]. Cross-classified multilevel models, in contrast, accommodate data where lower-level units are nested in two or more non-hierarchical higher-level units (Goldstein 1994). Cross-classified multilevel models facilitate the integration of data collected using different geographical frameworks, such as sociodemographic data available for census areas and civic data available for political boundaries, and allow for the analysis of multiple sets of overlapping higher-level areas that are each thought to have distinct crime-generating processes. Cross-classified models are also advantageous for policy applications as they can estimate the area-specific effects of both lower- and higher-level units and can quantify the degree to which each spatial scale explains the overall spatial and/or spatiotemporal variation of crime.

4 Study region and data

The Region of Waterloo is located in Ontario, Canada, and is composed of the cities of Cambridge, Kitchener, and Waterloo. The lower-level unit of analysis was the dissemination area (DA) and the higher-level units of analysis were the neighbourhood, electoral ward, and police patrol zone. DAs are the smallest census areas that cover the entirety of Canada and are delineated such that average residential populations are between 400 and 800. In the 2016 Canadian census, there were 656 DAs in the study region with an average residential population of 724 and an average size of 0.49 km² (Fig. 1). Crime, sociodemographic, and built environment data were

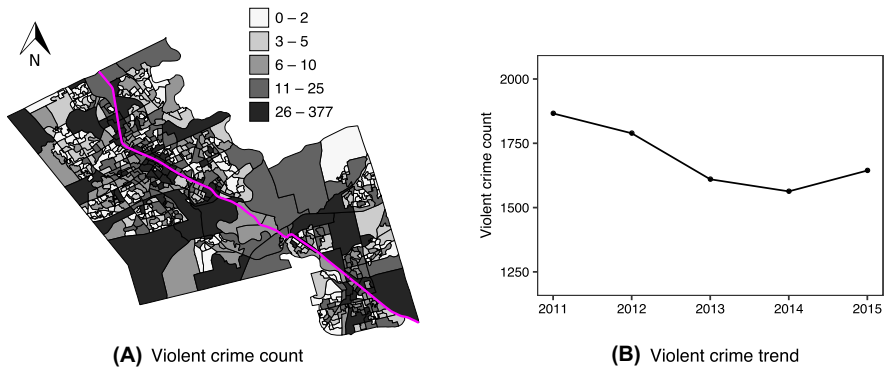


Fig. 1 Total violent crime counts at the dissemination area scale (a) and the five-year violent crime trend (b). The central commercial corridor is highlighted in (a)

analysed at the DA scale, and a covariate measuring civic engagement was included at the electoral ward scale.

4.1 Crime, sociodemographic, and built environment data

Reported violent crime incidents were retrieved from the Waterloo Regional Police Service for five years, from 2011 to 2015. Violent crime was calculated as the sum of homicide, assault, sexual offence, and robbery incidents (Statistics Canada 2015). Reported violent crime incidents were aggregated from street intersections to DAs for analysis. Total violent crime counts at the DA scale are shown in Fig. 1 (see “Appendix A” for descriptive statistics). In general, DAs with high counts of violent crime were clustered around the central commercial corridor as well as in peripheral areas located in the southwest and southeast. Temporally, violent crime decreased by about twelve per cent during the five-year study period, with a decline of approximately sixteen per cent between 2011 and 2014 and an increase of about five per cent from 2014 to 2015 (Fig. 1).

Fifteen sociodemographic and built environment variables were selected for analysis at the DA scale based on past research exploring the relationships between crime and the urban environment. Following past studies using generalized linear (Poisson) regression models for analyzing spatial and spatiotemporal crime counts (Ceccato et al. 2018; Quick et al. 2019), residential population was included as an explanatory variable for two reasons. One is because there was no clear definition of the population at risk as violent crime offenders and targets are mobile over space and time. Second is because assuming that areas with larger residential populations will have higher levels of crime, which is implied when residential population is used to construct crime rates for regression models requiring continuous dependent variables, may not be appropriate when crime is geographically clustered in areas with small residential populations (e.g. in central business districts, in suburban commercial areas, or in industrial areas) (Malleon and Andresen 2015).

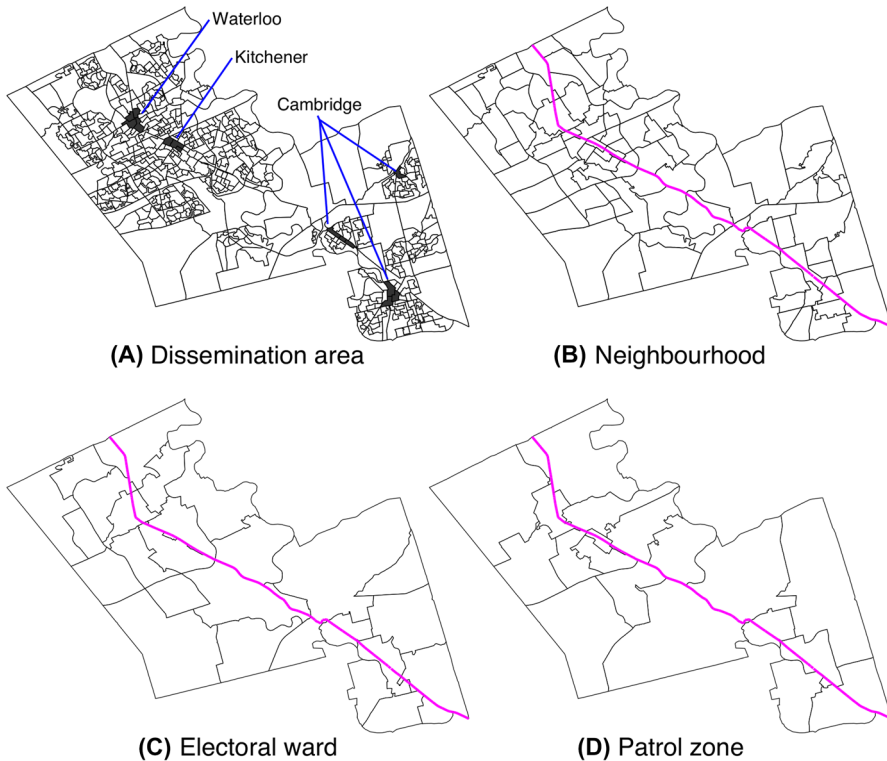


Fig. 2 The geographical boundaries of the lower- and higher-level units. The central business districts in Cambridge, Kitchener, and Waterloo are highlighted in (a). The central commercial corridor is highlighted in (b), (c), and (d)

Four variables were used to represent the built environment: the central business district, commercial land use, government-institutional land use, and recreational land use (Lockwood 2007; Stucky and Ottensmann 2009). The five central business districts in the study region were composed of 14 DAs and were operationalized as a binary variable, where DAs within the central business districts were assigned a value of one and all other DAs were assigned a value of zero (Fig. 2a). Commercial land use, government-institutional, and recreational land use variables were also analyzed as binary variables because these land uses were infrequent relative to the total number of DAs in the study region. Commercial land uses included retail stores and shopping malls, government-institutional land uses included government buildings and community services, and recreational land uses included parks and community centres.

Ten of the fifteen explanatory variables were treated as indicators of four latent factors representing residential instability, socioeconomic disadvantage, family disruption, and ethnic heterogeneity. Residential instability was operationalized as the per cent of renters and the five-year residential mobility rate, and socioeconomic disadvantage was measured via the median after-tax household income,

the per cent of low-income households, the unemployment rate, and the per cent of total income received from government transfer payments (Morenoff et al. 2001; Law and Quick 2013). Family disruption was a latent factor constructed from the per cent of single-parent families and the per cent of divorced or separated households, and ethnic heterogeneity was operationalized via the index of ethnic heterogeneity and the index of language heterogeneity (Sampson and Groves 1989; Veysey and Messner 1999; Hipp 2007). For reference, the indices of ethnic and language heterogeneity quantify the relative mix of ethnicities and languages spoken within DAs, respectively, and have values that range between zero (less heterogeneity) and one (more heterogeneity). Descriptive statistics for the explanatory variables are shown in “Appendix A” and details regarding the factor analytic models are shown in “Appendix B.”

4.2 Neighbourhoods, electoral wards, and patrol zones

Neighbourhoods, electoral wards, and police patrol zones were the higher-level units of analysis (Fig. 2). Each of the three sets of higher-level units covered the entirety of the study region, and so each DA was nested in one neighbourhood, one electoral ward, and one patrol zone. Because the boundaries of the three higher-level units were overlapping, however, the group of DAs within one neighbourhood could belong to multiple electoral wards and/or multiple patrol zones. Most DA boundaries directly aligned with the boundaries of neighbourhoods, electoral wards, and police zones; however, for the few DA boundaries that were misaligned with the higher-level boundaries, DAs were assigned to the neighbourhood, electoral ward, or police zone in which the largest proportion of area was located. In the study region, there were 97 neighbourhoods with an average area of 3.29 km², 25 electoral wards with an average area of 12.76 km², and 18 police zones with an average area of 17.73 km². Neighbourhoods included an average of 6.77 DAs, electoral wards included an average of 26.24 DAs, and police zones included an average of 36.44 DAs.

Each of the higher-level units is relevant for theoretical inference and for policy applications in urban planning (neighbourhoods), local government (electoral wards), and law enforcement (patrol zones). Defined by the three municipal governments and the three urban planning departments in the study region, neighbourhoods are used for a variety of administrative and policy purposes including neighbourhood associations, which are resident-led organizations that coordinate local programs and events, and secondary land use plans, which provide detailed guidelines for urban development, infrastructure, and environmental services. Past research has suggested that municipally defined neighbourhoods are suitable for analyzing public social control because this is often the scale at which the economic and political decisions of private and public actors are realized (e.g. via urban planning policies and investments in housing) (Wooldredge 2002; Sampson 2013). Also, because neighbourhoods are used for local land use policies, the DAs located in a given neighbourhood are likely to have more similar land use compositions and routine activity patterns than nearby DAs located in different neighbourhoods.

Electoral wards are delineated by the local governments in the study region and are used for the elections of regional representatives, city mayors, and city councillors. From a theoretical perspective, electoral wards are appropriate for operationalizing public social control and collective efficacy because they represent the geographical areas through which residents and communities engage with political representatives on local issues (e.g. emergency services, by-law enforcement) and work with government to secure external resources on behalf of the common good (e.g. public space amenities, funding for community programs) (Rosenfeld et al. 2001; Weisburd et al. 2012). From a practical perspective, data measuring the per cent of active voters in the 2014 municipal/regional elections were available for electoral wards and, as such, the effects of civic engagement on violent crime could be analyzed without changing the scale of the data.

Police patrol zones are defined by Waterloo Regional Police Services and are constructed to optimize the delivery of police services. Patrol zones were included in this study to account for potential geographical differences in law enforcement resources, policing tactics, and the relationships between community members and law enforcement (Hagan et al. 1978; Velez 2001). For example, it is possible that a higher proportion of crime events are reported to, or observed by, police in areas with a more visible police presence, such as in downtown areas where patrols are more frequently done on foot (Klinger and Bridges 1997). Furthermore, there may be differences in reported violent crimes that parallel the variations in resident-based police confidence and/or legal cynicism that are due to the interactions between police and community members within a patrol zone (Goudriaan et al. 2016; Kirk and Matsuda 2011).

5 Multilevel modelling of spatiotemporal crime patterns

Let O_{it} represent the observed violent crime counts for DA ($i = 1, \dots, 656$) and year ($t = 1, \dots, 5$). Each DA is nested in one neighbourhood ($j_1 = 1, \dots, 97$), one electoral ward ($j_2 = 1, \dots, 25$), and one police patrol zone ($j_3 = 1, \dots, 18$), and so the observed spatiotemporal violent crime counts in all lower- and higher-level units are denoted by $O_{it(j_1j_2j_3)}$. The parentheses surrounding j_1, j_2 , and j_3 indicate that neighbourhoods, electoral wards, and patrol zones are non-hierarchical and analyzed at the same level (Rasbash and Goldstein 1994). $O_{it(j_1j_2j_3)}$ were modelled as independent Poisson random variables conditional on mean $\mu_{it(j_1j_2j_3)}$. The Poisson model is often used in Bayesian spatial and spatiotemporal modelling of small area count data (Waller et al. 1997; Congdon 2003; Haining et al. 2009). The models used to analyze the multiscale patterns of violent crime are described below.

Model 1 is a single-level model that analyzes the spatiotemporal variation of crime across DAs (i.e. no terms indexed by j_1, j_2 , or j_3). In Model 1, the expected crime counts for each DA (μ_{it}) were modelled as the sum of an intercept (α), a set of covariates for the observed explanatory variables and latent constructs (βx_i 's and $\kappa \psi_i$'s), a set of spatially unstructured random effects terms (u_i), a set of temporal terms (ζ_t), and a set of space–time random effects terms (θ_{it}). The regression coefficients denoted by β_p ($p = 1, \dots, 5$) quantify the relationships between crime and

residential population, the central business district, commercial land use, government-institutional land use, and recreational land use. The regression coefficients denoted by κ_n ($n=1, \dots, 4$) quantify the relationships between crime and the four latent constructs representing residential instability, socioeconomic disadvantage, family disruption, and ethnic heterogeneity. The spatially unstructured random effects terms capture overdispersion of violent crime counts and the residual within-DA variability of crime, the temporal terms capture the overall crime trend for the study region, and the space–time random effects terms capture extra-Poisson heterogeneity that is not modelled via the purely spatial and purely temporal model parameters.

$$\log(\mu_{it}) = \alpha + \beta_p x_{pi} + \kappa_n \Psi_{ni} + u_i + \zeta_t + \theta_{it} \quad (1)$$

In Model 2, the multilevel structure of DAs nested in neighbourhoods, electoral wards, and patrol zones is modelled through the addition of three sets of higher-level random effects terms and one higher-level covariate. The higher-level covariate quantifies the relationship between violent crime and the per cent of voters at the electoral ward scale ($\lambda\omega_{(j_2)}$) and the random effects terms capture the variation of violent crime that is attributed to DAs being grouped in neighbourhoods ($\gamma_{1(j_1)}$), electoral wards ($\gamma_{2(j_2)}$), and patrol zones ($\gamma_{3(j_3)}$) (Langford et al. 1999). For interpretation, the intercept (α) captures the overall mean of violent crime across all lower- and higher-level units and the higher-level random effects terms ($\gamma_{1(j_1)}$, $\gamma_{2(j_2)}$ and $\gamma_{3(j_3)}$) capture the differences between the overall mean and the neighbourhood-, electoral ward-, and police zone-specific means of violent crime after accounting for the explanatory variables (Leckie 2013). For example, neighbourhoods with large values of γ_1 will tend to be composed of DAs that have relatively high violent crime, whereas police patrol zones with small values of γ_3 will tend to be composed of DAs that have relatively low violent crime. Note that Model 1 and Model 2 were tested with a set of spatially structured random effects (assigned intrinsic conditional autoregressive prior distributions with a first-order queen contiguity matrix) to capture residual spatial autocorrelation between dissemination areas; however, the spatially structured random effects terms did not converge, did not improve model fit, and were not included in the final models (Besag et al. 1991; Arcaya et al. 2012; Dong and Harris 2015).

$$\log(\mu_{it(j_1j_2j_3)}) = \alpha + \beta_p x_{pi} + \kappa_n \Psi_{ni} + u_i + \zeta_t + \theta_{it} + (\gamma_{1(j_1)}) + (\lambda\omega_{(j_2)}) + (\gamma_{2(j_2)}) + (\gamma_{3(j_3)}) \quad (2)$$

5.1 Variance partition coefficients

Variance partition coefficients (VPCs) quantify the degree to which the residual spatiotemporal variation of violent crime is explained by each set of random effects terms (Goldstein et al. 2002). The VPC calculating the proportion of variation explained by the lower-level random effects terms, for example, is equal to the sum of the empirical variances of μ_i and θ_{it} divided by the sum of the empirical variances

of μ_i , ζ_t , θ_{it} , $(\gamma_{1(j_1)})$, $(\gamma_{2(j_2)})$, and $(\gamma_{3(j_3)})$. Similarly, the VPC calculating the proportion of variation explained by the higher-level units is equal to the sum of the empirical variances of $(\gamma_{1(j_1)})$, $(\gamma_{2(j_2)})$, and $(\gamma_{3(j_3)})$ divided by the sum of the empirical variances of μ_i , ζ_t , θ_{it} , $(\gamma_{1(j_1)})$, $(\gamma_{2(j_2)})$, and $(\gamma_{3(j_3)})$ (Browne et al. 2001). For policy, it is also relevant to quantify the proportion of variation that is purely spatial (or stable over time), purely temporal (or stable over the study region), and spatiotemporal (or varies both in space and time). For example, if most of the variation of crime is spatial, then crime prevention initiatives may look to modify permanent geographical risk factors, but if the variation of crime is spatiotemporal, then policies and programs may look to target specific small areas with increasing violent crime (Johnson et al. 2008; Quick et al. 2017). The VPC calculating the proportion of variation that is purely spatial, for example, is equal to the sum of the empirical variances of μ_i , $(\gamma_{1(j_1)})$, $(\gamma_{2(j_2)})$, and $(\gamma_{3(j_3)})$ divided by the sum of the empirical variances of μ_i , ζ_t , θ_{it} , $(\gamma_{1(j_1)})$, $(\gamma_{2(j_2)})$, and $(\gamma_{3(j_3)})$.

5.2 Prior distributions

In Bayesian modelling, all parameters are treated as random variables and are assigned prior probability distributions. The intercept (α) was assigned an improper uniform prior distribution, and the regression coefficients (β 's, κ 's, and λ) were assigned vague normal prior distributions with means of zero and variances of 1000. The set of spatially unstructured random effects for DAs were assigned normal prior distributions with means of zero and a common unknown variance σ_u^2 .

The random effects terms for neighbourhoods $(\gamma_{1(j_1)})$, electoral wards $(\gamma_{2(j_2)})$, and police zones $(\gamma_{3(j_3)})$ were each assigned normal distributions with means of zero and common unknown variances $\sigma_{\gamma_1}^2$, $\sigma_{\gamma_2}^2$, and $\sigma_{\gamma_3}^2$, respectively (Browne et al. 2001). Because the variance parameters for the higher-level random effects parameters were estimated from the data, these prior distributions do not assume that neighbourhoods, electoral wards, or police zones were relatively more or less important for explaining the spatiotemporal patterning of violent crime. These prior distributions also assume that there was no spatial autocorrelation between the higher-level units. Preliminary models that included spatially structured random effects terms for each of the higher-level units (assigned intrinsic conditional autoregressive prior distributions as per Besag et al. (1991)) were tested, but these parameters did not converge, did not improve model fit, and were not included in the final model.

The temporal terms (ζ_t) were assigned a normal distribution with means of $(b_0 \cdot t^*)$ and an unknown variance σ_ζ^2 , where b_0 is a regression coefficient and t^* is the mean-centred time ($t^* = t - 3$) (Li et al. 2014). The regression coefficient b_0 was assigned a vague normal prior distribution. This parameterization estimates a linear violent crime trend over the five years via $b_0 \cdot t^*$ and allows for the overall time trend (ζ_t) to depart from the linear trend for each time period via the Gaussian noise modelled by σ_ζ^2 . The space–time random effects terms (θ_{it}) were assigned normal distri-

butions with means of zero and a common unknown variance σ_θ^2 . This prior distribution assumes that the residual space–time variability of crime was not correlated between small areas or between years.

To complete the Bayesian hierarchical model, prior distributions were assigned to the variance parameters of the random effects terms. The standard deviation of each set of random effects terms (σ_u , σ_{γ_1} , σ_{γ_2} , σ_{γ_3} , σ_ζ , and σ_θ) was assigned a positive half-Gaussian prior distribution, $Normal_{+\infty}(0, 10)$ (Gelman 2006). To test for the sensitivity of model results to the prior distributions of the random effects parameters, two alternative prior distributions were specified for the precisions of all random effects parameters, $Inverse\ Gamma(0.001, 0.001)$ and $Inverse\ Gamma(0.5, 0.0005)$ (Kelsall and Wakefield 1999; Browne et al. 2001). The results of these sensitivity analyses were nearly identical to the results shown here.

5.3 Model fitting

Model 1 and Model 2 were fitted using the Markov chain Monte Carlo algorithm (MCMC) in WinBUGS v.1.4.3. The WinBUGS code for Model 2 is shown in “Appendix D.” Two MCMC chains were initiated at dispersed initial values, and the first 200,000 iterations (for each chain) were discarded as burn-in. Convergence of model parameters was monitored via visual inspection of trace plots and via Brooks–Gelman–Rubin diagnostics. For inference, an additional 200,000 iterations were sampled for each MCMC chain, retaining every twentieth iteration to reduce autocorrelation of the posterior samples. The Monte Carlo errors for all model parameters were less than five per cent of the corresponding posterior standard deviations, indicating that the total number of iterations were sufficient to approximate the posterior distributions (Lunn et al. 2012). The Deviance Information Criterion (DIC) was used to compare Model 1 and Model 2. The DIC balances goodness of fit and model complexity, where goodness of fit is assessed via the posterior mean deviance and model complexity is assessed via the effective number of parameters (Spiegelhalter et al. 2002). The model with the smallest DIC value is considered to be the best fitting model.

6 Results

The DIC values for Model 1 and Model 2 are shown in Table 1. Model 2 had a smaller DIC value than Model 1. This provides evidence that model fit improved when accounting for the clustering of violent crime within neighbourhoods, electoral wards, and patrol zones. The posterior medians and 95% credible intervals (95% CI) of the intercept, the regression coefficients, and the variance parameters of the random effects terms from Model 1 and Model 2 are also shown in Table 1. The 95% CI is the interval that contains the true value of a parameter with 95% probability. The regression coefficients are interpreted as relative risks (i.e. exponential

Table 1 Model fit criterion and posterior medians and 95% credible intervals (in parentheses) for Model 1 and Model 2

	Model 1	Model 2
Deviance Information Criterion (DIC)	9950	9944
Intercept ($\exp(\alpha)$)	0.87 (0.77, 0.99)	0.92 (0.77, 1.15)
Population	1.28 (1.19, 1.36)	1.24 (1.15, 1.33)
Central business district	2.76 (1.79, 4.26)	2.75 (1.78, 4.29)
Commercial land use	1.51 (1.11, 2.08)	1.59 (1.16, 2.14)
Government-institutional land use	1.62 (1.27, 1.82)	1.46 (1.22, 1.75)
Recreational land use	1.22 (1.06, 1.39)	1.26 (1.10, 1.45)
Residential instability	1.14 (0.98, 1.32)	1.16 (1.00, 1.32)
Socioeconomic disadvantage	1.34 (1.01, 1.73)	1.23 (0.93, 1.62)
Family disruption	1.08 (0.86, 1.37)	1.08 (0.85, 1.37)
Ethnic heterogeneity	1.05 (0.96, 1.15)	1.07 (0.96, 1.18)
Per cent of active voters	NA	0.86 (0.69, 0.97)
Empirical variances of lower-level random effects		
Spatial (u_i)	0.97 (0.88, 1.08)	0.85 (0.75, 0.95)
Space-time (θ_{it})	0.06 (0.04, 0.09)	0.06 (0.04, 0.09)
Empirical variance of higher-level random effects		
Neighbourhood (γ_1)	NA	0.09 (0.03, 0.18)
Electoral ward (γ_2)	NA	0.02 (0.00, 0.08)
Patrol zones (γ_3)	NA	0.04 (0.00, 0.12)
Empirical variance of temporal effects (ζ_t)	0.01 (0.00, 0.02)	0.01 (0.00, 0.02)

transformations of β 's, κ 's, and λ), where coefficient values greater than one indicate positive associations between the explanatory variables and violent crime.¹

Violent crime was found to be positively associated with population size, the central business district, residential instability, and commercial, government-institutional, and recreational land uses within DAs. Broadly, these results support past research exploring the relationships between the urban environment and local crime patterns for a single spatial scale. From a social disorganization perspective, large population sizes are thought to increase the level of anonymity and distrust among residents and high levels of residential instability have been shown to limit the formation of social ties and reduce resident-based informal social control (Sampson and Groves 1989; Rosenfeld et al. 2001). From a routine activity perspective, areas with large populations and with commercial, government-institutional, and recreational land uses likely have higher number of potential targets and offenders, and, consequently, more frequent opportunities for violent crime offences. The central business district had the largest positive association with violent crime of all the covariates. In

¹ Following Congdon (2011), the regression coefficients were standardized in order to compare the relative effects of the observed explanatory variables and the latent factors (see Appendix B). Table 1 reports the posterior medians and uncertainty intervals from the standardized regression coefficients.

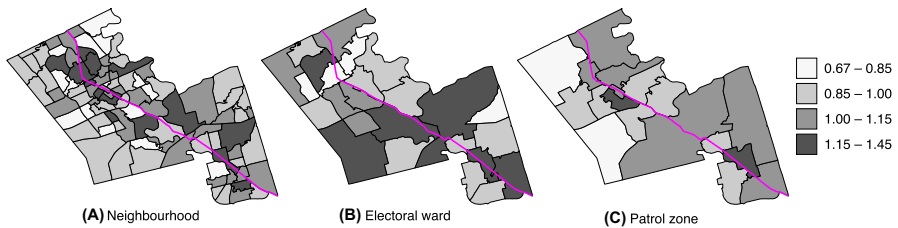


Fig. 3 Posterior medians of the higher-level neighbourhood ($\exp(\gamma_1)$) (a), electoral ward ($\exp(\lambda\omega_2 + \gamma_2)$) (b), and police patrol zone ($\exp(\gamma_3)$) terms (c)

this study region, the central business districts have high densities of business-centred non-residential land uses and attract large numbers of residents and non-residents during routine activities. Combined, these attributes have been found to limit social interaction among residents and challenge the formation informal social control as well as facilitate frequent convergences between targets and offenders (Taylor 1997).

Violent crime was also negatively associated with the per cent of active voters at the electoral ward scale. In particular, the per cent of active voters was found to have a contextual effect on violent crime such that small areas located within higher voting electoral wards had lower violent crime, on average, than small areas located within lower voting electoral wards after accounting for the lower-level sociodemographic and built environment variables. This advances past studies that focus on the single-level relationships between civic engagement and crime by directly analyzing the effect of per cent of active voters at the scale used by residents and communities to elect local representatives and work with government to secure place-based resources (Weisburd et al. 2012; Rosenfeld et al. 2001). The relationship between the per cent of active voters and violent crime is likely indirect insofar as it is an indicator of underlying public social control and collective efficacy that manifests through the relationships between, and actions taken by, residents living in electoral wards.

Focusing on the higher-level units, Fig. 3 maps the contextual effects of neighbourhoods, electoral wards, and patrol zones. Across the study region, neighbourhoods located close to the central commercial corridor generally had a positive influence on violent crime ($\exp(\gamma_1) > 1$) and neighbourhoods around the periphery generally had a negative influence ($\exp(\gamma_1) < 1$). Like neighbourhoods, electoral wards that had a negative influence on crime were typically located around the periphery of the study region, however, there was no clear grouping of high or low effects among electoral wards. For patrol zones, there were two areas located close to the central business districts in Kitchener and Cambridge with large positive effect on violent crime, however, the uncertainty intervals around the posterior medians of most patrol zone random effects terms included zero and so this spatial scale had no meaningful effect on violent crime. The visual heterogeneity of the higher-level effects in Fig. 3 reflects the ranking of the total empirical variances of the higher-level terms: neighbourhoods had the largest variance (0.09 with 95% CI 0.03–0.18), electoral wards had the second largest variance (0.06 with 95% CI 0.01–0.15), and patrol zones had the smallest variance (0.04 with 95% CI 0.00–0.12). Note that most

of the variance for electoral wards was due to the covariate for the per cent of active voters (0.04 with 95% CI 0.00–0.11) and that a relatively smaller proportion was due to the random effects terms (0.02 with 95% CI 0.00–0.08).

Examining the VPCs for the random effects parameters in Model 1, the lower-level spatially unstructured random effects terms explained approximately 93% (95% CI 91–95%) of the residual spatiotemporal variability of violent crime. Accounting for neighbourhoods, electoral wards, and patrol zones in Model 2, this decreased to approximately 79% (95% CI 69–86%). Combined, the three higher-level units explained about 15% (95% CI 8–24%) of the total residual variability of violent crime, of which, neighbourhoods accounted for the largest proportion (8% with 95% CI 3–16%). Patrol zone random effects terms captured about 4% (95% CI 0–10%), and there was effectively no variance explained by the electoral ward random effects terms (2% with 95% CI 0–7%). The VPCs for the space–time random effects (6% with 95% CI 4–8%) and the time trend (1% with 95% CI 0–2%) were consistent in both models. This broadly aligns with past multilevel analyses observing that the smallest geographical unit explains the greatest spatial variability of crime (Steenbeek and Weisburd 2016; Schnell et al. 2017). The posterior medians and uncertainty intervals for the VPCs are shown in “Appendix C.”

7 Discussion

This research has applied a Bayesian cross-classified multilevel modelling approach to analyze the multiscale spatiotemporal patterning of violent crime. Violent crime was measured at the small area scale (DAs), and small areas were nested in neighbourhoods, electoral wards, and police patrol zones. Violent crime was found to be positively associated with population size and built environment and sociodemographic characteristics within DAs and negatively associated with the per cent of active voters, an indicator of civic engagement, within electoral wards. Combined, the higher-level units explained approximately fourteen per cent of the spatiotemporal variation of violent crime, of which, neighbourhoods were the most important source of variability. The cross-classified model used in this study accommodates multiple non-hierarchical higher-level units that each influence where and when crime events occur. The variation jointly attributed to the three overlapping higher-level units would otherwise be overlooked in single-level analyses or misattributed to one of the scales included in multilevel analyses of purely hierarchical data (Moerbeek 2004; Leckie 2013).

Examining spatiotemporal crime patterns via cross-classified multilevel models advances theoretical understanding of the multiscale processes influencing crime and provides area-specific information for crime prevention policy. Focusing first on the theoretical contributions, this study found that DAs were the most important set of geographical units for explaining where and when violent crime occurred. This suggests that the local contexts surrounding crime events have a greater impact than broader social, political, and built environment contexts. Generally, the lower-level regression coefficients that were found to be associated with violent crime follows past research positing that small-area spatial units are more suitable than large areas

for capturing variation in routine activity patterns, or the convergences between offenders and targets around specific activity nodes and activity pathways (Brantingham and Brantingham 1993), as well as the variations of parochial social control, or the type of social control that arises from the non-intimate relationships among community members (Wooldredge 2002).

Among the three higher-level units, neighbourhoods were the largest source of variation after accounting for social disorganization, routine activity, and collective efficacy covariates. The importance of neighbourhoods for explaining violent crime patterns can be interpreted from analytical and theoretical perspectives. Analytically, neighbourhoods were composed of fewer DAs than electoral wards and patrol zones, and, as such, the DAs in a given neighbourhood were more likely to have similar levels of violent crime than the DAs nested in a given electoral ward or patrol zone. Theoretically, neighbourhoods in the study region are created by local planning departments and are meaningful spatial units for capturing similarities in routine activity patterns and variations in public social control. For example, many of the peripheral neighbourhoods that were found to have a negative effect on violent crime are suburban and characterized by residential and open space land uses (Fig. 3). These neighbourhoods are less likely to attract potential offenders and have frequent convergences between offenders and victims compared to neighbourhoods with many activity nodes, such as those that were found to have a positive effect on violent crime located close to the central commercial corridor (Greenberg et al. 1982). Likewise, because neighbourhoods generally align with the boundaries of neighbourhood associations and are used by local governments to allocate funding and resources for amenities and infrastructure, neighbourhoods are spatial units suitable for representing the relationships between community members as well as the relationships between neighbourhoods and extra-local organizations. Accordingly, a second explanation for the variation in neighbourhood-scale effects is that this pattern reflects the differing relationships between neighbourhoods and government and specifically the degree to which neighbourhoods mobilize to secure common good resources and influence the political processes that shape neighbourhood environments (Velez 2001; Kubrin and Weitzer 2003)

Electoral wards and patrol zones were found to explain the smallest proportions of the variation of violent crime among all four spatial scales (see “Appendix C”). That is, after accounting for the relationship between high civic engagement and low violent crime within electoral wards, the results of this study show that these two higher-level areas were generally composed of DAs with a mix of high and low violent crime and that any additional crime-generating processes operating within electoral wards and patrol zones did not have a substantial influence on crime. Despite the small proportion of variation explained by police patrol zones, the processes associated with crime within patrol zones are challenging to interpret and verify because they were captured via a set of random effects terms rather than observed covariates. In future research, the patterning of crime at the patrol zone scale should be further analyzed using spatiotemporal data measuring police activities and methodological approaches that quantify the impacts of these activities on crime (Li et al. 2013). For example, evaluating how multiscale crime patterns change in response to the initiation of a crime prevention intervention would help to understand if police

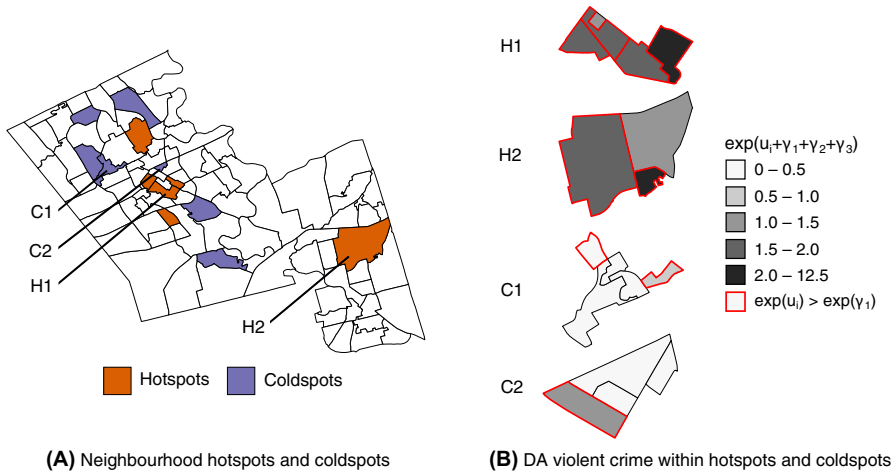


Fig. 4 Neighbourhood hot spots and coldspots (a) and total dissemination area violent crime within two hot spots and two coldspots (b). Dissemination areas with red boundaries had DA effects (posterior median of $\exp(u_i)$) that were greater than neighbourhood effects (posterior median of $\exp(\gamma_i)$). The insets of H1, H2, C1, and C2 are not to scale

activity acts as a deterrent and leads to decreasing numbers of crime incidents or if police activity strengthens the relationships between communities and law enforcement and leads to increasing numbers of reported incidents.

7.1 Policy applications of multiscale crime patterns

Analyzing the spatiotemporal patterning of crime via a cross-classified multilevel modelling approach also provides information regarding the location, the scale, and the spatial/spatiotemporal focus of crime prevention policies. Based on the regression coefficients and the VPCs, this study indicates that policies and programs for violent crime should be designed areas located in and around the central business districts that are characterized by large populations, activity node land uses, high levels of residential instability, and low civic engagement at the DA and neighbourhood scales. Moreover, because the geographical variation of crime between areas was relatively consistent over time, crime prevention initiatives should focus on the underlying spatial pattern of violent crime rather than spatiotemporal variations within lower- or higher-level units. For reference, the spatial random effects terms for all four scales accounted for 93% (95% CI 91–95%) of the total spatiotemporal variation of violent crime in Model 2 (see “Appendix C”).

Focusing specifically on the interactions between DAs and neighbourhoods, Fig. 4a identifies neighbourhood-scale violent crime hot spots and coldspots and Fig. 4b details the DA-specific violent crime in two neighbourhood hot spots and two coldspots. Hot spots and coldspots had a strong positive or negative contextual effect on violent crime and were identified using the posterior probability of

the neighbourhood random effects terms; $\Pr(\exp(\gamma_1) > 1 \mid \text{data}) > 0.8$ for hot spots and $\Pr(\exp(\gamma_1) < 1 \mid \text{data}) > 0.8$ for coldspots (Richardson et al. 2004). Broadly, Fig. 4 shows that a similar number of neighbourhoods had strong positive and negative contextual effects (six hot spots compared to seven coldspots), that neighbourhood hot spots and coldspots were dispersed throughout the study region, and that DA-specific violent crime varied considerably within the hot spot and coldspot neighbourhoods.

For neighbourhood hot spots, urban planners may look to increase public social control by facilitating relationships between residents and local government, and by providing resources that can be used to address community-oriented projects. Urban planners are also positioned to modify the built environment of both neighbourhood hot spots and high-crime DAs within neighbourhoods via land use zoning or urban design initiatives that increase perceptions of guardianship and reduce the likelihood that crime opportunities result in a crime offence (Johnson et al. 2008). Both of these strategies are primarily spatial insofar as public social control and the built environment are relatively stable over time. For DAs located within neighbourhood hot spots, and particularly for those with violent crime that exceeds the neighbourhood average, police may look to implement geographically focused initiatives such as hotspot policing (Braga et al. 2014). While policing interventions likely occur over a relatively short period of time, past studies have shown that crime reduction benefits can diffuse from targeted areas to nearby areas and influence the stable spatial pattern of crime as well as temporal fluctuations (Guerette and Bowers 2009).

7.2 Limitations and future research

One limitation of this research is that reported incident data was used to measure violent crime. While this type of data is common in past spatiotemporal analyses, and while the data used in this study includes only incidents with a filed police report, it is possible that victim or geographical characteristics influence the degree to which crime is reported to police and/or that reported crime types and locations were misclassified (Baumer 2002; Klinger and Bridges 1997). A second limitation is that crime data was aggregated from intersection points to small areas in order to have a common set of analytical units for crime, sociodemographic, and land use data. Future studies should look to integrate point and areal data using methods specifically designed for the change-of-support problem (Gelfand et al. 2001) and investigate how different data aggregation processes influence the results of analyses.

A third limitation of this study is that covariates were not available and were not included at the neighbourhood and patrol zone scales. One reason for the lack of neighbourhood data may be that this set of spatial units are under the jurisdiction of three separate municipalities and each municipality has a distinct data collection and dissemination approach. Future work should look to add higher-level covariates that operationalize public social control, collective efficacy, and the distribution of police resources to directly quantify the multiscale processes influencing crime. Note that the modelling approach used in this study can be extended to allow higher-level covariates to vary over space and/or over time, which would allow researchers to specifically examine

how spatiotemporal changes in the distribution of policing resources within patrol zones influences crime at both lower and higher-level units, for example (Wheeler and Waller 2008; Congdon 2003). A fourth limitation of this study is the modifiable areal unit problem, which highlights the influence of spatial scale and zonal boundaries on the analysis of aggregate data (Openshaw 1984). As such, this research should be taken in the context of the four spatial scales analyzed and alternative operationalizations of lower- and higher-level units should be investigated (Ratcliffe and McCullagh 1999).

Future research should also explore how the multiscale patterns of multiple crime types are similar and different. Past research has suggested that many crime types are correlated and may be associated with the same underlying spatial risk factors, yet no study has examined how multiple crime types are correlated across multiple scales (Quick et al. 2018; Yin et al. 2014). The cross-classified multilevel model used in this research can be extended to accommodate multiple outcomes and quantify the degree to which crime-general and crime-specific patterns are explained by each spatial scale. When reliable point-level crime data is available, future work should apply multilevel models to examine how specific crime, offender, or victim locations are influenced by small-area characteristics as well as multiple overlapping larger areas, while accounting for the correlation structures between points (i.e. via point process models) as well as higher-level areas (Rogerson and Sun 2001; Diggle et al. 2013). Finally, multilevel cross-classified models should be applied to evaluate the impacts of crime prevention initiatives because they can incorporate the multiple overlapping areas used for hotspot policing and/or planning interventions, quantify the impact of policy implementation over time, and account for covariates at multiple spatial scales.

Acknowledgements This work was supported by the Social Sciences and Humanities Research Council of Canada Grant Number [767-2013-1540]. All analyses and interpretation of this data are entirely that of the author.

Appendix A: Descriptive statistics for violent crime and explanatory variables

See Tables 2, 3.

Table 2 Total and annual violent crime counts at the dissemination area scale

	Total count	Mean	Standard deviation	Min	Max
Total	8472	12.92	24.40	0	377
2011	1866	2.85	5.18	0	59
2012	1789	2.72	5.17	0	75
2013	1610	2.45	4.99	0	77
2014	1563	2.38	5.18	0	81
2015	1644	2.51	5.37	0	85

Table 3 Descriptive statistics explanatory variables and the results of factor loadings. Posterior medians and 95% credible intervals of the factor loadings are shown

	Mean	Standard deviation	Factor loading
Residential population	712.69	609.03	NA
Per cent of voters (%)	26.49	6.48	NA
Built environment			
Central business district	0.02	NA ^a	NA
Commercial land use	0.04	NA	NA
Government-institutional land use	0.15	NA	NA
Recreational land use	0.38	NA	NA
Residential instability			
Five-year residential mobility (%)	37.81	15.13	1.00
Renters (%)	30.72	27.11	0.74 (0.68, 0.79)
Socioeconomic disadvantage			
Median after-tax household income (\$)	69,819.31	23,528.13	-1.00
Low-income individuals (%)	8.05	8.20	0.67 (0.61, 0.74)
Unemployment rate (%)	6.83	3.84	0.38 (0.31, 0.46)
Government transfer payments (%)	12.44	5.59	0.78 (0.72, 0.84)
Family disruption			
Single-parent families (%)	17.83	8.72	1.00 (0.91, 1.10)
Divorced or separated households (%)	9.33	4.43	1.00
Ethnic heterogeneity			
Index of ethnic heterogeneity (0 to 1)	0.34	0.15	1.03 (0.82, 1.30)
Index of language heterogeneity (0 to 1)	0.37	0.14	1.00

^aStandard deviation not reported for binary variables

Appendix B: Bayesian factor analysis models

Let Y_{ik} denote the standardized rates of the ten explanatory variables associated with one of the four latent constructs, where i indexes small areas ($i = 1, \dots, 656$) and k indexes the variable type ($k = 1, \dots, 10$). Y_{ik} were assumed to follow a normal distribution with mean η_{ik} and with an unknown variance for each variable $\sigma_{\eta k}^2$; $Y_{ik} \sim \text{Normal}(\eta_{ik}, \sigma_{\eta k}^2)$ (Congdon 2011). Model A1 estimates the four latent constructs representing residential instability, socioeconomic disadvantage, family disruption, and ethnic heterogeneity ($n = 1, \dots, 4$). Each of the ten census variables was modelled as the sum of a type-specific intercept ($\alpha_{\eta k}$) and a factor component ($\lambda_{nk} \cdot \psi_{ni}$), where λ_{nk} are the factor loadings and ψ_{ni} are the area-specific estimates of the four latent constructs. Following past research, each variable was a priori assigned to the latent constructs (Sampson et al. 1997; Sutherland et al. 2013). For example, $\lambda_{1,1}$ and $\lambda_{1,2}$ represent the factor loadings for the per cent of renters and the five-year mobility rate, respectively, and these variables were associated with the latent construct measuring residential instability, ψ_{1i} (see “Appendix A” and Sect. 4.1 for census variables and associated constructs).

$$\eta_{ik} = \alpha_{\eta k} + (\lambda_{nk} \cdot \psi_{ni}) \quad (3)$$

The standard deviations of type-specific variance parameters ($\sigma_{\eta k}$) were assigned positive half-normal prior distributions with means of zero and variances of 1000 (Gelman 2006). The type-specific intercepts ($\alpha_{\eta k}$) were assigned improper uniform prior distributions. To ensure model identifiability, one factor loading from each construct was set to positive (or negative) one, specifically for per cent of renters ($\lambda_{1,1} = 1$), median household income ($\lambda_{2,1} = -1$), separated/divorced families ($\lambda_{3,1} = 1$), and the index of language heterogeneity ($\lambda_{4,1} = 1$). The remaining factor loadings were assigned positive half-Gaussian prior distributions with means of zero and variances of 1000 (Congdon 2011). The four sets of random effects terms representing the latent variables were assigned a multivariate normal distribution with means set to zero and with a four-by-four variance–covariance matrix Σ . The multivariate normal distribution allows for correlation between the latent variables, where the diagonal elements of Σ are the conditional variances of the four sets of random effects and the off-diagonal elements are the covariances between the four constructs. The inverse of Σ was assigned a Wishart prior distribution with five degrees of freedom and diagonal and off-diagonal elements assigned values of 0.02 and 0, respectively (Thomas et al. 2004). Note that it is possible to impose spatial structure on the latent variables via a multivariate conditional autoregressive prior, as illustrated by Congdon (2008, 2011). Regression coefficients from analyses using spatially structured latent variables were virtually identical to the results shown in Table 1.

Following Congdon (2011), the equations to standardize the regression coefficients for the observed explanatory variables and latent explanatory variables are shown in Models A2, A3, A4, and A5 where $\beta_1^{(s)}$ is the standardized regression coefficient for population size; $\beta_{2:5}^{(s)}$ are the standardized regression coefficients for the binary variables (central business district, commercial land use, government-institutional land use, and recreational land use); $\kappa_n^{(s)}$'s are the standardized regression coefficients for the latent variables (residential instability, socioeconomic disadvantage, family disruption, and ethnic heterogeneity), and $\lambda^{(s)}$ is the standardized regression coefficient for the per cent of active voters. The standard deviations of the observed and latent explanatory variables are denoted by σ_{x_j} and σ_{ψ_n} , respectively, and ϕ is the square root of the variance of the modelled violent crime counts ($\phi = \text{var}(\log(\mu_{it}))^{0.5}$).

$$\beta_1^{(s)} = (\beta_1 \cdot \sigma_{x_1}) / \phi \quad (4)$$

$$\beta_{2:5}^{(s)} = \beta_2 / \phi \quad (5)$$

$$\kappa_n^{(s)} = (\kappa_n \cdot \sigma_{\psi_n}) / \phi \quad (6)$$

$$\lambda^{(s)} = (\lambda \cdot \sigma_{\omega}) / \phi \quad (7)$$

Appendix C: Posterior medians and 95% credible intervals (in parentheses) of the variance partition coefficients for the random effects terms in Model 1 and Model 2

	Model 1	Model 2
Lower-level random effects terms	0.99 (0.98, 1.00)	0.85 (0.75, 0.92)
Spatial	0.93 (0.91, 0.95)	0.79 (0.69, 0.86)
Space–time	0.06 (0.04, 0.08)	0.06 (0.04, 0.08)
Higher-level random effects terms	NA	0.15 (0.08, 0.24)
Neighbourhood	NA	0.08 (0.03, 0.16)
Electoral ward	NA	0.02 (0.00, 0.07)
Patrol zone	NA	0.04 (0.00, 0.10)
Temporal effects	0.01 (0.00, 0.02)	0.01 (0.00, 0.02)

Appendix D: WinBUGS code for Model 2

```

model{
for (i in 1:Narea_16) {
for (t in 1:T) {
O[i,t] ~ dpois(mu[i,t])
log(mu[i,t]) <- u[i] + zeta_model[t] + (beta[1] * pop[i]) + (beta[2] * cbd[i]) + (beta[3] * commercial[i])
+ (beta[4] * gov_ins[i]) + (beta[5] * recreation[i]) + (kappa[1] * instab[i]) + (kappa[2] * disadv[i]) +
(kappa[3] * family[i]) + (kappa[4] * heterog[i]) + theta[i,t] + gamma1[HOOD_ID[i]] +
gamma2[WARD_ID[i]] + gamma3[ZONE_ID[i]] + (lambda * voters[WARD_ID[i]])
theta[i,t] ~ dnorm(0, tau.r) # lower-level space-time random effects
}
}
for (p in 1:Nhood) { # neighbourhoods
gamma1[p] ~ dnorm(0, tau.hood)
hot_gamma1[p] <- step(exp(gamma1[p]) - 1) # neighbourhood hotspot
cold_gamma1[p] <- step(-(exp(gamma1[p]) - 1)) # neighbourhood coldspot
}
tau.hood <- 1 / pow(sd.hood, 2); sd.hood ~ dnorm(0, 10)|(0, )
for (p in 1:Nward) { # electoral wards
gamma2[p] ~ dnorm(0, tau.ward)
}
tau.ward <- 1 / pow(sd.ward, 2); sd.ward ~ dnorm(0, 10)|(0, )
for (p in 1:Nzone) { # police zones
gamma2[p] ~ dnorm(0, tau.zone)
}
tau.zone <- 1 / pow(sd.zone, 2); sd.zone ~ dnorm(0, 10)|(0, )
for (i in 1:Narea_16) { # lower-level (centered parameterization)
u[i] ~ dnorm(alpha, tau.u)
uns[i] <- u[i] - alpha
}
alpha ~ dflat()
(t in 1:T) { # temporal effects
time[t] <- t
zeta[t] ~ dnorm(mu.zeta[t], tau.t)
mu.zeta[t] <- bknot * (time[t] - meanzeta)
zeta_model[t] <- zeta[t] - meanzeta
}
bknot ~ dnorm(0, 0.001) # linear trend for temporal effects
meanzeta <- mean(zeta[1:T]);
meanzeta <- mean(zeta[1:T])
tau.u <- 1 / pow(sd.u, 2); tau.t <- 1 / pow(sd.t, 2); tau.r <- 1 / pow(sd.r, 2) # precision for random effects
sd.u ~ dnorm(0, 10)|(0, ); sd.t ~ dnorm(0, 10)|(0, ); sd.r ~ dnorm(0, 10)|(0, ) # sd for random effects
w ~ dnorm(0, 0.001) # electoral ward coefficient
for (n in 1:5) {
b[n] ~ dnorm(0, 0.001) # lower-level coefficient
}
for (n in 1:4) {
kappa[n] ~ dnorm(0, 0.001) # lower-level coefficient
}
}
}

```

References

- Andresen MA (2011) The ambient population and crime analysis. *Prof Geogr* 63(2):193–212
- Andresen MA, Malleon N (2011) Testing the stability of crime patterns: implications for theory and policy. *J Res Crime Delinq* 48(1):58–82

- Arcaya M, Brewster M, Zigler CW, Subramanian SV (2012) Area variations in health: a spatial multilevel modeling approach. *Health Place* 18(4):824–831
- Baumer E (2002) Neighborhood disadvantage and police notification by victims of violence. *Criminology* 40(3):579–616
- Besag J, York J, Mollie A (1991) Bayesian image restoration, with two applications in spatial statistics. *Ann I Stat Math* 43(1):1–59
- Boessen A, Hipp JR (2015) Close-ups and the scale of ecology: land uses and the geography of social context and crime. *Criminology* 53(3):399–426
- Braga AA, Clarke RV (2014) Explaining high-risk concentrations of crime in the city: social disorganization, crime opportunities, and important next steps. *J Res Crime Delinq* 51(4):480–498
- Braga AA, Papachristos AV, Hureau DM (2014) The effects of hot spots policing on crime: an updated systematic review and meta-analysis. *Justice Q* 31(4):633–663
- Brantingham PJ, Brantingham PL (1993) Nodes, paths and edges: considerations on the complexity of crime and the physical environment. *J Environ Psychol* 13:3–28
- Browne WJ, Goldstein J, Rasbash J (2001) Multiple membership multiple classification (MMMC) models. *Stat Model* 1:103–124
- Bursik RJ Jr, Grasmick HG (1993) *Neighborhoods and crime: the dimensions of effective community control*. Lexington Books, New York
- Ceccato V, Li G, Haining R (2018) The ecology of outdoor rape: the case of Stockholm, Sweden. <https://doi.org/10.1177/1477370818770842>
- Cohen LE, Felson M (1979) Social change and crime rate trends: a routine activity approach. *Am Sociol Rev* 44(4):588–608
- Coleman S (2002) A test for the effect of conformity on crime rates using voter turnout. *Sociol Quart* 43(2):257–276
- Congdon P (2003) Modelling spatially varying impacts of socioeconomic predictors on mortality outcomes. *J Geograph Syst* 5:161–184
- Congdon P (2008) The need for psychiatric care in England: a spatial factor methodology. *J Geograph Syst* 10:217–239
- Congdon P (2011) The spatial pattern of suicide in the US in relation to deprivation, fragmentation and rurality. *Urban Stud* 48(10):2101–2122
- Davies T, Johnson SD (2015) Examining the relationship between road structure and burglary risk via quantitative network analysis. *J Quant Criminol* 31(3):481–507
- Deryol R, Wilcox P, Logan M, Wooldredge J (2016) Crime places in context: an illustration of the multi-level nature of hot spot development. *J Quant Criminol* 32:305–325
- Diggle PJ, Moraga P, Rowlingson B, Taylor BM (2013) Spatial and spatio-temporal log-Gaussian Cox processes: extending the geostatistical paradigm. *Stat Sci* 28(4):542–563
- Dong G, Harris R (2015) Spatial autoregressive models for geographically hierarchical data structures. *Geogr Anal* 47:173–191
- Gelfand AE, Zhu L, Carlin BP (2001) On the change of support problem for spatio-temporal data. *Biostatistics* 2(1):31–45
- Gelman A (2006) Prior distributions for variance parameters in hierarchical models. *Bayesian Anal* 1(3):515–534
- Goldstein H (1994) Multilevel cross-classified models. *Sociol Methods Res* 22(3):364–375
- Goldstein H, Browne W, Rasbash J (2002) Partitioning variation in multilevel models. *Underst Stat* 1(4):223–231
- Goudriaan H, Wittebrood K, Nieuwebeerta P (2016) Neighbourhood characteristics and reporting crime: effects of social cohesion, confidence in police effectiveness and socio-economic disadvantage. *Brit J Criminol* 46:719–742
- Greenberg SW, Rohe WM, Williams JR (1982) Safety in urban neighborhoods: a comparison of physical characteristics and informal territorial control in high and low crime neighborhoods. *Popul Environ* 5(3):141–165
- Groff ER, 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(1):7–32
- Guerette RT, Bowers KJ (2009) Assessing the extent of crime displacement and diffusion of benefits: a review of situational crime prevention evaluations. *Criminology* 47(4):1331–1368
- Hagan J, Gillis AR, Chan J (1978) Explaining official delinquency: a spatial study of class, conflict, and control. *Sociol Quart* 19:386–398

- Haining R, Law J, Griffith D (2009) Modelling small area counts in the presence of overdispersion and spatial autocorrelations. *Comp Stat Data Anal* 53:2923–2937
- Hipp JR (2007) Block, tract, and levels of aggregation: neighborhood structure and crime and disorder as a case in point. *Am Sociol Rev* 27:659–680
- Johnson SD, Bowers KJ (2010) Permeability and burglary risk: Are cul-de-sacs safer? *J Quant Criminol* 26(1):89–111
- Johnson SD, Lab SP, Bowers KJ (2008) Stable and fluid hotspots of crime: differentiation and identification. *Built Environ* 34(1):32–45
- Kelsall JE, Wakefield JC (1999) Discussion of ‘Bayesian models for spatially correlated disease and exposure data’ by Best NG, Arnold RA, Thomas A, Waller L, Conlon EM. In: Bernardo JM, Berger JO, Dawid P, Smith AFM (eds) *Bayesian statistics*, vol 6. Oxford University Press, Oxford, pp 131–156
- Kirk DS, Matsuda M (2011) Legal cynicism, collective efficacy, and the ecology of arrest. *Criminology* 49(2):443–472
- Klinger DA, Bridges GS (1997) Measurement error in calls-for-service as an indicator of crime. *Criminology* 35(4):705–726
- Kubrin CE, Weitzer R (2003) New directions in social disorganization theory. *J Res Crime Delinq* 40:374–402
- Langford IH, Leyland AH, Rasbash J, Goldstein H (1999) Multilevel modelling of the geographical distributions of diseases. *J R Stat Soc C Appl* 48(2):253–268
- Law J, Quick M (2013) Exploring links between juvenile offenders and social disorganization at a large map scale: a Bayesian spatial modeling approach. *J Geograph Syst* 15(1):89–113
- Law J, Quick M, Chan PW (2015) Analyzing hotspots of crime using a Bayesian spatiotemporal modeling approach: a case study of violent crime in the Greater Toronto Area. *Geogr Anal* 47(1):1–19
- Leckie G (2013) Cross-Classified Multilevel Models. LEMMA VLE Module vol 12, pp 1–60. <http://www.bristol.ac.uk/cmm/learning/course.html>. Accessed 20 Sept 2018
- Li G, Haining R, Richardson S, Best N (2013) Evaluating the no cold calling zones in Peterborough, England: application of a novel statistical method for evaluating neighbourhood policing policies. *Environ Plann A* 45(8):2012–2026
- Li G, Haining R, Richardson S, Best N (2014) Space-time variability in burglary risk: a Bayesian spatiotemporal modelling approach. *Spat Stat* 9:180–191
- Lockwood D (2007) Mapping crime in Savannah: social disadvantage, land use, and violent crimes reported to the police. *Soc Sci Comput Rev* 25:194–209
- Lunn D, Jackson C, Best M, Thomas A, Spiegelhalter D (2012) *The BUGS Book: a practical introduction to Bayesian analysis*. Chapman and Hall/CRC Press, London
- Malleson N, Andresen MA (2015) The impact of using social media data on crime rate calculations: shifting hotspots and changing spatial patterns. *Cartogr Geogr Inf Sci* 42(2):112–121
- Miethe TD, Hughes M, McDowall D (1991) Social change and crime rates: an evaluation of alternative theoretical approaches. *Soc Forces* 70(1):165–185
- Moerbeek M (2004) The consequence of ignoring a level of nesting in multilevel analyses. *Multivar Behav Res* 39(1):129–149
- Morenoff JD, Sampson RJ, Raudenbush SW (2001) Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology* 39(3):517–560
- Openshaw S (1984) The modifiable areal unit problem. University of East Anglia, GeoAbstracts
- Ouimet M (2000) Aggregation bias in ecological research: How social disorganization and criminal opportunities shape the spatial distribution of juvenile delinquency in Montreal. *Can J Criminol* 42:135–156
- Quick M, Law J, Li G (2017) Time-varying relationships between land use and crime: a spatio-temporal analysis of small-area seasonal property crime trends. *Environ Plann B*. <https://doi.org/10.1177/2399808317744779>
- Quick M, Li G, Brunton-Smith I (2018) Crime-general and crime-specific spatial patterns: a multivariate spatial analysis of crime types at the small-area scale. *J Crim Just* 58:22–32
- Quick M, Li G, Law J (2019) Spatiotemporal modeling of correlated small-area outcomes: analyzing the shared and type-specific patterns of crime and disorder. *Geogr Anal* 51(2):221–248
- Rasbash J, Goldstein J (1994) Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model. *J Educ Behav Stat* 19(4):337–350
- Ratcliffe JH, McCullagh MJ (1999) Hotbeds of crime and the search for spatial accuracy. *J Geograph Syst* 1:385–398

- Richardson S, Thomas A, Best N, Elliott P (2004) Interpreting posterior relative risk estimates in disease-mapping studies. *Environ Health Persp* 112(9):1016–1025
- Rogerson P, Sun Y (2001) Spatial monitoring of geographic patterns: an application of crime analysis. *Comput Environ Urban* 25:539–556
- Rosenfeld R, Messner SF, Baumer EP (2001) Social capital and homicide. *Soc Forces* 80(1):283–310
- Sampson RJ (2013) The place of context: a theory and strategy for criminology's hard problems. *Criminology* 51(1):1–31
- Sampson RJ, Groves WB (1989) Community structure and crime: testing social disorganization theory. *Am J Sociol* 94(4):774–802
- Sampson RJ, Raudenbush SW, 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
- Shaw CR, McKay HD (1942) *Juvenile delinquency and urban areas*. University of Chicago Press, Chicago
- Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002) Bayesian measures of model complexity and fit. *J Roy Stat Soc B Met* 64(4):583–639
- Statistics Canada (2015) *Juristat Definitions*. <https://www150.statcan.gc.ca/n1/pub/85-002-x/2010002/definitions-eng.htm#v1>. Accessed 19 June 2019
- 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* 32:449–469
- Stucky TD, Ottensmann JR (2009) Land use and violent crime. *Criminology* 47(4):1223–1264
- Sutherland A, Brunton-Smith I, Jackson J (2013) Collective efficacy, deprivation, and violence in London. *Brit J Criminol* 53:1050–1074
- Taylor RB (1997) Social order and disorder of street blocks and neighborhoods: ecology, microecology, and the systemic model of social disorganization. *J Res Crime Delinq* 34(1):113–135
- Taylor RB (2015) *Community criminology: fundamentals of spatial and temporal scaling, ecological indicators, and selectivity bias*. New York University Press, New York
- Thomas A, Best N, Lunn D, Arnold R, Spiegelhalter D (2004) *GeoBUGS User Manual*. <http://www.mrc-bsu.cam.ac.uk/wp-content/uploads/geobugs12manual.pdf>. Accessed 20 July 2017
- van Wilsem J, Wittebrood K, De Graaf ND (2006) Socioeconomic dynamics of neighborhoods and the risk of crime victimization: a multilevel study of improving, declining, and stable areas in The Netherlands. *Soc Probl* 53(2):226–247
- Velez MB (2001) The role of public social control in urban neighborhoods: a multi-level analysis of victimization risk. *Criminology* 39:837–864
- Veysey BM, Messner SF (1999) Further testing of social disorganization theory: an elaboration on Sampson and Grove's 'Community structure and crime'. *J Res Crime Delinq* 36(2):156–174
- Waller LA, Carlin BP, Xia H, Gelfand AE (1997) Hierarchical spatio-temporal mapping of disease rates. *J Am Stat Assoc* 92(438):607–617
- 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
- Wheeler DC, Waller LA (2008) Comparing spatially varying coefficient models: a case study examining violent crime rates and their relationships to alcohol outlets and illegal drug arrests. *J Geograph Syst* 11(1):1–22
- Wilcox P, Eck JE (2011) Criminology of the unpopular: implications for policy aimed at payday lending facilities. *Criminology and Public Policy* 10(2):473–482
- Wooldredge J (2002) Examining the (ir)relevance of aggregation bias for multilevel studies of neighborhoods and crime with an example comparing census tracts to official neighborhoods in Cincinnati. *Criminology* 40:681–710
- Yin P, Mu L, Madden M, Vena JE (2014) Hierarchical Bayesian modeling of spatio-temporal patterns of lung cancer incidence risk in Georgia, USA: 2000–2007. *J Geograph Syst* 16:387–407