

# Examining the Relationship Between Road Structure and Burglary Risk Via Quantitative Network Analysis

Toby Davies · Shane D. Johnson

Published online: 11 October 2014

© The Author(s) 2014. This article is published with open access at Springerlink.com

## Abstract

**Objectives** To test the hypothesis that the spatial distribution of residential burglary is shaped by the configuration of the street network, as predicted by, for example, crime pattern theory. In particular, the study examines whether burglary risk is higher on street segments with higher usage potential.

**Methods** Residential burglary data for Birmingham (UK) are examined at the street segment level using a hierarchical linear model. Estimates of the usage of street segments are derived from the graph theoretical metric of betweenness, which measures how frequently segments feature in the shortest paths (those most likely to be used) through the network. Several variants of betweenness are considered. The geometry of street segments is also incorporated—via a measure of their linearity—as are several socio-demographic factors.

**Results** As anticipated by theory, the measure of betweenness was found to be a highly-significant predictor of the burglary victimization count at the street segment level for all but one of the variants considered. The non-significant result was found for the most localized measure of betweenness considered. More linear streets were generally found to be at lower risk of victimization.

**Conclusion** Betweenness offers a more granular and objective means of measuring the street network than categorical classifications previously used, and its meaning links more directly to theory. The results provide support for crime pattern theory, suggesting a higher risk of burglary for streets with more potential usage. The apparent negative effect of linearity suggests the need for further research into the visual component of target choice, and the role of guardianship.

---

T. Davies (✉) · S. D. Johnson  
Department of Security and Crime Science, University College London, 35 Tavistock Square,  
London WC1H 9EZ, UK  
e-mail: toby.davies@ucl.ac.uk

T. Davies  
Department of Mathematics, University College London, 25 Gordon Street, London WC1H 0AY, UK

**Keywords** Burglary · Street network · Environmental criminology · Crime

## Introduction

The role of urban configuration in shaping patterns of crime has been a theme of research for some time, and is a fundamental component of several prevailing criminological theories. In particular, it represents a central issue within environmental criminology (Brantingham and Brantingham 1981; Andresen 2014), one aspect of which seeks to reconcile the distribution of crime with the way in which the activities of individuals are shaped by the structure of the built environment.

When considering urban form and its relationship with crime, a natural object of study is the street network. The street network defines the paths which can be taken when moving between any two locations, and is therefore a primary determinant of both pedestrian and vehicular movement patterns. In addition, in modern cities it acts as the ‘skeleton’ of an urban area, in the sense that it is the structure around which elements of the built environment are arranged. Thus, it provides a convenient instrument for analysis of the more general environmental context. This is of particular significance in the case of crimes for which the target is an element of the built environment, and therefore located at a fixed point on the network; burglary, which will be considered here, is such an example.

When examining the street network, a natural unit of analysis is the street segment: the section of street that connects any pair of neighboring junctions. Use of this unit is analytically expedient for a number of reasons. On one hand, the fact that it is clearly and objectively defined means that it is methodologically convenient. On the other, that street segments represent micro level units of analysis means that they are well-aligned with the need to retain spatial granularity in crime analysis (Brantingham et al. 2009). Perhaps more importantly, the street segment is of theoretical significance. It has, for example, previously been shown that clustering at the street segment level is a significant driver of more general spatial patterns (Weisburd et al. 2004; Andresen and Malleson 2011). Furthermore, it is at the level of the street segment that numerous social processes (including criminal activity) take place. Street segments can be reconciled with relevant social concepts: they correspond to notions of community, for example, and can be classified in terms of the nature of activity taking place upon them. Indeed, segments of different types, in this sense, have been shown to display distinctive criminological character (Weisburd et al. 2012).

In discussion of the structural properties of street segments, reference is typically made to concepts such as permeability and accessibility. Although no universal definitions of these terms exist, they are broadly concerned with the ease with which places can be travelled to, or the readiness with which people are likely to travel through them to reach other destinations. Such ideas play a central role in the theory of routine activities (Felson 1987) and pattern theory (Brantingham and Brantingham 1993b). These concepts are, however, difficult to measure quantitatively, and the network properties that have typically been analyzed in the past have only oblique correspondence to the concepts of theoretical interest. This represents a significant shortcoming of previous research.

The aim of the present work is to formalize the study of street network properties in the context of crime by applying an approach to the analysis of networks that is more sophisticated than those previously employed. In particular, we examine the relationship between burglary risk and the network metric ‘betweenness’, which measures the frequency with which street segments feature in the shortest paths (i.e. those most likely to be used) for journeys through the network. In terms of urban studies, this represents an

estimate of the popularity or usage potential of particular street segments (Hillier 1996; Crucitti et al. 2006) for the journeys (pedestrian or vehicular) collectively made by the resident or ambient population through the street network. From a criminological perspective, it provides a representation of the collective awareness spaces that people, including offenders, may develop during the course of their routine activities. The use of this metric therefore provides a means by which concepts such as offender awareness—a core element of crime pattern theory—can be properly represented in empirical analyses. In addition to this structural analysis, we examine the physical properties of streets; in particular, their linearity, which is of relevance to theories of target choice and affects the ability of capable guardians to observe criminal activity along street segments.

The paper is organized as follows. In the next section, we discuss existing theoretical and empirical research on this topic. We then introduce a framework for the analysis of street networks, before discussing the statistical techniques employed and the data analyzed. Finally, we present our findings and discuss their theoretical and policy implications.

## Background

Routine activity theory (Cohen and Felson 1979) considers the ecological conditions necessary for crime to occur and has been highly influential in theories of crime pattern formation. In its simplest form, the theory states that direct contact predatory crimes can only take place when a motivated offender encounters a suitable target in the absence of a capable guardian (that would otherwise prevent the crime). Crimes take place at particular locations at specific times and so these three elements must simultaneously converge for a crime to occur. The street network is a fundamental determinant of this convergence. In the case of burglary, targets are positioned at fixed locations on the network, and their density varies from street to street. Furthermore, potential offenders encounter opportunities while moving around the network and must travel to and from offences; in both cases, the routes they take are constrained by the configuration of the network. Likewise, non-criminal pedestrian activity (which may affect guardianship), and that of formal guardians such as the police, is not random and is shaped by their journeys through the network. As suggested by Brantingham and Brantingham (1993b), the network therefore captures many of those aspects of the ‘urban backcloth’ relevant to the location of criminal activity, and hence influences the ‘potential’ for crime to occur.

The ways in which offender awareness and guardianship are shaped by the configuration of the street network can be understood by appealing to crime pattern theory (Brantingham and Brantingham 1993a). In the context of burglary, crime pattern theory asserts that offenders typically choose to victimize properties encountered during the course of non-criminal activities. More specifically, offenders are said to form awareness spaces, represented as cognitive maps, during their routine activities, and it is where these activity spaces intersect with suitable opportunities for crime that offenders are most likely to offend. Considered in these terms, the question of risk to a given property can therefore be translated to one concerning the extent to which it features in the awareness spaces of potential offenders. If it can be assumed that greater awareness corresponds to higher risk, then understanding the aggregation of awareness spaces becomes crucial in assessing the relative risk to each viable target. This, in conjunction with information concerning the distribution of opportunities, can be used to evaluate the overall spatial distribution of risk.

Since awareness spaces are formed during the course of regular activities, their location and extent are determined by the travel patterns that arise from such activities. Crime

pattern theory emphasizes the role of individuals' primary 'activity nodes' (e.g. homes, workplaces and leisure premises) in shaping these, with awareness spaces thought to be centered around these or along the routes between them. A straightforward corollary of this is that places that are activity nodes for many people are therefore the intersections for many individual awareness spaces and accordingly ought to have characteristic crime patterns. Empirical research has supported this, showing that such locations experience elevated levels of various crime types, including robbery (Bernasco and Block 2011), drug crime (McCord and Ratcliffe 2007) and liquor-related crime (Block and Block 1995). By similar reasoning, certain roads are likely to be used more than others during routine trips (the hierarchical nature of street networks means that certain roads feature in many journeys) and therefore ought to feature in the awareness spaces of many people, including offenders, and hence determine the criminal opportunities of which they become aware.

Turning to the issue of guardianship, the theoretical context is significantly more complicated, and the subject of some debate (Mawby 1977; Hillier and Shu 2000; Reynald and Elffers 2009). Pedestrian activity is, again, a crucial issue, since the number of potential guardians at a given time and place will be determined to a large extent by the movement of people and the routes they take. However, although the reasoning discussed above—that patterns can be understood in terms of journeys through urban space—can again be applied to understand variation in this, the effect on crime is not necessarily clear.

On one hand, it is suggested that increased pedestrian activity has the effect of increasing vigilance and therefore acting to reduce crime. This effect, succinctly encapsulated by Jacobs (1961) as the notion of "eyes on the street", may counteract the effect associated with increased offender awareness of busy places by providing a ready supply of guardians, the prospect of whose intervention dissuades potential offenders from committing crimes. Were this to be the case, less crime would be expected on those streets with the highest usage; this is a testable hypothesis in network terms. In addition, the implications of this from the perspective of planning are clear: that urban design should encourage pedestrian throughput as an indirect and natural method of crime control.

Some authors have, however, suggested that this effect is overstated and insufficient to offset the increase in risk associated with greater exposure. The willingness and ability of passers-by to intervene has been shown to be a complex issue (Reynald 2010), and may not be as significant an impediment to criminal acts as might be expected. In fact, an alternative school of thought suggests that the converse is true: that transient pedestrian activity (in sufficient volume) simply serves to diminish the territoriality of places and therefore renders them less daunting to potential offenders (Newman 1972). Ethnographic studies with burglars (Cromwell et al. 1991; Ashton et al. 1998) suggest that exploratory search strategies are relatively rare, and that offenders are generally unwilling to venture to areas where they will 'stand out'. Consequently, locations that are already known to them and for which the ambient population is transient may be particularly attractive.

In line with these arguments concerning the nature of guardianship, Brantingham and Brantingham (1975) considered the configuration of neighborhoods, finding that crime tended to be higher at the boundaries, where community structure is likely to be less well-defined. This finding is also consistent with theories of collective efficacy (e.g. Sampson et al. 1997), which argue that stronger, more cohesive communities—where residents know each other and share common goals—are better equipped to defend against and deter intruders, or to control the behaviour of residents. Most studies of collective efficacy are conducted at the area level (e.g. Sampson and Raudenbush 1999), but Weisburd et al. (2012) examine this at the street segment level, and report findings consistent with the theory. Thus, in contrast to Jacobs' concept of eyes on the street, the effectiveness of

guardianship from this perspective may be determined not just by the number of people moving along a street, but also by who those people are. For busier streets, for which the ambient population is likely to be transient, the mix of people using them may be less conducive to the production of effective guardianship.

Several empirical studies have examined various aspects of the relationship between the street network and burglary. The earliest is the work of Bevis and Nutter (1977), who examined burglary levels on isolated streets and studied the effect of network density at the census tract level, using data from Minneapolis, USA. The concept of density is defined in that work as the ratio of the number of street segments to the number of junctions. This is a relatively coarse-grained metric, but nevertheless a relevant one: a minimal functioning street network would be tree-like in structure, with relatively long trips required to reach most destinations, and would yield a low score on this metric. Increasing density via the addition of more segments adds redundancy and therefore plurality in terms of travel path choices, with the distribution of throughput becoming more balanced as a result. A regression analysis presented in the study revealed a positive relationship between area-level density, defined in this way, and burglary levels, after controlling for various social and demographic variables. Individual streets which were categorized as inaccessible on the basis of the number of connecting segments were also found to experience less victimization, on average, than their more connected counterparts.

Again working at the area level, White (1990) examined burglary risk across 86 neighbourhoods in Massachusetts, USA. In this case, the independent variable was the number of roads in each neighborhood that were connected to (categorically-defined) major roads. While offering relatively little insight into traffic flow, such a metric most obviously corresponds to permeability in that it provides an indication of the ease of access of neighborhoods. As with the Bevis and Nutter study, after controlling for various other relevant factors, it was found that network structure was a statistically significant factor in determining area-level crime rates, with more accessible areas having higher burglary rates.

At a lower spatial scale—the street segment—Beavon et al. (1994) considered levels of crime risk (including, but not limited to, burglary) in Ridge Meadows, Canada. They used two metrics to capture different network effects: for a given segment, accessibility was measured as the number of other segments with which it shared a junction, whereas flow was estimated via a hierarchical road classification system. After again controlling for various other factors in the statistical analysis, both factors were found to be positively associated with crime: risk increased with the number of connected streets, and roads classified as being more important also experienced additional victimization.

The role of the street network in urban processes is one of the primary foundations of ‘space syntax’, an approach that has been applied in a variety of urban contexts (Hillier 1996). Much of this work rests on the fundamental observation that individual street segments cannot properly be understood in isolation, but must be considered in the context of the rest of the network and their position within it. In addition, it places particular emphasis on lines of sight as a determinant of connectivity (and, in some cases, as the definition of the street segment itself), which, as noted previously, is of particular relevance in the context of crime. Several metrics have been developed using these ideas, such as ‘integration’, which measures how close (or easy to access) a given location is to all others in terms of paths through the network. Applying these methods to crime, and analyzing data from London, UK, and an urban area in Australia, it was found that crime was positively related to connectivity but negatively related to integration (Hillier 2004). The conclusions drawn from this are that permeable designs are favorable, but that where

redundant connectivity is present (that which does not increase integration) the effect can be reversed, perhaps because of the provision of extra entry or escape routes.

The space syntax work also focuses on one particular class of street segment: cul-de-sacs. These represent a neat example for analysis since they can be universally defined, relatively independently of spatial context, and represent an extreme case in theoretical terms, since they will usually be remote and experience little through traffic. In the analyzed data, risk on these segments was found to be considerably higher than on others, although it was emphasized that physical shape must be taken into account. Linear and well-connected segments—which may offer good lines of visibility to capable guardians—were found to be safe, whereas those which were sinuous and relatively secluded were at higher risk (Hillier 2004). Clearly this is a finding that runs contrary to what would be anticipated according to pattern theory.

Hillier's work makes several other observations at the level of individual properties, in particular relating to their modes of access (e.g. the direction they face or their proximity to alleys). Such features were also considered by Armitage (2007), whose work involved detailed assessment of the physical features of individual houses. These observations included the type of road on which they were situated and a subjective estimate of its usage, both of which were then individually compared with crime levels. Once again it was found that increased activity and permeability was associated with higher risk, with the difference between isolated cul-de-sacs and those serviced by pedestrian alleys being a case in point.

A weakness of many of the studies conducted previously regards the statistical framework employed. Numerous studies have presented only descriptive statistics or have used simple bivariate tests of correlation. Where statistical models have been employed, street segments have usually been treated as independent units of analysis, which, of course, they are not. A recent paper by Johnson and Bowers (2010) explored many of the issues discussed above, but using a statistical framework specifically designed to account for the hierarchical structure of the data—street segments are nested within neighborhoods, and neighborhoods are nested within larger areas—and the fact that, while crimes are rare, incident counts may exceed 1 (an issue also not accounted for in previous studies). Using burglary data from Merseyside, UK, they considered variables similar to those seen in the previous work at the street segment level—number of connections, road classification and physical shape—with a particular emphasis on cul-de-sacs. Their hierarchical linear analysis showed that, in addition to significant variation at higher levels of spatial aggregation, there was a positive effect of connectivity on crime, and that higher counts of victimization were observed on major roads (all other things being equal). Cul-de-sacs were found to be the safest type of street segment, but with a marked difference between those which were sinuous in form (which had lower crime counts) and those that were linear. The findings of this study support the general hypothesis that permeability is associated with higher crime, but studies of this scale and that use appropriate statistical methods are rare. Moreover, the metrics used to measure permeability, itself a proxy for offender awareness, were still relatively crude.

Though the distribution of crime was not the primary focus, the work of Iwanski et al. (2012) is also notable for its treatment of the street network. Working with data for offenders' homes and offence locations, the work examined the hypothesis that crimes were committed while the offender was in the process of a longer journey. Estimates were produced for the onward paths offenders may have been taking, where navigational choices were based on a combination of directionality and the likely popularity of roads (estimated by simulating a number of journeys through the network). Results show that these paths are

biased towards ‘crime attractors’—locations that concentrate criminal activity because of known opportunities for crime (Brantingham and Brantingham 1995)—in the form of major shopping centres. The findings are, however, also consistent with a somewhat simpler hypothesis that crime attractors tend to be located on roads with high activity (see Hillier 1997).

Computer simulation has also been used as an experimental approach in itself, most notably in the work of Groff (2007). In the study in question, it is argued that the influence of the street network on offender spatial behavior is so great that its exclusion from models represents a significant shortcoming. Indeed, it is demonstrated that the behavior of an ABM situated on a realistic street network (imported via GIS) differed significantly from that observed when an abstract grid is used instead. Although the model relates to street robbery rather than burglary, the fact that the theoretical focus of the work concerned routine activities suggests that the findings are applicable in the present context. No attempt is made within the study to relate the patterns observed to network properties, but the fact that urban structure influenced the distribution of crime is significant in itself.

As noted, one of the limitations of previous work in this area concerns the methods used to measure network properties; specifically, the level of detail afforded by variables and their appropriateness as measures of the underlying quantity of interest. To take one example, categorical road classification (e.g. major, minor, and so on) is used as a proxy measure for the likely activity levels on a street in a number of studies (White 1990; Beavon et al. 1994; Johnson and Bowers 2010), but suffers from a number of shortcomings. The first of these is technical, and concerns the fact that the number of categories defined by traditional classification systems is typically small. Few systems include more than ten, and this is reduced further if certain types can be discounted as potential crime locations (such as motorways, in the case of burglary). Considerable variation is likely to be disregarded when such a coarse partition is used. This is particularly problematic when it is noted that research (Johnson 2010; Weisburd et al. 2012) suggests that crime risk exhibits sharp spatial discontinuities, such that adjacent street segments can vary considerably in the level of risk observed.

The use of categorical systems also poses conceptual problems, since classifications do not necessarily correspond directly to the properties of interest to a researcher. While broadly indicative of likely usage, for example, road classifications are primarily an administrative construct, and numerous other factors are taken into account in their specification. In addition, they are defined for roads in their entirety (as opposed to individual street segments), and do not capture variations between different sections of the same road, which can be considerable. To assume that roads are homogeneous entities is a form of the ecological fallacy (Robinson 1950), and real-world examples will be readily available to the reader of pairs of street segments for which the relative levels of usage conflict with those which would be expected on the basis of classification. Finally, the use of such categories severely compromises the wider applicability of research, particularly with respect to inter-national comparison. Road classification systems differ from country to country, often with no natural correspondence between categories, and comparison between nations on this basis is therefore futile in all but the most general terms.

The selection of appropriate alternative network metrics, however, requires careful consideration of precisely the sense in which urban activity is to be measured. Although theories based on routine activities emphasize the role of journeys between locations, the effect of the network can also be manifested in other ways. For urban movements in general, for example, certain areas will experience less activity if they are more difficult to access, in some sense, and again this is likely to imply lower criminal activity. Although the effect is similar, it is not precisely the same as that due to routine activities, which are shaped by the

need to travel down certain paths, rather than the ease of doing so. The extent to which an area is readily accessed—typically referred to as its ‘permeability’—is not necessarily equivalent to how regularly it is used by pedestrians: an area may be well-connected via streets, but if there is no reason to travel along them, it is unlikely to see substantial activity. At the same time, these streets may be more likely to be found during exploratory behavior than their more secluded counterparts (or, indeed, invite such behavior), and may be attractive for reasons of access or escape. Such exploratory behavior by offenders, however, is thought to be rare (Cromwell et al. 1991; Ashton et al. 1998), and our focus here is primarily on the effect of the network in shaping routine journeys. The principal metric we choose (see below) is thus one that is defined in terms of such journeys, and therefore corresponds directly to this issue. As such, it is more closely aligned with theory—routine activity and crime pattern theory in particular—than those used in the previous work reviewed above, whose primary interpretation is in the sense of ease of access.

The rapidly-developing field of network science (for an introduction, see Newman 2010) provides an array of metrics for the rigorous analysis of networks. A network is, generically, a collection of discrete components (nodes) and the connections between them (links), and can be adapted to various contexts. Spatial networks—where nodes have some definite location—represent a distinctive subfield (reviewed by Barthélemy 2011), requiring bespoke treatment, and street networks have received considerable attention (Porta et al. 2006a, b; Masucci et al. 2009; Chan et al. 2011). Several metrics tailored to the requirements of urban analysis have been developed (Crucitti et al. 2006), and the properties of a diverse array of real-world examples have been measured and compared (Jiang 2007; Strano et al. 2013).

Clearly, the motivation for this work reaches beyond the structure of networks themselves, and focuses additionally on social processes taking place upon them. Obvious examples of this include applications to traffic modeling, but other processes likely to have a relationship with urban movement are also of interest. The potential of this avenue has been demonstrated by Porta et al. (2009, 2012), who found economic activity at the street level to be positively associated with an integrated measure of network centrality in data from Bologna, Italy, and Barcelona, Spain.

The aim of the present work is thus to extend understanding of the influence of the street network on crime by employing an explicitly quantitative analytical framework for the study of urban streets, thereby formalizing several of the general ideas which have been explored to date. This will be achieved by analyzing the relationship between the locations of burglaries and several highly-granular network metrics which have not previously been employed in this setting, but which align more closely with the theories examined. The use of such techniques not only strengthens the relationship between empirical observation and theory, but suggests an approach which can be employed in future analyses and that might inform practical interventions. In what follows we describe the data analyzed and the analytic framework. Specific hypotheses associated with the different network measures are also stated explicitly.

## Methods and Results

### Data

Analyses are conducted for the city of Birmingham (UK), which is the second-largest city of the UK and is, in general, unremarkable in terms of its socio-demographic



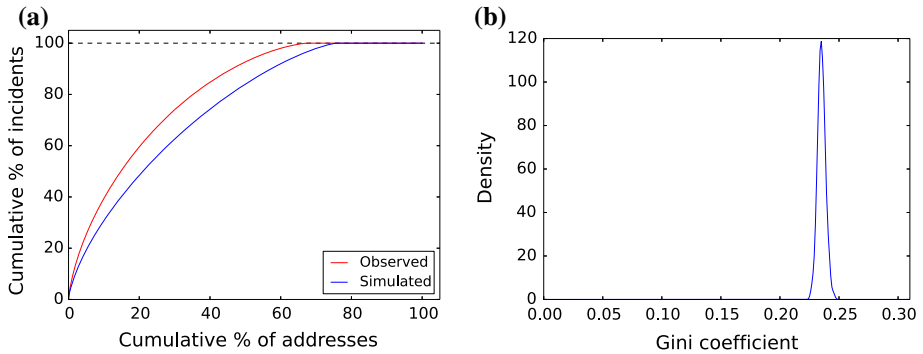
composition. Police recorded crime data were provided by West Midlands Police, and comprise all recorded incidents of residential burglary for the 4-year period from April 2009 to March 2013, inclusive. There were 27,383 such incidents. For each incident, the location of the victimized property is recorded both in terms of its full address and a grid reference. Since the analysis requires that we control for the opportunity for burglary, Ordnance Survey (OS) data, which provide the exact location of every home in Birmingham, are also used. In order to maintain consistency between the datasets, and to ensure the accuracy of the grid references associated with the burglary events, all offenses were geocoded by the first author using a bespoke text-matching algorithm and the OS data. This algorithm parses each incident address, separating it into its component postal code (if available), road name and house number, and seeks a match amongst all known addresses. Using this approach, 26,614 (or 97.2 % of) incidents were successfully geocoded (with no systematic aspect to the attrition), and it is these events that are analyzed here.

Data for the street network for the city of Birmingham were obtained from the OS. These data include the geometry of every street segment, along with other information, including: a hierarchical classification ('Motorway', 'A Road', 'B Road', 'Minor Road', 'Local Street' or 'Private Road'), the nature of each street segment (e.g. 'Single Carriageway', 'Traffic Island Link', 'Roundabout') and the name of the street. In the OS data, the network is defined using junctions, where a junction is defined as any point at which two roads intersect (in a physical sense; that is, regardless of street name or continuity) and a 'street segment' is any portion of road which connects two junctions. Prior to mathematical and statistical analysis, some pre-processing was required to correct features liable to distort the analysis. For example, roundabouts, which appear as several segments but are more practically thought of as single intersections, were collapsed to single points, and redundant link roads at junctions were removed. In addition, since some of the metrics computed are susceptible to 'edge effects' (where misleading results are found near the spatial extremities, see Porta et al. 2006a), all streets within a 2 km buffer of the study area were also included during the calculation.

The final stage of data preparation concerned the association of the address points with the street segments on which they lie. Since the accuracy of this process is crucial to the validity of the analysis, the association was not based purely on distance (which leads to frequent errors), but also involved matching on the basis of street names. To be precise, for each address point, a 40 m radius was searched for all segments with a matching street name: if any existed, the closest of these was taken to be the match; if none existed, the closest segment was selected, regardless of street name. A suitable segment with matching street name was found in 94.2 % of cases, and the remainder displayed no systematic spatial pattern. For address points located near junctions, the street name was used to determine the segment with which the point should be associated.

### Clustering of Crime on Networks

To establish the premise for the remainder of the analysis, we test the most basic hypothesis (Hypothesis 1) that crime is distributed heterogeneously around the network. To do this, we replicate the approach taken by Johnson and Bowers (2010) to compare the observed concentration of crime events at the segment level with that expected assuming the null hypothesis. Since houses themselves are not evenly distributed across the network, segments do not have an equal opportunity for crime. Consequently, the analysis is more involved than simply checking for an absolute departure from homogeneity, and so a Monte Carlo simulation approach is used. For the observed data, all street segments with at



**Fig. 1** Clustering of burglary incidents at street segment level: **a** Lorenz plots for both observed and simulated data, and **b** the distribution of the Gini coefficient across all 99 simulated datasets

least one home are ordered according to their crime rate (from highest to lowest), and a Lorenz plot produced of the cumulative fraction of addresses against the cumulative fraction of incidents. This plot is shown in Fig. 1a.

To produce an expected distribution, multiple synthetic sets of incidents are created, under the assumption that all addresses are equally likely to be victimized. To do this, for each synthetic dataset, we randomly select  $K$  homes (with replacement) from all those possible (where  $K$  is the number of burglaries in the observed data). Homes are selected using a uniform random number generator, and by making the selection with replacement we allow for simulated repeat victimization of the same household (see Pease 1998). A Lorenz curve is then plotted for each synthetic dataset and compared with the observed distribution. A full permutation is practically infeasible (and, frankly, unnecessary), so in this case we generate 99 synthetic datasets. The mean of these Lorenz curves is shown as Fig. 1a.

A quantitative measure of the difference between the observed and synthetic distributions is the Gini coefficient (Gini 1921), which measures the extent to which some quantity is distributed unequally in a population (in this case incidents across street segments). The Gini coefficient ranges from zero to one, with larger values indicating increasing departure from equality. In terms of the Lorenz curves, the Gini coefficient corresponds to the area between the observed and expected curves, with values above zero indicating an inequality. In this case, the observed distribution differs from that expected assuming the risk of victimization is uniformly distributed across homes. Figure 1b shows the distribution of the Gini coefficient for all 99 synthetic distributions produced; it is clear that the statistic is densely clustered around a value of 0.24 and does not approach the value of 0, which would be the case under the null hypothesis. The statistical significance of the difference between the observed and expected distributions can be estimated by comparing the distribution of the Gini coefficient with the expected point estimate, assuming that the observed and expected distributions do not differ (i.e. that the Gini coefficient is 0). To do this, all 99 values of the Gini coefficients are placed in a rank-ordered list and the value  $r$  is taken to be the position which 0 would take in this list (North et al. 2002). The statistical significance  $p$  is then defined as

$$p = \frac{n - r + 1}{n + 1} \quad (1)$$

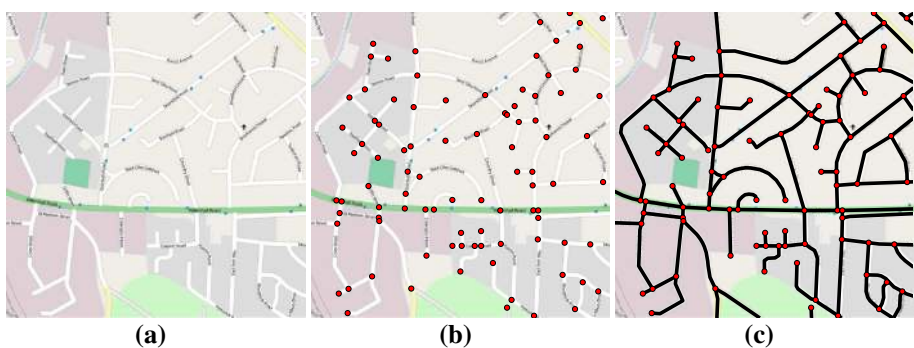
where  $n$  is the number of synthetic distributions. In this case, we can reject the hypothesis that the clustering of incidents at street segment level is simply due to chance, with 0.01 significance.

### Street Network Analysis

In order to facilitate their rigorous analysis, and to make sense of the metrics used here, the structure of the street network must first be expressed using the terminology of network theory, which is based upon ideas and techniques from the mathematical field of graph theory (for a formal introduction, see Bollobás 2002). A *network*  $G = (V, E)$  is a set of *nodes*,  $V$ , and a collection of *links*,  $E$ , between pairs of nodes (also referred to as *edges* within graph theory). The number of nodes in a given network, known as its *order*, is denoted  $N$  and the number of links (its *size*) is represented as  $M$ . The nodes are labelled using the integers  $1, \dots, N$ , where the ordering is unimportant as long as the labelling is consistent and unique, and each node is then referred to by its label. Where a link exists between two nodes  $i$  and  $j$ , the nodes are said to be *adjacent* and the link is represented by the unordered pair of nodes  $(i, j)$  (although the single letter  $e$  is also used to refer to a generic link).

The translation of a street network to these terms can be done in more than one way, and the choice of which is most appropriate is dependent upon the issue at hand. One popular choice is the ‘dual representation’ (Porta et al. 2006a), in which whole roads (sets of street links which have been associated on the basis of street name or geometry) are represented as nodes, and links are placed between any two which share a junction. Although the process of associating streets is useful and realistic, it is reliant on consistent street naming, and may not cater well for situations where streets are topologically the same but have different names (or vice versa). In addition, spatial information, such as the length of streets or their location, is discarded. Perhaps a more natural representation, and the one we consider here, is the ‘primal representation’ (Porta et al. 2006b), in which each junction is represented by a node, and a link is added between any two nodes which are connected by a road (there is therefore a direct correspondence between street segments and links). An example of this construction is shown in Fig. 2.

As stated previously, one complication in the study of street networks, and spatial networks more generally, is that many of the classical metrics used in network analysis are



**Fig. 2** The construction of the primal representation of a street network: **a** the original map, **b** nodes placed at each junction, **c** links added between any pair of junctions connected by a street segment

of limited relevance. The *degree* of a node, for example, is the number of links coincident with it (in this context, the number of streets meeting at a junction) and is the most widely-studied network metric. For many real-world networks this can take a wide range of values and offers great insight; however, the physical constraints of street networks mean that it can take a very limited number of values—to find junctions at which 6 or more segments meet is extremely rare, and the vast majority of nodes have degree 1, 3 or 4—and so (despite its frequent use in previous criminological studies) it is not an effective discriminant.

Many of the more meaningful metrics are concerned in some way with travel through the network, and we therefore introduce the concept of a *path*. A path in a network is any ordered sequence of nodes such that every consecutive pair of nodes is connected by a link (that is, a sequence of nodes which can be traversed by following links). The length,  $l$ , of such a path can be defined in two ways: in purely topological terms, as the number of links it comprises (which is 1 fewer than the number of nodes in the path), or in metric terms, as the sum of the physical length of all constituent links. The choice of which to use depends on interpretation: metric distance is a true measure of cost, but topological distance has been suggested to be more representative of distance as perceived by an individual navigating the network (Hillier and Iida 2005). In this paper, we test hypotheses using both variants. All analysis is identical under either definition, but the distinction will be re-visited later.

Clearly, for any pair of nodes  $i, j \in V$  it can be determined whether a path between the two exists, and indeed there may be more than one. A *shortest path* between  $i$  and  $j$  is one such path of minimal length (though, again, there may be more than one), and this length is denoted  $d_{ij}$ . To recapitulate in real-world terms, then,  $d_{ij}$  is the shortest distance (either in topological or metric terms) one would have to cover to travel between two junctions  $i$  and  $j$  through the street network.

If we define  $\sigma_{ij}$  as the total number of shortest paths between  $i$  and  $j$ , and then, more specifically,  $\sigma_{ij}(e)$  as the total number of shortest paths between  $i$  and  $j$  which pass through the link  $e \in E$ , we can define the *betweenness centrality*  $B_e$  (Freeman 1977) of a given link  $e$  as:

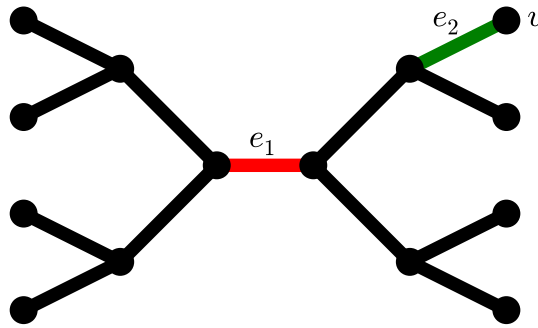
$$B_e = \sum_{i,j \in V, i \sim j} \frac{\sigma_{ij}(e)}{\sigma_{ij}} \quad (2)$$

where  $\sim$  here represents the relation ‘there exists a path between  $i$  and  $j$ ’. Re-expressed in simple, non-mathematical terms, and in the language of streets, the process of its calculation is:

1. For a given pair of junctions A and B in the network, find the shortest path(s) between them;
2. For every street segment that appears in the shortest path(s), increment its betweenness centrality by the proportion of shortest paths that it appears in (so if there is only one shortest path between A and B, add 1 to the value of every segment in it);
3. Repeat steps 1 and 2 for every pair of junctions.

The significance of the metric is perhaps best clarified by considering an example, and the stylised network shown in Fig. 3 is illustrative of its discriminatory value. The two links identified—one peripheral and one highly central—are those with betweenness values at each extreme.

Betweenness has a relatively clear interpretation in real-world terms as an estimate of the use of any given link by traffic (pedestrian or vehicular) passing through the network. Although the premise of single trips between all junctions is crude, it nevertheless



**Fig. 3** Stylised illustration of link betweenness: link  $e_1$  (shown red) features in any path between one of the seven nodes on the ‘left’ of the network and the seven on the ‘right’, and therefore has a relatively high betweenness value of 49. Link  $e_2$  (green), on the other hand, is only traversed by paths starting/ending at  $v$ ; there are 13 such paths and it therefore has a relatively low value of 13 (Color figure online)

represents a well-motivated first-order heuristic for urban movements. Of particular note is the fact that the value for every segment depends entirely on its role in the network as a whole. This, however, is a feature that can be refined somewhat: rather than considering all possible trips through the network, it may be beneficial to count only trips shorter than a certain length, particularly when the social processes we have in mind take place over a short range. If we take this maximum *radius* to be  $r$ , we therefore define

$$B_e^r = \sum_{i,j \in V, i \sim_j} \frac{\sigma_{ij}(e)}{\sigma_{ij}}, \tag{3}$$

where  $\sim$  has been modified to represent ‘there exists a path of length at most  $r$  between  $i$  and  $j$ ’. This can be defined for any value of  $r$ , topological or metric.

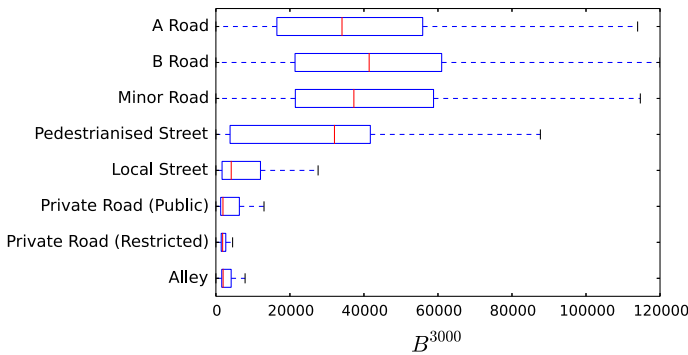
By limiting the length of trips to short distances of around 500m (or few topological steps) the approach can be used to identify those segments that are likely to feature in journeys through the network that represent ‘local’ trips (e.g. Hillier 1996). For such trip lengths, therefore, the metric of betweenness may be thought of as estimating the ‘usage potential’ of segments for mostly pedestrian movement, or trips made by those who likely live or have routine activity nodes in the area. By extending the length of the journeys considered to distances of 5km or more (or many topological steps), the paths identified can be thought of as approximating more ‘regional’ movements, or those most likely to be taken in vehicles. By this reasoning, the approach can also be used to estimate the usage potential of segments for vehicular journeys or trips made by those from outside the area.

The concept of betweenness is particularly pertinent to the study of crime through the prism of pattern theory, since it approximates exactly those journeys through the network which would be expected to shape awareness spaces. It is also notable that many of the factors investigated in previous work are implicit within its definition. To take the example of cul-de-sac segments, these will be characterised by low betweenness values since the only trips in which they feature will be those to the end of the cul-de-sac (of which there are relatively few); through roads, on the other hand, will feature in many trips.

In addition, instead of relying on road classification as a surrogate estimate of activity, we now have a method which relates much more directly (and with finer granularity) to the underlying concept of interest: the usage potential of street segments, which we take to represent an estimate of (offender) collective awareness spaces. While the classification of

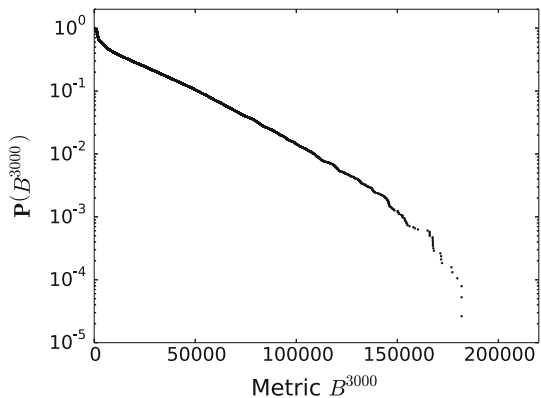
roads as ‘major’ is, of course, primarily because they act as conduits for a large volume of journeys, there will naturally be considerable variation in how this is realised. Figure 4 shows the typical values of betweenness for the various classes of street in the data, showing a general trend in the expected direction but, crucially, with considerable overlap between categories.

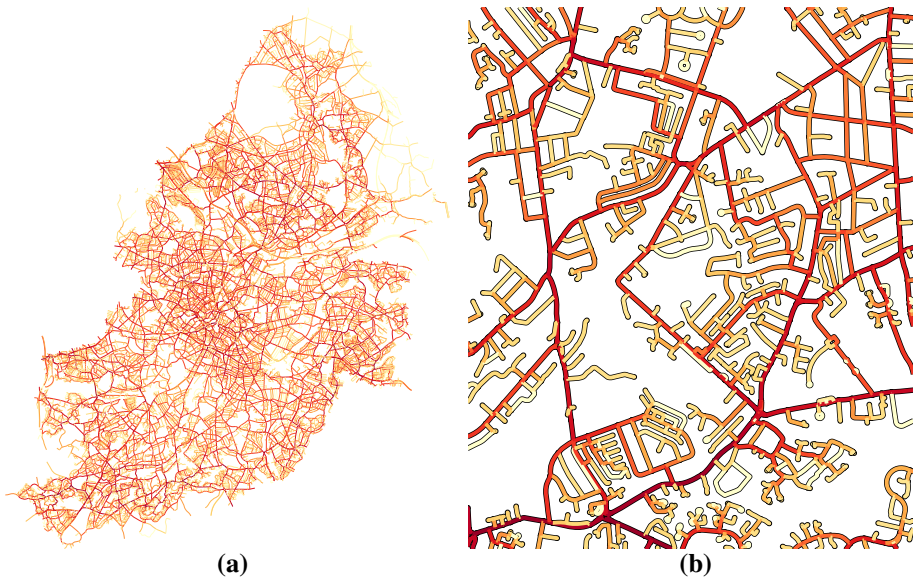
Figure 5 shows the distribution of betweenness values in Birmingham, represented as a complementary cumulative density function and plotted on semi-logarithmic axes. The general shape of the distribution is indicative of an exponential form, as has been found in previous published work by Crucitti et al. (2006) (although it should be noted that those results concerned betweenness for nodes, rather than links). Since the extremely large values in the tail are somewhat problematic for the later statistical analysis, and since our interest in betweenness is based primarily on its ability to suggest a hierarchy of streets, we rescale these values by taking their natural logarithm, and it is this value that is used in the remaining analysis. These values are used to produce the maps shown in Fig. 6, where it can be seen that betweenness identifies a hierarchical, almost-skeletal, structure for the network as a whole. Our predictions concerning the relationship between these properties and crime are based on the material reviewed in the introduction. With all else equal, we expect the level of burglary to be higher on street segments that are more likely to feature



**Fig. 4** Boxplots of betweenness values,  $B^{3,000}$ , for various OS road classifications, using a metric radius of 3,000 m. Categories are ordered on the basis of approximate hierarchical position

**Fig. 5** Distribution of betweenness,  $B^{3,000}$ , for the street network of Birmingham, using a radial limit of 3,000 m. The plot shows the complementary cumulative distribution of  $B^{3,000}$ ; formally,  $P(B^{3,000})$  is defined by  $P(B^{3,000}) = \int_{B^{3,000}}^{\infty} \frac{M(B')}{M} dB'$ , where  $M$  is the total number of links and  $M'$  the number of links with betweenness equal to  $B$





**Fig. 6** Street segments coloured according to their betweenness for **a** the entire street network of Birmingham, and **b** one smaller section. Betweenness values are calculated on the basis of a limited radius of 3,000 m

in the awareness spaces of offenders; that is, those with higher betweenness values (Hypothesis 2).

As discussed above, the geometry of a street segment can influence the lines of sight along it (Hillier 1996), which can, in turn, affect the extent to which capable guardians can observe what is happening on them. Consequently, we also consider a simple measure of the geometry of individual street segments. This quantity is the *linearity* of the segment, denoted  $L$ , and measures the extent to which it deviates from a straight line. For a given segment,  $e$ , it is defined as the ratio of the Euclidean (as-the-crow-flies) distance between its two end-points,  $s_e^{Euc}$ , to the true length of the segment,  $s_e^{true}$ , so that its value is

$$L_e = \frac{s_e^{Euc}}{s_e^{true}}. \tag{4}$$

A perfectly linear segment, for which the two lengths will be equal, will therefore take the value 1, with the value decreasing as the deviation from a straight line increases. This is a continuous and objective alternative to the binary classification used in previous work (Johnson and Bowers 2010). Our third hypothesis, then, is that levels of burglary will be lower on more linear street segments.

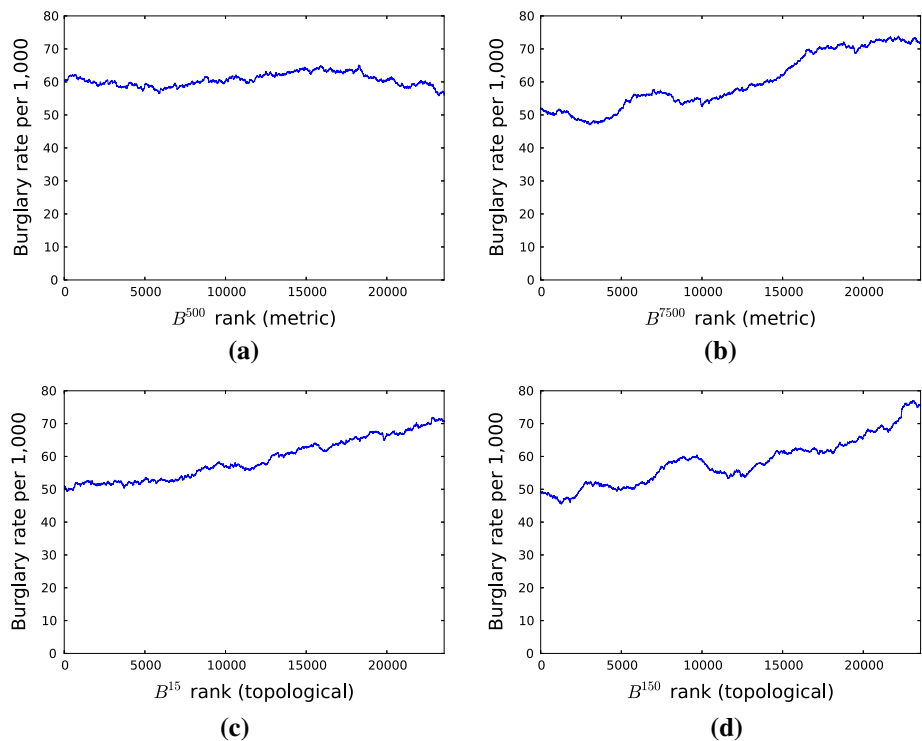
### General Relationships

Before quantifying the relationship statistically, considerable insight into the variation of burglary rate with network properties can be gained through exploratory data analysis. One simple approach is to order all street segments (with at least one dwelling) according to their increasing betweenness and to examine the burglary rate (the count of crimes per

1,000 dwellings) as we move through this ordering. Since burglary is a rare event, any trend is likely to be somewhat opaque if individual segments are examined, and so the data are smoothed by considering aggregated groups of 3,000 segments. For each frame of Fig. 7, therefore, the value plotted for a given betweenness rank  $r$  is the total burglary rate for the street segments ranked between  $r$  and  $(r + 3,000)$  over the 4 year period.

The apparent positive association of betweenness and burglary rate when the data are presented in this way is compelling, though it is clear that the precise form (and strength) of this is dependent upon the specific way in which betweenness is calculated. The trend is most clearly seen when betweenness is calculated in topological terms, with the smoothness of the relationship at the relatively low radius of 15 (see Eq. 3) particularly striking. When the physical length of street segments is taken into account, the expected relationship is also present, but only becomes apparent at higher radii (and, indeed, is almost absent for a radius of 500 m).

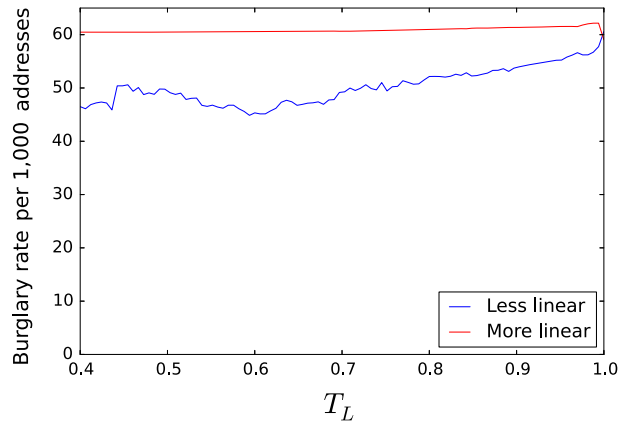
Considering linearity, its use as a means of ordering street segments is relatively uninformative, since the vast majority of segments have values of 1 or very close to 1 (street segments with linearity of 0.99 or greater account for 65 % of address points). Instead, we relate it to previous work by classifying segments dichotomously on the basis of linearity and comparing the burglary rates of each group. For a given threshold value,  $T_L$ , the aggregate burglary rate is found for all street segments with linearity less than  $T_L$ , and the



**Fig. 7** The variation of burglary rate when moving through a list of street segments ordered according to betweenness. The values plotted for any given rank represent the aggregate burglary rate for the 3,000 segments at that position in the list. The panels show the relationship for different variants of  $B^r$ : for metric  $r$  of **a** 500 m, or **b** 7,500 m, and for topological radii of **c** 15 or **d** 150



**Fig. 8** Comparison of crime rates when street segments are divided into two groups on the basis of linearity,  $L$ , falling above or below  $T_L$



same quantity found for all segments with values greater than or equal to  $T_L$  (recalling that lower linearity corresponds to greater deviation from a straight line). The crime rates for each of these groups as the threshold value is varied are plotted in Fig. 8.

Two things are immediately notable from this figure: that more linear street segments experience higher rates of burglary, and that that is true across almost all definitions of what constitutes ‘more linear’. This provides support for the similar conclusions found for cul-de-sacs by Johnson and Bowers (2010), and seems to suggest that the linearity measure here is capable of identifying the same streets as were found by manual classification in the earlier work.

### Regression Analysis

Although fairly clear patterns are identified by the above analysis, the possibility exists that the apparent trends are spurious. Many socio-demographic variables have been shown to influence burglary patterns, and even in purely geographical terms it is possible that the patterns displayed graphically are simply artefacts of variation at some other geographical level. To control for these, and to quantify the strength of any relationships which may exist, we use a hierarchical linear model (HLM). This is a statistical object explicitly designed to handle multi-level data; that is, data which can be grouped in a meaningful and systematic way.

In this case, the levels in question are geographical units at various spatial scales. It is to be expected that segment-level patterns will vary with the characteristics of the area in which they are situated, and to a certain extent this variation can be controlled for by including area-level independent variables for each data point. Nevertheless, some effects will not be captured explicitly, as data will be unavailable for them and they will hence go unobserved. However, the inclusion of random error terms in the (multi-level) regression model can be used to estimate their effects. The fact that such effects are unlikely to act uniformly across areas, however, means that they require special treatment. In statistical terms, this issue is accounted for by allowing the unknown effects in question to vary from area to area, with the extent of this variation calculated using methods of maximum likelihood estimation.

The spatial structure of the data is considered here at three levels: the basic street segment unit, and two nested areas. The first of these is the output area (OA), a UK census

unit which typically contains approximately 150 households. This is taken as an approximate encoding of the notion of ‘neighborhood’, and there are 3,127 in the city of Birmingham. In some cases, the two sides or different parts of a street segment may be located in different OAs. In such cases, the street segment was assigned to that within which the largest proportion of its total length lay. The second unit, the medium super output area (MSOA), is another UK census unit that is defined by the aggregation of OAs, incorporating approximately 25 on average. There are 131 of these in Birmingham. The particular HLM we use is therefore a three-level variant.

In selecting a suitable dependent variable for the analysis of ‘count’ data such as is considered here, where there is a requirement to control for opportunity, there are three well-motivated choices. These are: to calculate the rate for each segment and model that directly; to model the count data and include the opportunity variable as an explanatory variable, but constrained to have a certain coefficient (this is known as an ‘offset’ term); or to model the counts with the opportunity variable as an orthodox explanatory variable. We choose the latter approach here, with the reasoning being that potential targets are not sufficiently isolated from each other as to justify their treatment as entirely independent opportunities; that is, there is likely to be some non-linearity in the relationship between household count and opportunity. For example, there is a limit to how many times a certain number of houses can plausibly be victimized. We therefore model the total count of burglary observed across the 4 years considered at the street segment level. A standard choice when dealing with count data concerning relatively rare events such as these is to model them as a Poisson distribution, and this is the approach taken here.

### *Street Segment Independent Variables*

The explanatory variables included in the model comprise street network metrics, simple indicators of opportunity and a limited number of socio-demographic descriptors. Betweenness and linearity—the two network properties of primary interest—are included at the lowest level of the model: the street segment. As referred to previously, we first take the natural logarithm of betweenness before including it in the model, in order to moderate the long-tailed nature of its distribution. Furthermore, values in each case are normalized to the range  $[0, 1]$  by dividing by the maximum value, to allow meaningful comparison between results for different definitions of betweenness (topological/metric etc.).

Two further variables are included at street segment level. The number of address points associated with each segment is included as a simple indicator of opportunity: all residential dwellings, including individual flats within blocks, are counted, and this quantity therefore represents a measure of the abundance of possible targets. In addition, the spatial density of address points is also considered. This is included as previous research has suggested that street segments with higher target density experience a lower relative risk of burglary (Johnson and Bowers 2010), although the effect is notably minor. This density variable was computed by dividing the number of address points located on a street segment by the segment length (in meters).

### *Output Area (OA) Independent Variables*

It is also necessary to control at the OA level for other possible causes of variation in the risk of burglary. To this end, while these are not the focus of the current study, we include a number of statistics from the 2001 UK census which have been shown to be associated with burglary risk in previous empirical work. In line with research concerned with the

peak age of offending (e.g. Farrington 1983), the percentage of residents between the ages of 10 and 15 was included.

As discussed, theories of social disorganization suggest that crime is higher in disorganized neighborhoods (Sampson and Raudenbush 1999), and empirical research supports this. Ethnic heterogeneity, measured using a standard index (Simpson 1949; Blau 1977), is commonly used as a measure of social disorganization, and has been found to be a significant predictor of burglary risk in previous work (e.g. Hirschfield and Bowers 1997). It is thus included here. Unemployment (e.g. Johnson and Bowers 2010) and the prevalence of vacant housing (e.g. Spelman 1993) are further related social factors thought to have an effect on burglary risk, and are hence also included.

The percentage of students in the neighborhood population has also been found to be positively associated with burglary risk (see Tilley et al. 1999) and is consequently included in the model. This is particularly relevant in the case of Birmingham, since it is a city with several identifiable student districts. To estimate the influence of those factors that are unobserved in our model but that vary at the OA level, a random effects intercept term is included in the model. No socio-demographic variables are included at MSOA level; instead, any effects at this spatial scale are estimated using a random effects term.

The structure of the HLM can be written down in relatively simple mathematical terms. An indexing system is constructed for the various spatial units:  $i$  for street segments,  $j$  for OAs and  $k$  for MSOAs. The model is then fully described by

$$\pi_{ijk} = \exp(\beta_0 + u_{jk} + v_k + \beta_1 x_{1ijk} + \dots + \beta_m x_{mijk} + \beta_{m+1} x_{(m+1)jk} + \dots + \beta_n x_{njk}) \quad (5)$$

$$u_{jk} \sim N(0, \Omega_u) \quad (6)$$

$$v_k \sim N(0, \Omega_v) \quad (7)$$

where  $\pi_{ijk}$  is the burglary count on segment  $i$  (in OA  $j$ , in MSOA  $k$ ),  $x_1, \dots, x_m$  are the explanatory variables defined at segment level, and  $x_{m+1}, \dots, x_n$  are those defined at OA level. The terms  $u_{jk}$  and  $v_k$  are the ‘random intercepts’ at OA and MSOA level respectively, both normally distributed. Before estimation, the independent variables were tested for evidence of multicollinearity; correlation coefficients were all found to be in the range [0.08, 0.35] and so multicollinearity was not considered a threat to statistical inference.

### Findings

Regression coefficients for six HLMs are given in Table 1. These differ in only the precise definition of betweenness used in the calculations. As described previously (see Eq. 3), an upper limit can be placed on the length of the network paths considered in the calculation of betweenness; similarly, this distance can be defined in either topological or metric terms. The various definitions can be interpreted as representing the concept of betweenness at different scales, and six alternatives are presented here for comparison.

The results as a whole provide persuasive evidence that street network properties, in particular betweenness, are a highly significant predictor of burglary. In almost all cases, higher betweenness is associated with greater victimization (Hypothesis 2), and the associated  $z$ -scores are particularly high. The one anomalous case occurs when between-

**Table 1** Regression coefficients for HLMs of street segment burglary risk. Estimated coefficients for several HLMs, where the various columns correspond to different definitions of betweenness (in terms of the radius used in its calculation and whether distance is defined in topological or metric terms)

Segment level	Topological				Metric			
	5	50	150	500	3,000	7,500		
Address count	1.02* (146.06)	1.02* (139.88)	1.02* (141.41)	1.03* (139.83)	1.03* (158.50)	1.03* (149.12)		
Addresses per 100 m	1.00* (4.06)	1.00* (8.16)	1.00* (9.95)	1.00* (16.75)	1.00* (13.37)	1.00* (13.63)		
Linearity	0.52* (10.65)	0.57* (9.18)	0.62* (7.76)	1.02 (0.29)	0.61* (7.95)	0.62* (7.80)		
Betweenness value	4.65* (22.28)	3.86* (18.60)	3.14* (16.09)	1.00 (0.10)	2.25* (16.24)	2.27* (16.27)		
OA level								
Ethnic heterogeneity	1.03* (2.85)	1.03* (2.59)	1.03* (2.83)	1.03* (2.99)	1.03* (2.84)	1.03* (2.70)		
Aged 10–15 (%)	1.02* (4.04)	1.02* (4.31)	1.02* (4.38)	1.02* (4.07)	1.02* (4.42)	1.02* (4.50)		
Unemployed (%)	1.01* (2.35)	1.01* (2.57)	1.01* (2.56)	1.01* (2.34)	1.01* (2.29)	1.01* (2.38)		
Students (%)	1.01* (4.90)	1.01* (4.97)	1.01* (4.95)	1.01* (5.12)	1.01* (5.16)	1.01* (5.04)		
Houses vacant (%)	1.00 (0.25)	1.00 (0.33)	1.00 (0.33)	1.00 (0.55)	1.00 (0.28)	1.00 (0.27)		
Random intercepts								
OA random variance	0.26* (24.83)	0.25* (24.76)	0.25* (24.76)	0.26* (24.92)	0.26* (24.84)	0.26* (24.85)		
MSEA random variance	0.09* (6.82)	0.09* (6.80)	0.09* (6.82)	0.09* (6.81)	0.09* (6.79)	0.09* (6.78)		

The values given are exponentiated regression coefficients, and therefore represent the factor by which the mean count on a segment is estimated to change as a result of a one-unit change in the explanatory variable. Z-scores are given in brackets, and \* denotes significance at  $p = 0.05$  level

ness is calculated on the basis of metric distance, with a relatively short radius of 500 m, for which the relationship is non-significant. Although only a selection are presented here, models were run for a wide range of radii, with the results consistent in all but that case. One possible explanation for the behaviour at that radius may be that the effects it measures are simply too local to discriminate sufficiently between segment types in the system as a whole.

Although once again there is an anomalous result in the case of 500m radius, there is also a consistent relationship between the street segment burglary count and linearity across all other models (Hypothesis 3). The direction of this highly significant relationship—that more linear roads experience lower victimization—is in line with Hypothesis 3, but unexpected given the results shown in Fig. 8. One possible explanation for this concerns address point counts and their variation with network properties: in the data as a whole, there is a highly significant inverse relationship between linearity and address point count (so that more sinuous streets have higher address counts). As it happens, the positive effect of address count on incident count predicted by the model is sub-linear at the scale we are dealing with, and so the lower crime rate on sinuous streets apparent in Fig. 8 may simply be an artefact of their higher address count (since the calculation of rate is predicated upon an assumption of equality of opportunity).

With regard to the socio-demographic variables included in the model, with the exception of the proportion of vacant houses, all variables included appear to be significantly positively associated with burglary, as expected. The most striking relationships appear to be those with the size of student population and the proportion of residents aged 10–15 in an area, both of which have been found in previous work and which accord with what is known about the particular circumstance of Birmingham. Speaking of area-level effects more widely, the decision to adopt the multi-level structure in the model appears to be justified on empirical (as well as theoretical) grounds, since the parameter estimates for the random intercepts are clearly non-zero. These also, of course, indicate that a considerable amount of variation in burglary counts cannot be accounted for by the variables considered here.

## Discussion

That street networks should play a role in shaping patterns of crime is an immediate corollary of various theories of environmental criminology (Brantingham and Brantingham 1981), and the form of this effect has been the subject of much conjecture. Though rigorous empirical studies on the topic have been carried out, reaching (largely) consistent conclusions, they are limited to a certain extent by the level of refinement possible in the analysis of the network itself, and many studies use small sample sizes (for a review, see Johnson and Bowers 2010). The work presented here uses analytical techniques from network science, not previously used in this domain, to formalise several of the ideas previously proposed and to relate them more explicitly to basic theory. The results confirm the existence of a relationship between street network properties and burglary, and suggest that the use of network metrics as predictors of crime has considerable potential.

A notable result from the statistical model presented here is the highly significant relationship between burglary and betweenness centrality. Bearing in mind how betweenness is calculated, this suggests that those street segments which have the potential to be more frequently used in movement around the network will be at higher risk, and the result is robust to almost all changes to the definition of which paths are used in its

calculation. Insofar as betweenness broadly corresponds to street classification and, more obliquely, to the number of connections of a segment, this result agrees with previous empirical work, and is as would be expected from theory. Most obviously, it provides evidence in support of crime pattern theory and its emphasis on the role of the urban backcloth in crime clustering.

The use of betweenness, however, represents a significant improvement upon the street measurements used in previous work. Unlike an ordinal classification, it is a direct and objective estimate of the usage of streets and therefore has immediate correspondence to the consideration of urban awareness spaces. Secondly, it is a highly-granular measurement and offers a level of detail significantly beyond previous approaches (although, of course, care must be taken that detail and explanatory value are not conflated). Both of these points are demonstrated clearly by the analysis of betweenness across road types shown in Fig. 4: the large intra-class variation and significant overlaps between classes imply that they carry relatively little information and define a far-from-consistent hierarchy in material terms.

Several realizations of our statistical model were produced, with the way in which betweenness was calculated varying between cases. This variation took two forms: the maximum length of trips considered in the analysis, and whether this length was calculated in topological or metric terms. When metric distance was used, with a radius of 500 m, the results were non-significant, suggesting that any extremely local effects which were present were accounted for by other factors; however, in all other cases, betweenness was found to be a highly significant predictor. In the case of topological distance, however, the significance of betweenness as a predictor of risk was consistent across all radii. One possible explanation for the minor discrepancy between the results for the two methods of measuring distance might be that, as proposed by Hillier et al. (2010), pedestrians are more likely to select less complex routes, with fewer turns, when navigating around an urban area: these routes would have the lowest topological distance, but not necessarily be the shortest in metric terms.

Comparing the results for the three models which use topological distance, the highest z-scores were found using low radii, which implies that relatively local travel is of greatest significance in the context of burglary risk. Journeys of this length are likely to be pedestrian, and the results therefore suggest that this mode of travel plays a substantial role in offending. It is important, however, that these findings are interpreted in the context of the mechanisms by which street network effects are hypothesised to act: the role of these journeys is in the accumulation of awareness space, rather than in the commission of the offence itself. Similarly, the fact that the significance of betweenness is persistent across a variety of radii suggests that travel patterns across a range of scales are of relevance. Further exploration of these issues represents an interesting avenue for future research.

Concerning the case of linearity, the results show consistently that, when measured in this way, more sinuous street segments are at higher risk of burglary after controlling for other variables. This contradicts some previous empirical work (which was limited to cul-de-sacs), but can still be reconciled with existing theory. One possible inference may be that the 'cover' afforded by sinuous streets is favourable for crime and that, equivalently, the visual guardianship on more linear streets is indeed a significant impediment. This accords with the findings of qualitative interviews with burglars (e.g. Wright and Decker 1994), in which the desire to evade detection is frequently cited as a key factor in target choice. Caution should, however, be exercised when interpreting these results and determining the suitability of linearity as a metric. Defined as it is, it does not necessarily correspond to a street's visual properties as, perhaps, an angular measurement would, and it is also measured for segments in isolation, taking no account of visibility from connecting

segments. Coupled with the fact that linearity also appears to be associated with higher betweenness and lower address counts, the picture in this case is not necessarily clear and is certainly worthy of further investigation.

There are several ways in which the relationships observed here might be employed in a practical context. The first of these concerns lessons which might be learned from the perspective of planning and urban design. The measured characteristics of street networks are clearly a result of how they are constructed and can be made to take different values by modifying design principles: the extreme examples of grid-like systems (where all streets have similar properties) and tree-like structures (which are hierarchical) are convenient illustrations of this. The implications of the apparent association of betweenness and burglary, though, depend crucially on whether the additional crime on more central streets would otherwise have taken place elsewhere. To simplify, if we assume all else to be equal, it may be the case that the total volume of crime is constant and that the effect of inter-segment variation is simply to create inequality in its distribution. In this case, where the effect of a more homogeneous structure would be to distribute the same volume of crime more evenly, such a design would be of questionable value (indeed, since it removes the possibility of targeted policing, it may be disadvantageous).

On the other hand, if the crimes occurring on more central streets take place, in some sense, *because* the street is more central, then changes in street network structure may well have the potential to reduce crime. The issue is fundamentally one of displacement, or rather placement (see Barr and Pease 1990): assuming that crime does not move elsewhere, our results concerning betweenness suggest that reducing the concentration of activity on certain streets should lead to an overall reduction in risk. Of course, to modify existing networks is, in general, unrealistic, so the primary relevance of this finding is in the design of new urban areas. Hypotheses concerning such design principles would be difficult to test, since any comparison between urban areas with differing structure would be compromised to a certain extent by many other variable factors. Nevertheless, the use of road closure as a crime prevention tactic is not without precedent (see Clarke 2004), and may provide a means of evaluating the impact of possible design changes. Since even small structural modifications can cause large perturbations in network metrics, such an experiment would represent a direct test of the effect of varying usage levels and permeability. Though not concerned explicitly with network structure, two previous studies of road closure, carried out by Matthews (1993, 1997) using a quasi-experimental design, are particularly notable for having brought about an apparent reduction in burglary, despite being focussed primarily on prostitution. A further, and complementary, possibility, would be to explore these issues further via agent-based modelling (see Groff and Mazerolle 2008). This represents a means by which the consequences of potential changes to network structure could be explored quantitatively, at low marginal cost, prior to their deployment in the real world.

In terms of practical value at somewhat shorter time-scales, the results here suggest that street network metrics can, with some confidence, be added to the array of variables used by police in directing general policing effort and targeting specific interventions. In fact, they may be of particular value since they are an objective physical quantity which can be measured without consideration of any local context. As such, they are not subject to any of the inherent unreliability or ethical bias associated with other predictors, such as socio-demographic factors, and can be calculated in real-time, being updated as changes—temporary or otherwise—occur.

The above point is particularly relevant in the wider context of recent efforts in the mathematical modelling of burglary (Short et al. 2008; Pitcher 2010; Berestycki and Nadal

2010). One limitation of these models (and of predictive methods more generally) is their failure to take into account a realistic representation of urban form. The analysis presented here both emphasizes the need to do this and suggests that graph theory offers a route by which it might be achieved mathematically.

Significant scope for further work in this area exists, given the apparent potential of network analysis techniques. Many metrics beyond those considered here have been proposed in the network analysis literature, and further work is required to determine a wider toolkit appropriate to the crime domain. Indeed, betweenness itself is a relatively naïve measure of travel patterns, and could be refined by incorporating additional aspects of urban form in its calculation. For example, journeys to and from known centers of activity (e.g. dense residential areas, entertainment districts) could be assigned higher weighting, and the selection of routes could take into account factors other than absolute distance, such as travel speed and likely congestion. In doing this, however, care must be taken not to forego the parsimony and universality of the basic definition. It will also be interesting to explore the interaction between network effects and social context: while social influences were beyond the scope of the present analysis, it may be the case that the effect of physical structure is dependent upon social aspects of the environment.

With a significant volume of empirical evidence now suggesting a relationship between network effects and the spatial clustering of crime, it will also be interesting to consider their effect on other known phenomena, such as repeat (Pease 1998) and near repeat victimization (e.g. Johnson et al. 2007), whereby crime is seen to cluster in both space and time. The activity patterns on which much of the preceding theory is based will, of course, vary temporally, and therefore the effect of the network might be expected to differ by time of day. In the case of near-repeat victimization, theories which involve a same-offender hypothesis (e.g. Johnson et al. 2009) typically invoke ideas of awareness space (perhaps in the form of escape routes) in order to explain why nearby properties are at higher risk. The street network will naturally be expected to influence this.

In conclusion, relationships have been found between network properties and burglary levels which relate directly to ideas of urban permeability and movement, and which have potential to influence crime reduction policy. The general framework introduced suggests a fertile area of research by which the effect of urban form in a criminal context can be better quantitatively understood.

**Acknowledgments** The authors wish to thank West Midlands Police, and Chief Superintendent Alex Murray in particular, for the provision of the recorded crime data analysed in this paper. We would also like to thank David Weisburd and the three anonymous reviewers for their insightful comments, which helped to improve the paper. We also acknowledge the financial support of the Engineering and Physical Sciences Research Council (EPSRC) under the Security Science Doctoral Training Centre, reference EP/G037264/1.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Andresen MA (2014) Environmental criminology: evolution, theory, and practice. Routledge, Abingdon
- Andresen MA, Malleon N (2011) Testing the stability of crime patterns: implications for theory and policy. *J Res Crime Delinq* 48(1):58–82



- Armitage R (2007) Sustainability versus safety: confusion, conflict and contradiction in designing out crime. In: Farrell G, Bowers KJ, Johnson SD, Townsley M (eds) *Imagination for crime prevention: essays in honour of Ken Pease*. Willan, Cullompton, pp 81–110
- Ashton J, Brown I, Senior B, Pease K (1998) Repeat victimisation: offender accounts. *Int J Risk Secur Crime Prev* 3:269–280
- Barr R, Pease K (1990) Crime placement, displacement, and deflection. *Crime Justice* 12:277–318
- Barthélemy M (2011) Spatial networks. *Phys Rep* 499(1–3):1–101
- Beavon D, Brantingham P, Brantingham P (1994) The influence of street networks on the patterning of property offences. In: Clarke R (ed) *Crime prevention studies*, vol 2. Criminal Justice Press, Monsey, pp 115–148
- Berestycki H, Nadal J-P (2010) Self-organised critical hot spots of criminal activity. *Eur J Appl Math* 21(Special Double Issue 4–5):371–399
- Bernasco W, Block R (2011) Robberies in Chicago: a block-level analysis of the influence of crime generators, crime attractors, and offender anchor points. *J Res Crime Delinq* 48(1):33–57
- Bevis C, Nutter J (1977) Changing street layouts to reduce residential burglary. Minnesota Crime Prevention Center, Minnesota
- Blau PM (1977) *Inequality and heterogeneity: a primitive theory of social structure*. Free Press, New York
- Block RL, Block CR (1995) Space, place, and crime: hot spot areas and hot places of liquor-related crime. *Crime Prev Stud* 4:145–184
- Bollobás B (2002) *Modern graph theory*, 2nd edn. Springer, New York
- Brantingham PJ, Brantingham PL (1981) *Environmental criminology*. Sage, Thousand Oaks
- Brantingham PL, Brantingham PJ (1993a) Environment, routine and situation: toward a pattern theory of crime. *Adv Criminol Theory* 5:259–294
- Brantingham PL, Brantingham PJ (1995) Criminality of place. *Eur J Crim Policy Res* 3(3):5–26
- Brantingham PL, Brantingham PJ (1975) Residential burglary and urban form. *Urban Stud* 12(3):273–284
- Brantingham PL, Brantingham PJ (1993b) Nodes, paths and edges: considerations on the complexity of crime and the physical environment. *J Environ Psychol* 13(1):3–28
- Brantingham PL, Brantingham PJ, Vajihollahi M, Wuschke K (2009) Crime analysis at multiple scales of aggregation: a topological approach. In: Weisburd D, Bernasco W, Bruinsma GJ (eds) *Putting crime in its place*. Springer, New York, pp 87–107
- Chan SHY, Donner RV, Lämmer S (2011) Urban road networks—spatial networks with universal geometric features? *Eur Phys J B* 84(4):563–577
- Clarke RVG (2004) Closing streets and alleys to reduce crime: should you go down this road?. US Department of Justice, Office of Community Oriented Policing Services, Washington
- Cohen LE, Felson M (1979) Social change and crime rate trends: a routine activity approach. *Am Sociol Rev* 44(4):588–608
- Cromwell PF, Olson JN, Avary DW (1991) *Breaking and entering: an ethnographic analysis of burglary*. Sage, Thousand Oaks
- Crucitti P, Latora V, Porta S (2006) Centrality measures in spatial networks of urban streets. *Phys Rev E* 73(3):036125
- Farrington DP (1983) Offending from 10 to 25 years of age. In: Van Dusen KT, Mednick SA (eds) *Prospect Stud Crime Delinq*. Kluwer-Nijhoff, Dordrecht, pp 17–37
- Felson M (1987) Routine activities and crime prevention in the developing metropolis. *Criminology* 25(4):911–932
- Freeman LC (1977) A set of measures of centrality based on betweenness. *Sociometry* 40(1):35–41
- Gini C (1921) Measurement of inequality of incomes. *Econ J* 31(121):124–126
- Groff ER (2007) ‘Situating’ simulation to model human spatio-temporal interactions: an example using crime events. *Trans GIS* 11(4):507–530
- Groff E, Mazerolle L (2008) Simulated experiments and their potential role in criminology and criminal justice. *J Exp Criminol* 4(3):187–193
- Hillier B (1996) *Space is the machine*. Cambridge University Press, Cambridge
- Hillier B (1997) Cities as movement economies. In: Droege P (ed) *Intelligent environments: spatial aspects of the information revolution*. Elsevier, North Holland
- Hillier B (2004) Can streets be made safe? *Urban Des Int* 9(1):31–45
- Hillier B, Iida S (2005) Network and psychological effects in urban movement. In: Cohn AG, Mark DM (eds) *Spatial information theory*, vol 3693., Lecture notes in computer science. Springer, Berlin, pp 475–490
- Hillier B, Shu S (2000) Crime and urban layout: the need for evidence. In: Ballantyne S, MacLaren V, Pease K (eds) *Secure foundations: key issues in crime prevention, crime reduction and community safety*. University College London Institute for Public Policy Research, London, pp 224–248

- Hillier B, Turner A, Yang T, Park H-T (2010) Metric and topo-geometric properties of urban street networks: some convergences, divergence and new results. *J. Space Syntax* 1(2):258–279
- Hirschfield A, Bowers KJ (1997) The effect of social cohesion on levels of recorded crime in disadvantaged areas. *Urban Stud* 34(8):1275–1295
- Iwanski N, Frank R, Reid A, Dabbaghian V (2012) A computational model for predicting the location of crime attractors on a road network. In: Memon N, Zeng D (eds) Proceedings of the 2012 European intelligence and security informatics conference (EISIC). IEEE, Los Alamitos, pp 60–67
- Jacobs J (1961) The life and death of Great American cities. Random House, New York
- Jiang B (2007) A topological pattern of urban street networks: universality and peculiarity. *Phys A* 384(2):647–655
- Johnson SD (2010) A brief history of the analysis of crime concentration. *Eur J Appl Math* 21(Special Double Issue 4–5):349–370
- 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(3):201–219
- Johnson SD, Bowers KJ (2010) Permeability and burglary risk: are cul-de-sacs safer? *J Quant Criminol* 26(1):89–111
- Johnson SD, Summers L, Pease K (2009) Offender as forager? A direct test of the boost account of victimization. *J Quant Criminol* 25:181–200
- Masucci AP, Smith D, Crooks A, Batty M (2009) Random planar graphs and the London street network. *Eur Phys J B* 71(2):259–271
- Matthews R (1993) Kerb-crawling, prostitution and multi-agency policing. Crime prevention unit series paper no 43. Home Office, London. [http://www.popcenter.org/problems/street\\_prostitution/PDFs/fcpu43.pdf](http://www.popcenter.org/problems/street_prostitution/PDFs/fcpu43.pdf)
- Matthews R (1997) Developing more effective strategies for curbing prostitution. In: Clarke RV (ed) Situational crime prevention: successful case studies, 2nd edn. Criminal Justice Press, Monsey
- Mawby RI (1977) Defensible space: a theoretical and empirical appraisal. *Urban Stud* 14(2):169–179
- McCord ES, Ratcliffe JH (2007) A micro-spatial analysis of the demographic and criminogenic environment of drug markets in Philadelphia. *Aust N Z J Criminol* 40(1):43–63
- Newman M (2010) Networks: an introduction. Oxford University Press, Oxford
- Newman O (1972) Defensible space; crime prevention through urban design. Macmillan, New York
- North BV, Curtis D, Sham PC (2002) A note on the calculation of empirical p values from Monte Carlo procedures. *Am J Hum Genet* 71(2):439–441
- Pease K (1998) Repeat victimisation: taking stock. Home Office Police Research Group, London
- Pitcher AB (2010) Adding police to a mathematical model of burglary. *Eur J Appl Math* 21(Special Double Issue 4–5):401–419
- Porta S, Crucitti P, Latora V (2006a) The network analysis of urban streets: a dual approach. *Phys A* 369(2):853–866
- Porta S, Crucitti P, Latora V (2006b) The network analysis of urban streets: a primal approach. *Environ Plan B* 33(5):705–725
- Porta S, Latora V, Wang F, Rueda S, Strano E, Scellato S, Cardillo A, Belli E, Cárdenas F, Cormenzana B, Latora L (2012) Street centrality and the location of economic activities in Barcelona. *Urban Stud* 49(7):1471–1488
- Porta S, Latora V, Wang F, Strano E, Cardillo A, Scellato S, Iacoviello V, Messori R (2009) Street centrality and densities of retail and services in Bologna, Italy. *Environ Plan B* 36(3):450–465
- Reynald DM (2010) Guardians on guardianship: factors affecting the willingness to supervise, the ability to detect potential offenders, and the willingness to intervene. *J Res Crime Delinq* 47(3):358–390
- Reynald DM, Elffers H (2009) The future of Newman's defensible space theory: linking defensible space and the routine activities of place. *Eur J Criminol* 6(1):25–46
- Robinson W (1950) Ecological correlations and the behavior of individuals. *Am Sociol Rev* 15(3):351–357
- Sampson RJ, Raudenbush SW, Earls F (1997) Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science* 277(5328):918–924
- Sampson RJ, Raudenbush SW (1999) Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. *Am J Sociol* 105(3):603–651
- Short M, D'Orsogna M, Pasour V, Tita G, Brantingham P, Bertozzi A, Chayes L (2008) A statistical model of criminal behavior. *Math Models Methods Appl Sci* 18(S1):1249–1267
- Simpson EH (1949) Measurement of diversity. *Nature* 163:688
- Spelman W (1993) Abandoned buildings: magnets for crime? *J Crim Justice* 21(5):481–495
- Strano E, Viana M, da Fontoura Costa L, Cardillo A, Porta S, Latora V (2013) Urban street networks, a comparative analysis of ten European cities. *Environ Plan B* 40(6):1071–1086

- Tilley N, Pease K, Hough M, Brown R (1999) Burglary prevention: early lessons from the crime reduction programme. Home Office Policing and Reducing Crime Unit, London
- Weisburd D, Bushway S, Lum C, Yang S-M (2004) Trajectories of crime at places: a longitudinal study of street segments in the city of Seattle. *Criminology* 42(2):283–322
- Weisburd D, Groff ER, Yang S-M (2012) *The criminology of place*. Oxford University Press, New York
- White GF (1990) Neighborhood permeability and burglary rates. *Justice Q* 7(1):57–67
- Wright RT, Decker SH (1994) *Burglars on the Job*. Northeastern University Press, Boston