

# The ecology of outdoor rape: The case of Stockholm, Sweden

**Vania Ceccato**

Royal Institute of Technology, Sweden

**Guangquan Li**

Northumbria University, UK

**Robert Haining**

University of Cambridge, UK

## Abstract

The objective of this article is to report the results of an ecological study into the geography of rape in Stockholm, Sweden, using small area data. In order to test the importance of factors indicating opportunity, accessibility and anonymity to the understanding of the geography of rape, a two-stage modelling approach is implemented. First, the overall risk factors associated with the occurrence of rape are identified using a standard Poisson regression, then a local analysis using profile regression is performed. Findings from the whole-map analysis show that accessibility, opportunity and anonymity are all, to different degrees, important in explaining the overall geography of rape - examples of these risk factors are the presence of subway stations or whether a *basområde* is close to the city centre. The local analysis reveals two groupings of high risk of rape areas associated with a variety of risk factors: *city centre areas* with a concentration of alcohol outlets, high residential population turnover and high counts of robbery; and *poor suburban areas* with schools and large female residential populations where subway stations are located and where people express a high fear of crime. The article concludes by reflecting upon the importance of these results for future research as well as indicating the implications of these results for policy.

## Keywords

Geographical Information Systems (GIS), profile regression, public places, rape, sexual violence

---

### Corresponding author:

Vania Ceccato, Royal Institute of Technology, Drottning Kristinas väg 30, Stockholm, 10044, Sweden.

Email: [vania.ceccato@abe.kth.se](mailto:vania.ceccato@abe.kth.se)

## Introduction

In 2013, the Swedish newspaper *Dagens Nyheter* reported the case of a young woman raped in one of the most popular public gardens in the centre of Stockholm. In 2016, another young woman was sexually attacked, but in that incident the assault happened on the city's outskirts (*Dagens Nyheter*, 2013, 2016). About two-thirds of rape cases reported to the police in Sweden happen indoors, committed by someone the woman knows; only about 30 percent take place outdoors, in public places. But outdoor rape is a distinctive form of this crime. First, in the case of outdoor rape, the rapist is usually someone unknown to the woman. Second, this type of sexual violence often takes place in interstitial public spaces such as back streets and hidden places and its occurrence depends, in part at least, on local environmental circumstances (Ceccato, 2014, and Ceccato et al., 2017). Third, outdoor rape has an impact on the community and may affect local levels of fear of crime. Fourth, unlike domestic rape, improvements in city planning and changes in policing methods might prevent, or at least greatly reduce, the numbers of cases of outdoor rape.

What characterizes the environments where rape takes place? Might a better understanding of the kinds of places where rape occurs be helpful in preventing it? We argue that place matters for our understanding of where the crime of rape happens in two senses. First, place matters in the sense of a specific, highly localized, micro-scale setting such as the existence of a secluded area in a park or badly lit alleyway where the criminal act takes place. Second, place matters in the sense of a more broadly defined meso-scale environment or geographical context such as a neighbourhood or small area of a town or city within which the crime is situated. In the past, where an attack happened was regarded as a marginal aspect of the rape act (see, for example, Belknap, 1987), but more recent work has shown evidence of the importance of these settings (for a review, see Beauregard et al., 2005, and Hewitt and Beauregard, 2014).

This article reports the results of modelling police-recorded rape cases in Stockholm, Sweden, by small areas (*basområdes*), complementing earlier work that focused on the micro-places where individual rape incidents took place (Ceccato, 2014). These small areas have, on average, a population of 2200 individuals. In focusing on what we will term the meso-scale of analysis, the purpose is to emphasize the ecological characteristics of places, their land use, demographic composition and socio-economic characteristics, and their association with varying numbers of cases. Ecological models assess the statistical significance and substantive importance of different small area characteristics to our understanding of the geography of crime. However, in undertaking analysis at this scale and particularly with a predatory crime such as rape that involves very direct and immediate contact between offender and target, it is also necessary to recognize that small area (meso-scale) counts will be influenced by how many suitable 'micro-scale spaces' exist within each area. These are areas offering adequate opportunity as well as anonymity to the attacker. This presents a challenge for ecological modelling but one that is not unique to the study of the crime of rape (see, for example, Sherman et al., 2006, and Weisburd et al., 2012).

Data analysis involves two phases of modelling (Li et al., n.d.). First, a hierarchical regression model is fitted, the estimates from which show how each risk factor modifies

the ‘whole map’ or ‘global’ crime risk with all other factors held constant. The second phase of analysis uses profile regression (Molitor et al., 2010) to reveal the association with ‘local’ ecological characteristics. The combination of these two modelling techniques offers a better understanding of risk factor effects, by estimating the *whole map* effects associated with each individual risk factor and also identifying how these risk factors impact *at the local level*.

Both modelling phases adopt a Bayesian methodology. Bayesian hierarchical regression modelling separates the regression model into three linked models: a data model, a process model and one or more parameter models.<sup>1</sup> Regression modelling applied to spatial data raises a number of particularly complex challenges, including the presence of spatial dependence in data and spatial heterogeneity (for example, parameters varying spatially). Because spatial units can be small and cases of rape relatively uncommon, the data are often sparse (small numbers of cases per spatial unit; often many spatial units with no cases). Finally, uncertainty is endemic – uncertainty in respect of the data (for example, unreported cases, the small number problem), the model (for example, contested theories or explanations of where rapes are most likely to occur) and the model’s parameters (for example, the unknown regression coefficients). Such complexity requires a flexible modelling framework that enables us to address all these challenges. Hierarchical modelling provides flexibility and the Bayesian approach to hierarchical modelling, unlike the frequentist approach, allows us to jointly incorporate uncertainty *at all three model levels*. Spiegelhalter et al. (2004) argue that, compared with frequentist methods, the Bayesian approach is more flexible in adapting to each unique situation, more efficient in using all the available evidence and more useful in providing relevant quantitative summaries.

Profile regression not only produces a classification of areas based on their risk factor (covariate) values but also uses regression to associate these groups to the risk of rape. Other clustering methods classify areas based only on the former. The Bayesian aspect of this clustering process has some further advantages over traditional clustering approaches: it allows the number of groups to vary; it reveals groups (of areas) and examines the risk factor pattern and its association with the risk of rape within each group; the risk of rape in an area, in addition to its covariate values, influences the assignment of the area to the groups.

Bayesian inference is more relevant than frequentist inference to observational (non-experimental) science. Bayesian inference updates our prior beliefs (however vague or non-informative they might be) given a new set of data. Frequentist inference, by contrast, is founded on the idea of ‘infinite hypothetical replications of the situation under consideration’ (Wakefield, 2013: 23).

## **Literature review and hypotheses**

### *The definition of outdoor rape*

The legal definition of rape has varied over time, by country and even within national borders. In the United States, the Justice Department defines rape as ‘forced sexual intercourse, including vaginal, anal, or oral penetration. Penetration may be by a body part or

an object. Sexual assault is unwanted sexual contact that stops short of rape or attempted rape. This includes sexual touching and fondling' (RAINN, 2014).

In European countries, rape may involve a wide spectrum of behaviours (for a comparative country review, see Lovett and Kelly, 2009). In Sweden, marital rape was criminalized in 1965 (in the 1970s in the USA and in the 1980s in the UK). The definition of rape was extended to include any sexual act committed by a stranger, an acquaintance, a partner or a family member of the person. Rape is not confined to the penetrative act, but comprises 'all sexual behaviour comparable in nature to violation having regard for the prevailing conditions. Having intercourse with a person who is unconscious, sleeping, intoxicated, handicapped or in a similarly helpless state shall be regarded as equivalent to rape by threat or violence' (Ministry of Justice, 1999: Ch. 6 §1). The maximum penalty is 6 years in prison, or 10 years in the case of 'gross rape' determined by the level of violence involved and whether more than one person assaulted the individual. In 1984, the definition of rape became gender-neutral, incorporating both male rape and assaults by women against men. The Swedish definition of rape is close to what some states in the USA (for example, the state of New York) define as 'predatory sexual assault'. A person is guilty of predatory sexual assault when he or she commits the crime of rape, a criminal sexual act or aggravated sexual abuse. In this case, rape is not confined to the penetrative act as previously mentioned and may also have a premeditated dimension to it (DCJS, 2014; NYSCASA, 2007).

'Rape' in this article refers to all police-recorded cases of complete (sexual intercourse) and incomplete rape (attempted sexual intercourse and other types of sexual behaviour as defined by law) that happen outdoors (such as, for example, in streets, parks or transportation settings). In most cases, the offender was someone who was unknown to the victim. See Table 2 in Appendix A for crime codes and for details see BRÅ (2012).

### *An overview of socio-spatial research into the geography of rape*

In common with other ecological studies into crime patterns (see, for example, Bernasco and Luykx, 2003), the starting point for our analysis is to conceptualize where an act of rape takes place as the outcome of a two-stage rational choice process on the part of the offender. At the first stage, although the act may be primarily impulsive or opportunistic, we start from the position that the rapist will be rational to the extent that they will be willing to take the risk only in certain areas or neighbourhoods. At the second stage, the offender selects a specific target within the chosen neighbourhood. Rationality at this second stage involves assessing perceived target vulnerability and whether the time and location offer an acceptable level of risk to the offender.

At the first stage, neighbourhoods are chosen from within the offender's activity field (Brantingham and Brantingham, 1984) or awareness space, which from experience or instinct offer opportunity with a low 'subjective probability of detection' in committing rape (Hechter and Kanazawa, 1997: 201). The activity field or awareness space of an individual comprises their anchor points in the urban landscape, such as place of residence, place of work, public meeting places including bars and frequently used subway stations and the lines of travel connecting these anchor points. Most importantly, because the offender knows these areas they will be familiar with escape routes from the crime

scene. *Accessibility* in the sense of getting to, and knowing, the crime scene and being able to flee it quickly forms an essential part of the decision process. Previous research into the geography of serial rape has found that offenders tend to use the same geographical space repeatedly (for example, Lebeau, 1985; Pyle, 1974; Rhodes and Conly, 1981).

Other area characteristics are also important. Of necessity, the neighbourhood must offer *opportunities* for rape, that is the presence of women, perceived as vulnerable, who either live in the neighbourhood or travel through it on their way home or to work or leisure. Recent research shows that place factors affect the offender's behaviour more than offence timing (Hewitt and Beauregard, 2014). Finally, the neighbourhood must provide sufficient *anonymity* for the criminal act, as we now discuss.

Acts of rape tend to occur on isolated streets, in interstitial spaces, parks, fields, places with footpaths with many sites for concealment and with building and street layouts with poor lighting and offering poor natural surveillance (Belknap, 1987; Canter and Larkin, 1993; Rhodes and Conly, 1981; Rosay and Langworthy, 2003; Pyle, 1974). Such spaces provide *micro-scale* anonymity for the motivated offender (Newman, 1972; Zelinka and Brennan, 2001). Ceccato studied 76 cases of rape in Stockholm municipality during the period 2012–13 (Ceccato, 2014). She concluded that the 'micro-spaces' of rape share three commonalities. First, they are at or very close to areas with vegetation such as parks, forested areas or interstitial places that are easy to hide in. Second, they offer an easy escape route from the crime scene, because they are located close to public transportation such as bus stops or subway/train stations. Third, they tend to be secluded, with, for example, tunnels, ditches or stairs. The study did not answer the question of why other similar types of spaces in the city that share these commonalities reported no cases of rape during that period. Part of the answer to this question might lie in examining the types of neighbourhoods where these micro-places are located, that is, placing micro-scale spaces in their meso-scale context.

Neighbourhoods where people have a high fear of crime or where there are few pedestrians on the streets provide what we term *meso-scale* anonymity for the offender. Opportunities for rape are greatest in neighbourhoods where social control is weakest and where individuals (both men and women) have become reluctant to go out (Foster and Giles-Corti, 2008; Gray et al., 2011). Evidence shows that socially disorganized neighbourhoods have disproportionately more crimes, including cases of rape (for example, Amir, 1971; Maume, 1989; Pyle, 1974; Tewksbury and Mustaine, 2006). Social Disorganization Theory (Shaw and McKay, 1942, and, later, Kornhauser, 1978; Bursik and Grasmick, 1993; and Sampson and Groves, 1989) provides a significant part of the theoretical underpinning for many studies that focus on the impact of area and neighbourhood characteristics on crime. Shaw and McKay (1942) linked neighbourhood social disorganization to delinquency and crime. The mechanisms linking socially disorganized neighbourhoods and crime relate to people's inability to exercise social control in their neighbourhoods and solve jointly experienced problems. In neighbourhoods with high population turnover, guardianship is often low, thereby further enhancing the ecological conditions associated with rape. Other neighbourhood characteristics are important to an understanding of the geography of rape, including population overcrowding (Maume, 1989); the number of bars in an area (Rosay and Langworthy, 2003); the number of vacant dwelling units; and expressways cutting the neighbourhood (Pyle, 1974).

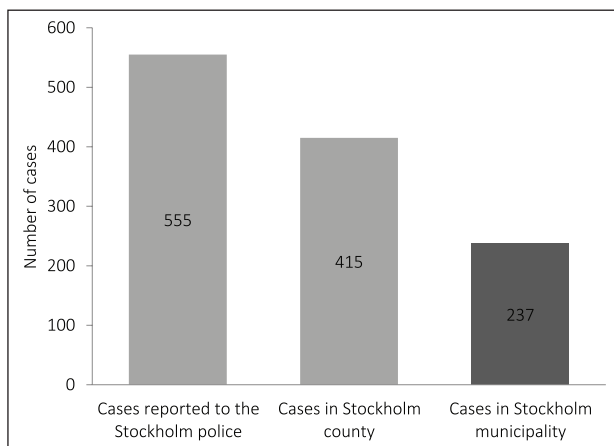
Person–environment interactions are central to our understanding of where crimes take place at the small area level, including, we argue, the crime of rape – a crime that necessarily involves direct contact between an offender and their victim. The ecological characteristics found in a place define the area-level risk of rape occurring in that place. Given the above review we define such places in terms of three constructs. First *accessibility* – the ease of getting to and away from the crime scene for the offender because of familiarity, but also bringing potential victims into an area. Second *opportunity* – the potential of an area to provide suitable, vulnerable targets. Third *meso- and micro-scale anonymity* – the potential of an area, at the meso-scale, to reduce the risks an offender faces and, at the micro-scale, to provide cover and concealment for the actual criminal act. We shall use these three conceptual categories in our model.

### *Hypotheses of study*

Two-stage rational choice theory provides a conceptual basis for separating the ecological level (numbers of cases of rape by small area) from the collection of individual criminal acts. It provides an argument to support the value of studying rape as counts of cases at the small area level. Here we test the following hypotheses at the small area (*basområde*) scale:

- (1) High counts of rape are associated with areas offering *opportunity* to the offender. These are areas with a relatively large number of vulnerable female targets either frequenting or just passing through, perhaps associated with their place of residence or with the locations of particular institutions or amenities (for example, Ceccato, 2014; Hewitt and Beauregard, 2014; Pyle, 1974). (*Opportunity hypothesis*)
- (2) High counts of rape are associated with areas with poor surveillance and/or poor social control where residents have a high fear of crime. Such areas may provide anonymity for the offender at the meso-scale (Canter and Larkin, 1993; Maume, 1989; Tewksbury and Mustaine, 2006). Neighbourhoods with buildings and street layouts providing areas for concealment, perhaps because of inadequate lighting or quiet alleyways, may provide a number of suitable ‘micro-spaces’ for the offence to be committed, thereby providing anonymity at the micro-scale (for example, Pyle, 1974; Rosay and Langworthy, 2003). (*Meso- and micro-scale anonymity hypothesis*)
- (3) High counts of rape are associated with areas that have good accessibility because they bring potential targets into an area and offer a quick and easy escape route from the crime scene for the offender. An offender is likely to be familiar with at least a subset of such areas because, the more accessible a place is, the more likely it is to form part of their awareness space (for example, Lebeau, 1985; Rhodes and Conly, 1981; Rosay and Langworthy, 2003). (*Accessibility hypothesis*)

We defer discussion of the variables used to test these hypotheses to the next section (‘Area risk factors’ subsection).



**Figure 1.** Rape cases used in this study, 2008–9 ( $N = 237$ ).

*Data source:* Police-recorded data, Stockholm Police Headquarters, 2010.

## The Stockholm rape study

This case study is limited to the Stockholm municipality (population 880,000 in 2012), which means the inner-city area and those suburbs that administratively belong to the city of Stockholm. The municipality is an archipelago well connected by bridges, roads and an efficient public transport system consisting of subways, trams, commuting trains and buses. It is in the inner-city area that the main public transport junction is located, concentrating a large number of people daily in the city's centre. Stockholm's inner city comprises residences where citizens enjoy a good quality of life. Housing, which may be rented, privately owned or occupied under a cooperative agreement, is of a high standard. Residents in these areas, on average, report a 30 percent higher annual income than those residing in the 10 peripheral areas (Statistik Stockholm, 2012). The mass-produced blocks of flats built in the 1960s and 1970s do not all perform equally well: low prices are often linked to poor architecture, lack of amenities and social problems.

### *Data on rape cases*

Stockholm's Police Headquarters provided records for the years 2008 and 2009. They contained crime code, day and hour of the crime and the geo-coordinates of where the rape happened. The records showed 555 cases of outdoor rape (including attempted rape) in this period. Outdoor locations are places where police surveillance and police patrols can limit or discourage acts of crime and disorder. Outdoor locations include: streets; transit environments; areas and premises adjacent to railways; metro and traffic terminals the public has access to; other facilities for public events that have the character of an outdoor facility such as an open-air sports ground (BRÅ, 2012). Of the 555 cases, 415 took place in Stockholm county and 237 cases in Stockholm municipality (140 cases took place elsewhere in the country or abroad). This analysis involves the 237 cases in Stockholm municipality (Figure 1).



A Geographical Information System (GIS) was used to map rape cases and then aggregate to the *basområde* level. This is the smallest spatial unit of analysis giving access to demographic, social and economic statistics in Sweden. Note that, of the 407 *basområdes* in the Stockholm municipality, 63 percent have no reported cases. We discuss the implications of this for modelling in the ‘Statistical models’ subsection.

### *Area risk factors*

We describe the covariates used to measure small area level accessibility, meso- and micro-scale anonymity and opportunity. Table 1 summarizes the 11 area-level covariates by category: two measuring aspects of accessibility; five measuring meso-scale anonymity and one measuring micro-scale anonymity; three measuring opportunity (detail as to data sources is provided in Table 2 in Appendix A).

We argue that areas exhibiting high levels of *opportunity* are those with large numbers of women (targets) and those with institutions (such as schools) or amenities (such as alcohol outlets). Women may be putting themselves at additional risk if they merely pass through areas with bars, where there may be more potential perpetrators. Areas of Stockholm meeting these criteria will include high-density residential neighbourhoods near the city centre and some suburban neighbourhoods.

We identify areas exhibiting high levels of *meso-scale anonymity* as those neighbourhoods with poor social control. We identify areas with poor social control using data on the percentage of people who fear crime and levels of population turnover. High values of these variables are indicators of poor social control. Neighbourhoods where the fear of crime is high tend to lack informal surveillance (‘eyes on the street’) because people may be reluctant to spend leisure time out of their homes. High turnover is often associated with neighbourhoods that are less affluent, less desirable and less safe. Young males with no family and lacking steady employment may comprise a significant element of the resident population. We include average income as an indicator of affluence. Areas of low average income may show poor social control. Finally, we also define industrial and forested areas as exhibiting *meso-scale anonymity*. Such places, particularly at certain times of the day, can be lonely and isolated with little or no surveillance.

We identify areas with good public transportation or that are located in the city centre of Stockholm as areas that enjoy good *accessibility*. Such areas are likely to form part of a would-be offender’s awareness space. We use access to subway stations, rather than, say, the bus network, as our measure of accessibility because previous work (Ceccato, 2013, 2017) has drawn attention to the role of the subway system in providing targets for would-be perpetrators and providing a rapid means of escape from the crime scene. The city centre is arguably the most accessible area in any large city but its role in any study of rape victimization clearly overlaps other categories including, at certain times of the day, meso-scale anonymity and opportunity as we have defined them.

We have argued above that anonymity needs to be defined at two scales, but we have no direct data on the number of ‘micro-spaces’ in each *basområde* that offer sites that give the attacker anonymity in the act of rape (*micro-scale anonymity*). There are clear challenges to providing a measure of the number of such spaces. In this study we have used a surrogate measure – the number of cases of street robbery – because many forms



**Table 1.** Posterior estimates (posterior mean and 95% credible interval) of the risk factor effects from the Poisson regression (Models 1 and 2).

Hypotheses	Risk factors	Posterior estimates of risk		
		Model 1	Model 2	
Accessibility Anonymity Meso-scale Micro-scale Opportunity	Whether subway station is present or not	1.57 (1.09, 2.27)	1.43 (1.00, 2.14)	
	Whether area is close to the city centre or not	1.90 (1.22, 3.04)	1.90 (1.21, 3.00)	
	Average income (in 10,000 Kronor)	0.98 (0.97, 1.00)	0.98 (0.96, 1.00)	
	Percentage of the population who fear crime and who avoid going out in the neighbourhood	1.00 (0.99, 1.02)	1.01 (0.98, 1.03)	
	Population turnover	2.06 (1.16, 3.53)	1.86 (1.02, 3.22)	
	Forest present or not	1.07 (0.45, 2.18)	1.08 (0.45, 2.19)	
	Whether area is an industrial area or not	0.65 (0.35, 1.19)	0.69 (0.37, 1.19)	
	Counts of street robbery	0	Reference category	1.05 (1.02, 1.07)
		1–2	1.14 (0.69, 1.93)	
		3–4	1.63 (0.92, 2.89)	
		≥ 5	2.67 (1.55, 4.60)	
	Number of women aged 13–65 (in '000)	1.16 (0.90, 1.52)	1.34 (1.05, 1.73)	
	Whether educational institution(s) is(are) present or not	1.29 (0.92, 1.82)	1.34 (0.94, 1.80)	
	Whether alcohol outlet(s) is(are) present or not	0	Reference category	1.03 (1.00, 1.05)
		1–2	0.97 (0.58, 1.60)	
3–6		1.17 (0.69, 2.01)		
≥ 7		1.55 (0.88, 2.78)		

of street robbery also require micro-scale anonymity for carrying out the criminal act. It is not, of course, ideal. Cases of street robbery also occur in places with many people, which offer opportunities for stealthy, distracted, rapid, offender-on-victim interactions. We suggest that future research should look into constructing a more 'direct' measure of micro-scale anonymity at the *basområde* scale. Perhaps remotely sensed imagery of Stockholm's land use (forestry, industry, built form) by day and luminosity by night might provide a useful resource.

## Statistical models

An ecological analysis of rape aims to assess how risk factors *individually as well as jointly* affect the risk of outdoor rape. To achieve this, we describe an approach that encompasses two statistical techniques: a Bayesian Hierarchical Poisson regression model for a ‘whole map’ analysis and a Bayesian profile regression (BPR) model (Molitor et al., 2010) for a ‘local’ analysis. Often employed in ecological crime studies (for example, Congdon, 2013, and Matthews et al., 2010), a Bayesian Poisson regression model estimates the ‘whole map’ effect of individual covariates. Li et al. (2013) discuss the benefits of adapting the Bayesian approach in modelling small area crime data. In this paper, we perform a further localized analysis using BPR to understand how the covariates *jointly* affect the area-level risk. The combination of the results from both methods provides a better understanding of the ecology of outdoor rape. Because of its novelty to crime analysis, we will first highlight the benefits of employing BPR in the present study then provide the technical details of the two methods.

BPR is a Bayesian clustering method that seeks to cluster or group areas based not only on the similarity of rape occurrence but also on the similarity of the *risk factor profile*. This modelling approach offers several advantages. First, by using risk factor profiles, BPR considers all risk factors *jointly* when assessing similarity. For example, two areas are considered to have similar risk factor profiles if they have similar levels of population turnover, income and female resident population. This clustering of risk factor profiles places no restriction on the correlation structure amongst the covariates, hence resolving the multicollinearity problem that is often encountered when fitting a standard regression model. Moreover, by using risk factor profiles, BPR does not require the specification of an explicit covariates–outcome relationship (as required when fitting a standard regression model), hence allowing for potential interactions *amongst* the covariates.

Second, BPR clusters areas based on their risk factor profiles *and* their risk levels, allowing us to differentiate, for example, two groups of high-risk areas, one group having a set of distinctly different *local characteristics* from the other.

Third, the ‘local focus’ of BPR enables us to deal with spatially heterogeneous covariate effects. For example, a small female population may not always be associated with a low risk of rape. Although there are other methods to allow for such heterogeneous covariate effects (for example, geographically weighted regression – Fotheringham et al., 2002 – and Bayesian spatially varying coefficient models – Gelfand et al., 2003), these methods may suffer from the multicollinearity issue (Wheeler and Tiefelsdorf, 2005).

Fourth, to model the observed rape cases, BPR allows the inclusion of random effects, which are often used to deal with the so-called overdispersion problem, whereby the variability in the observed count data exceeds that from a Poisson distribution. Often encountered in analysing small area crime data, overdispersion can be a result of effects arising from unmeasured/unobserved risk factors. Standard clustering methods do not include random effects and thus fail to account for such an important source of uncertainty in modelling small area crime data.

Finally, placed in the Bayesian framework, BPR allows the number of clusters to be estimated from the data and this uncertainty is fully accounted for in estimating other parameters in the model. Non-Bayesian clustering methods fix the cluster number – such uncertainty is ignored.

*A ‘whole map’ analysis using Poisson regression.* In this part of the analysis, we use Poisson regression (McCullagh and Nelder, 1989; Osgood, 2000) to estimate the overall effects of the risk factors on the number of cases of outdoor rape. Denoting  $O_i$  the number of outdoor rape cases in *basområde*  $i$  ( $i = 1, \dots, 407$ ), a Poisson regression models  $O_i$  using a Poisson distribution (the data model), i.e.  $O_i \sim \text{Poisson}(\mu_i)$ . The Poisson mean  $\mu_i$  (the process model) is expressed as a function of area-level risk factors and random effects, where the latter are included to account for effects that arise from unobserved/unmeasured covariates. The process model is given by

$$\log(\mu_i) = \alpha + \sum_{l=1}^{11} \beta_l X_{li} + u_i + v_i$$

The expected cases  $\mu_i$  are log transformed to ensure that they are non-negative. The parameter  $\alpha$  is the (log) overall expected number of cases in Stockholm municipality;  $\sum_{l=1}^{11} \beta_l X_{li}$  represents the effects arising from the risk factors (see next subsection) and  $u_i$  and  $v_i$  are the random effect terms. The regression coefficient  $\beta_l$  ( $l = 1, \dots, 11$ ) quantifies the impact of the risk factor  $X_{li}$  on rape occurrence. For the random effects,  $u_i$  is modelled as a set of spatially structured random effects, representing a set of unobserved risk factors whose joint effects on the occurrence of rape display spatial structure, and  $v_i$  is modelled as a set of spatially unstructured random effects, representing the unobserved risk factors whose effects on the occurrence of rape do not display spatial structure. To impose spatial structure, the intrinsic conditional autoregressive (ICAR) model is applied to  $u_i$  such that neighbouring areas have similar values of  $u_i$  (Besag et al., 1991). This models spatial dependence in the Poisson mean.<sup>2</sup> For  $v_i$ , independent normal distributions,  $N(0, \sigma_v^2)$ , are assigned. Incorporating the two random effect terms helps account for variability that is not explained by the 11 covariates that are already included in the model. In the present dataset, a source of such extra variability comes from a large proportion of areas with 0 cases, and the two random terms help deal with these 0. Note that, compared with the modelling approach discussed in Osgood (2000), an offset term is not included for two reasons. First, it is difficult to define the ‘at risk’ population for outdoor rape. Second, using female residential population as an offset implies that a smaller female population is associated with a lower risk of rape. However, this assumption is not true for some of the *basområdes* (as results will show). Therefore, we directly modelled the underlying expected counts as opposed to modelling the rates but included the female population as a covariate, whose effect on risk is then estimated from the data.

Under the Bayesian approach, prior distributions are assigned to all model parameters (the parameter model). Here, non-informative priors were used so that the testing of the hypotheses discussed above is minimally affected by the prior choice. This reflects the

small amount of prior evidence we have from which to start the current modelling. Specifically, for the regression coefficients, a normal prior  $N(0, 10000)$  was used. For the intercept  $\alpha$ , we used the improper uniform distribution defined on the whole real line as its prior distribution (GeoBUGS, 2004). To  $\sigma_v$  and  $\sigma_u$ , the standard deviations of the two sets of random effects, a normal prior  $N(0, 1)$  is assigned; to ensure that a standard deviation parameter is non-negative, this normal prior is truncated so that it takes only non-negative values (Gelman, 2006). Parameter estimation was carried out in WinBUGS (Lunn et al., 2000) via iterative Markov chain Monte Carlo (MCMC) methods and the WinBUGS code is provided in Appendix B.

*A 'local' analysis using Bayesian profile regression.* This part of the analysis employs BPR to understand the local environments where outdoor rapes occur. BPR allocates *basområdes* into clusters (or groups<sup>3</sup>) so that *basområdes* within each cluster tend to have similar risks of rape **and** similar risk factor profiles. Specifically, the risk factor profile of an area is simply a list of the observed risk factor values for the area. Given a cluster formation, the risk factor profile for each cluster is estimated based on the observed risk factor profiles of its member areas. In our study, the risk factors are both continuous- and discrete-valued (binary and categorical). Within a cluster, the continuous-valued risk factors are modelled using a multivariate normal distribution and the binary and categorical risk factors are modelled respectively using a vector of independent binary and categorical random variables. Independent of the cluster allocation, we also included two sets of random effects, spatially structured and spatially unstructured, to account for unmeasured/unobserved risk factors as well as dealing with the many areas with 0 cases. The specifications of these two sets of random effects are the same as those in the Poisson regression model. The BPR model was fitted using the R package PREmM, which provides a simple platform to fit the BPR model using MCMC algorithms (see Liverani et al., 2015, for more detail).

It is important to note that BPR performs cluster allocation probabilistically, meaning that the *number* of clusters and the *composition* of each cluster (which areas are included in any cluster) may differ from one MCMC iteration to the next. This probabilistic way of clustering can be thought of as exploring different ways of grouping the areas and reflecting the uncertainty associated with the number of clusters and their composition. However, the challenge is to summarize the rich MCMC outputs and, ultimately, to decide on the number of clusters and their composition. To post-process these MCMC outputs, we follow the three steps suggested in Molitor et al. (2010). Briefly, the first step constructs the posterior similarity matrix  $\mathbf{S}$ , which summarizes different partitions of areas over all MCMC iterations. Each off-diagonal entry in  $\mathbf{S}$ ,  $s_{ij}$  (for  $i \neq j$ ), quantifies how likely it is for two *basområdes*  $i$  and  $j$  to be grouped together over all MCMC iterations. The second step finds a representative partition that has the shortest distance to the similarity matrix  $\mathbf{S}$ . Using all MCMC iterations, the third step then produces the estimates for the risk and risk factor profile associated with each of the clusters within the representative partition. The results presented in the next section are from the representative partition.

Note also that, because BPR models the continuous-valued risk factors using multivariate normal distributions, each continuous-valued risk factor was transformed to be approximately normally distributed (see Table 2 in Appendix A). The two discrete-valued covariates – number of alcohol outlets and of robbery cases – were converted into categorical

covariates (as defined in Table 1) because no suitable normality transformation was found for the highly positively skewed distributions. For the whole map analysis, the Poisson regression was fitted based on the untransformed continuous-valued risk factors, with alcohol outlets and robbery cases being either in categorical form (Model 1) or in count form (Model 2) (see Table 1).

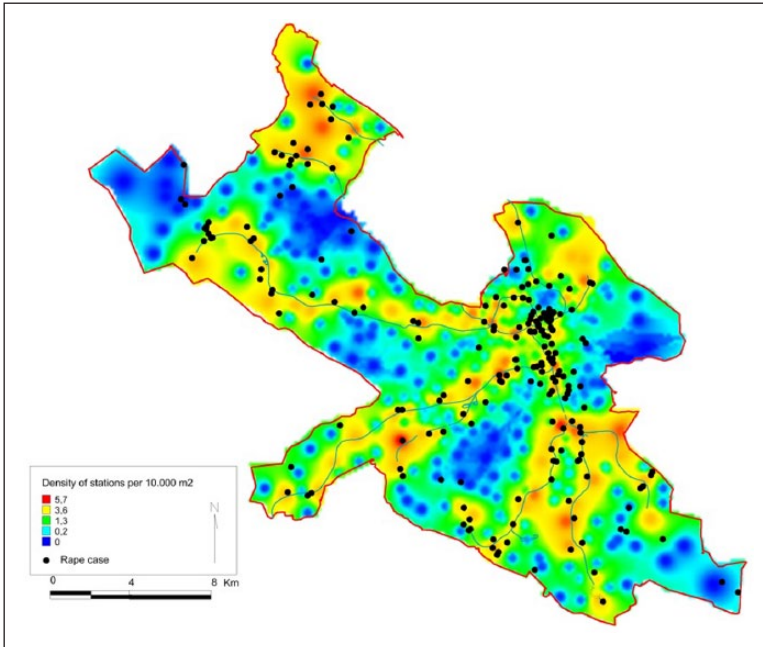
## Results

*'Whole map' associations for the risk factors.* The estimated effects of the 11 risk factors from the Poisson regression are summarized in Table 1. Here  $\exp(\beta_l)$  represents the relative risk owing to a one-unit change in the  $l^{\text{th}}$  risk factor (or compared with the reference category for binary/categorical risk factors) while keeping all other risk factors fixed.<sup>4</sup>

The presence of a subway station in a *basområde* generally increases the risk of outdoor rape in an area by over 50 percent (with 95% credible interval (denoted as 95% CI hereafter): 9% – 127%), perhaps because of the opportunities it offers for escape. An alternative explanation is that public transportation sites are the places where motivated offenders encounter the women whom they will follow and assault. Figure 2 displays, in raster form, data on the density of stations as continuous colour gradations. Overlaying the cases of rape suggests an association between the locations of subway stations and the locations of rape cases. In addition to the effect of subway stations, the risk of rape is nearly doubled in the 'city centre'. The importance of both these two risk factors provides evidence supporting the *accessibility hypothesis*.

A *basområde* with a high level of population turnover tends to have an increased risk of rape (posterior mean of 2.06 with 95% CI: 1.16–3.53). The results also show some, albeit weak,<sup>5</sup> evidence of an increased risk of rape associated with a falling level of average income (posterior mean of 0.98 with 95% CI: 0.97–1.00), suggesting that a reduction of SEK10,000 (equivalent to US\$1175) leads to a 2 percent (95% CI: 0%–3%) increase in risk. By contrast, the other three meso-scale anonymity risk factors (people's fear of crime and the presence of forested or industrial areas) do not appear to be associated with the risk of rape. These results suggest that *meso-scale anonymity*, in the sense of poorer areas with high population turnover, does contribute to explaining variation in the number of rape cases. The covariate on street robbery counts, used as a surrogate measure of *micro-scale anonymity*, is found to be positively associated with the risk of rape. Under Model 2, an increase in robbery count is found to be associated with an increase in the risk of rape (posterior mean of 1.05 with 95% CI: 1.02–1.07), indicating that one additional robbery case leads to a 5 percent (with 95% CI: 2%–7%) increase in the risk of rape. Such a linear trend is also seen when the covariate is included as a categorical variable in Model 1, although the effect is significant only when the number of street robberies is five or more, suggesting that the relationship is strongest where micro-scale anonymity is greatest. These results, taken together, indicate that both forms of anonymity contribute to our understanding of small area variation in the number of cases of rape.

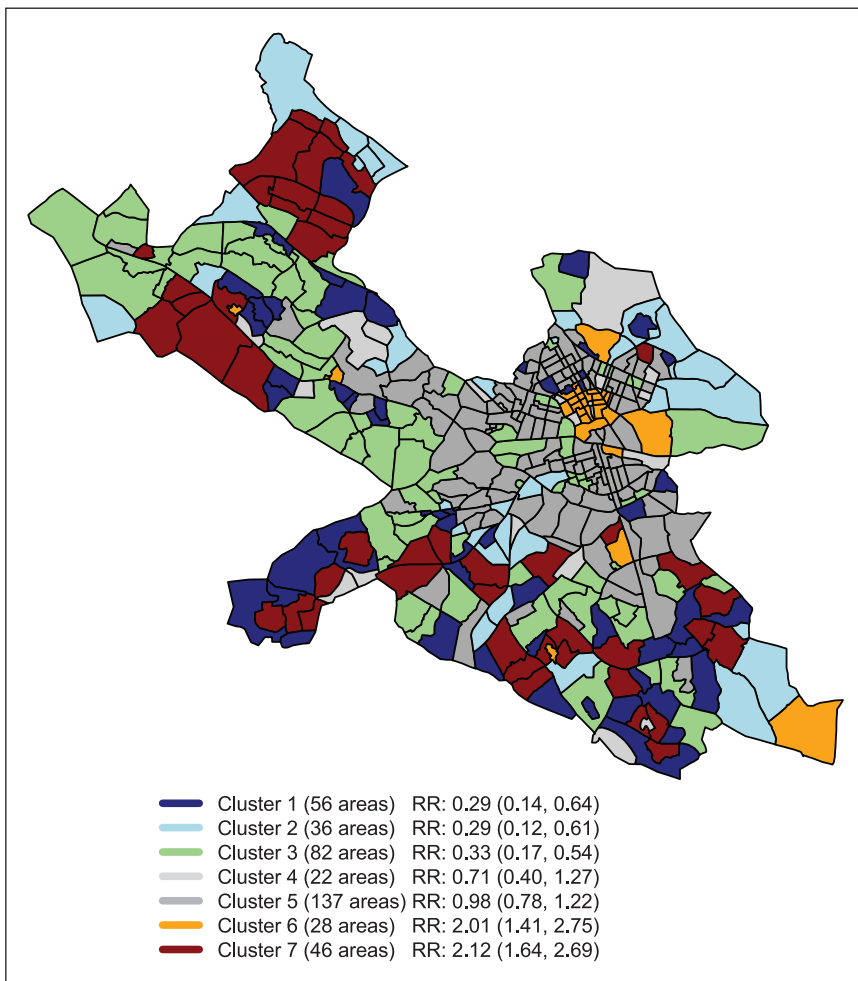
The whole map analysis does not appear to offer a clear message regarding the importance of the three risk factors we have associated directly with the *opportunity hypothesis*, that is, the size of the female population, the number of alcohol outlets and the presence of educational establishments. The first two risk factors are significant in Model



**Figure 2.** Subway station density and cases of rape, 2008–9 (cells: 100 × 100 metres).  
 Data source: Police-recorded data, Stockholm Police Headquarters, 2010.

2 but not under Model 1. However, it is worth noting the clear increasing trend between the number of alcohol outlets in an area and the risk of outdoor rape, suggesting that risk in an area does increase when a larger number of alcohol outlets are present. Also, in the subsection above on ‘Area risk factors’, we remarked that the city centre can be associated with the opportunity hypothesis and as noted above this covariate is significant. As we shall see in the next section, the application of BPR allows us to gain further and clearer understanding of various risk factors within a local context.

*‘Local’ associations for the risk factors.* BPR was applied to reveal the localized characteristics of places where rape took place. Over the 200,000 MCMC iterations, about 99.7 percent of them consist of seven clusters with 0.3 percent having eight clusters, showing some variability (or uncertainty) in determining the cluster number. To summarize different clustering options of the areas, the representative partition of the *basområdes* was produced following the steps outlined in the subsection ‘A “local” analysis using Bayesian profile regression’. Figure 3 shows the seven clusters in the representative partition. There are three low-risk clusters (Clusters 1–3, all with the 95% CI for relative risk below and excluding 1). There are two medium-risk clusters (Clusters 4 and 5, with relative risks of 0.71 (95% CI: 0.40–1.27) and 0.98 (95% CI: 0.78–1.22), respectively. Finally, there are two high-risk clusters (Clusters 6 and 7, both with the 95% CI for relative risk above and excluding 1).

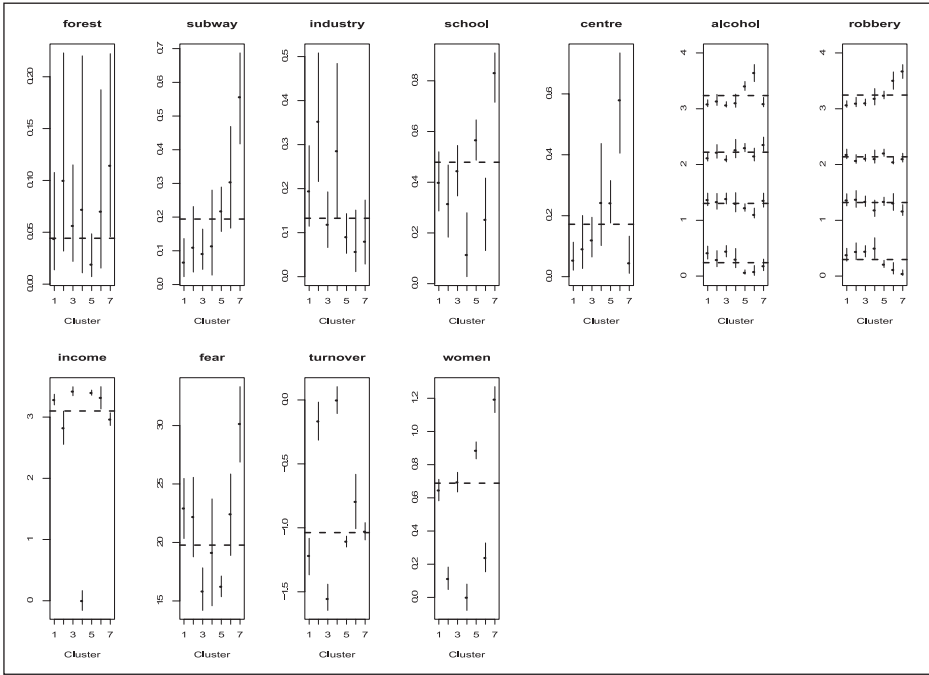


**Figure 3.** Membership of the *basområdes* under the representative partition using the Bayesian profile regression model.

Note: For each cluster, or grouping of areas, the legend reports the number of *basområdes*, the posterior mean of the cluster-specific relative risk (RR) and the 95% credible interval.

Figure 4 describes the estimated cluster-specific risk factor profiles. For the two high-risk clusters, Cluster 6 represents areas mainly in the city centre with high population turnover, small female residential populations and small numbers of schools. There are also large numbers of alcohol outlets (a high proportion of *basområdes* are in category 4) and large numbers of street robbery cases in this cluster. The characteristics of this cluster illustrate that it is inappropriate to use the female population as an offset. Cluster 7, on the other hand, represents relatively deprived suburban areas with large numbers of subway stations and schools, relatively low income, and high levels of fear of crime. The





**Figure 4.** The estimated risk factor profile for each of the 7 clusters in the representative partition.

Notes: The dots correspond to the posterior means and the vertical bars represent the 95% credible intervals. The horizontal dashed line shows the Stockholm average for each risk factor. For the two categorical covariates, alcohol and robbery, the dots and the vertical bars, representing the posterior means and the 95% credible intervals respectively, show the estimated proportions of areas in each of the four categories across the seven clusters and the horizontal dashed line shows the Stockholm average for each category.

relatively large female residential populations in these areas present would-be offenders with a large number of potential targets. All factors combined contribute towards the high risk of rape in these areas. Similar to Cluster 6, relatively large numbers of alcohol outlets and street robbery cases are found in these areas. The finding of these two groups of high-risk areas, each with distinctly different local characteristics, provides a better description of the clusters, illustrating a benefit of the BPR approach.

There are two medium-risk clusters, Clusters 4 and 5. Cluster 5 consists of areas mainly on the outskirts of the city centre (Figure 3). These areas tend to have a relatively large female residential population, more schools, a generally low fear of crime amongst residents and good subway connections. There are large numbers of alcohol outlets present in these areas but numbers of robbery cases are relatively low, compared with the levels in the two high-risk clusters. Cluster 4, the other medium-risk cluster, consists of 22 *basområdes*. These *basområdes* contain an airport and shopping malls that have 0-values for income, population turnover and size of the female residential population (Table 2 in Appendix A). Although the above three risk factors are all 0, areas within this cluster are mainly industrial areas and schools are present in only a few of these areas.

Clusters 1–3 are the three low-risk clusters. Cluster 2 represents primarily industrial areas whereas Clusters 1 and 3 are the affluent suburban areas with few alcohol outlets and few robbery cases, suggesting few micro-sites offering anonymity for the criminal act.

## Discussion of the results

We hypothesized that small area variation in the number of cases of rape can be understood by reference to three theoretical constructs: accessibility, opportunity and (meso- and micro-scale) anonymity. To discuss the risk factors identified by the ‘whole map’ Poisson regression modelling, we consider both models (1 and 2) reported in Table 1. We observe that all the theoretical constructs are significantly associated with small area variation although not with the same degree of empirical support. The two covariates measuring accessibility – city centre; presence of a subway station – are significant in both sets of results. Two of the five covariates measuring meso-scale anonymity are also significant in both sets of results: population turnover and average income, which taken together suggest that poor areas with low levels of social control (Maume, 1989) present the highest risk. The covariate measuring micro-scale anonymity – the number of street robberies – was significant in both sets of results, and the results of the two analyses taken together suggest that the association is most pronounced in areas with the largest numbers of street robberies. Not all street robberies necessarily depend on high levels of micro-scale anonymity, but those areas with high numbers of cases are likely to be the areas offering the largest number of suitable sites. On the other hand, the evidence for the importance of opportunity (as defined by the internal characteristics of a *basområde*) is less compelling. The presence of schools is not significant and the size of the resident female population and the number of alcohol outlets in the *basområde* are significant only when the alcohol covariate is measured as a continuous covariate. Here it is important to recall that *basområdes* represent artificial partitions of the urban space. The effects of some attributes may be felt ‘at a distance’ from their actual location, that is, when women are on the move. This was noted by Ceccato (2014: 103), who observed that, although more than half of the rape cases she analysed occurred within 1 km of the woman’s residence, many of these, as well as the other cases beyond 1 km, might be recorded in *basområdes* other than where the women lived. The same is likely to be true in the case of attacks linked to areas with alcohol outlets.

It is also interesting to note the risk factors that were not significant: *basområdes* where industry or forest was present and areas where fear of crime was high. We linked these covariates to the meso-scale anonymity hypothesis. The location of parks and forested areas may be associated with where *some* rapes occur (Ceccato, 2014), but such areas seem not, in general, to be associated with a significant increase in area-level risk. This may be because other characteristics need to be present in order to constitute a ‘whole map’ risk factor. In the case of fear of crime, it is high levels of violent and property crimes that lie at the core of many people’s fear of crime (Uittenbogaard and Ceccato, 2012). People in better-off areas may declare high levels of anxiety about crime and express concerns for their personal safety (Da Costa and Ceccato, 2015), but they may have the resources to provide themselves with protection that lowers risk, including the risk of sexual assault.

Profile regression reveals the localized characteristics of places where rapes occur and in particular the role of risk factor interactions in determining local area risk. Risk factors that do not emerge as significant in a whole map analysis may emerge in this form of regression modelling because of local context. Two types of areas stood out where risks were high. Cluster 6 represents areas in the city centre with high population turnover, large numbers of alcohol outlets and large numbers of cases of street robbery. This is not surprising. Cluster 7 represents relatively deprived suburban areas with large numbers of subway stations and schools, relatively low income levels, high levels of fear of crime and relatively large female residential populations. Relatively large numbers of robbery cases occur in areas where there are also many alcohol outlets. In the case of the areas in cluster 7, three covariates emerge as important risk factors locally but not in the ‘whole map’ analysis. These are areas where many schools are located, the fear of crime levels are high and there are large female residential populations. This observation highlights the importance of understanding the effect of a risk factor within the local context, one of the original motivations for this study, and jointly within the setting provided by other risk factors.

We conclude this discussion of findings by acknowledging that police data on rape are far from problem free. A general estimate from the National Council of Crime Prevention shows that only 10–15 percent of all sexual crimes are reported to the police (BRÅ, 2014). Of the cases that do go to trial, a quarter result in acquittal. Rapes committed by a stranger that happen in public places fit the ‘stereotypes of rape and rapists’ and constitute the majority of cases that reach court. If a reported case happened in an open or forested area or in a place unknown to the raped woman, then the precise location of the rape usually cannot be determined. For a further discussion of data quality on rape in Sweden, see Ceccato et al. (2017).

## Conclusions

This article has reported findings from an analysis of the geographical distribution of cases of rape in Stockholm in 2008–9 at the small area (*basområde*) scale. Theories from urban criminology provided the basis for testing three hypotheses based on opportunity, accessibility and anonymity using land use, demographic and socio-economic characteristics of the small area units. Bayesian regression methods and GIS technology underpin the work in this study. The main contribution of this study is that it indicates the areas that present the highest risk of rape. Previous research on an individual case by case basis in Stockholm (Ceccato, 2014) drew attention to certain environmental commonalities relating to anonymity but was unable to explain why other areas, despite offering similar levels and forms of anonymity for example, did not report any rape cases. The implication is that certain classes of risk factor, especially land use and service outlet variables, may be highly place or context specific. This study sheds light on this contextual effect and the need to consider contextual effects when crime prevention initiatives are put into practice.

Findings from the ‘whole map’ analysis have confirmed the importance of the accessibility dimension to an understanding of the geography of rape. Both of the accessibility risk factors (the presence of subway stations; proximity to the city centre) are significant in both Models 1 and 2. The importance of the opportunity dimension is captured by the presence of alcohol outlets, particularly where there are large numbers of such outlets.

The number of women resident in an area is also significant in one form of the model. There is support too for the importance of both forms of anonymity. Population turnover is significant in both forms of the model, as is the number of cases of street robbery. Our findings with respect to micro-scale anonymity must be qualified by our dependence on surrogate data. However, notwithstanding these reservations, there is clear evidence that areas experiencing the highest levels of street robbery are also the areas experiencing the highest risk of rape – a finding with clear crime prevention implications.

Findings from the local analysis in this study reveal two types of high-risk area within Stockholm. The first are the parts of the city centre where there are large numbers of alcohol outlets, high residential population turnover and high counts of robbery. The second type of high-risk areas comprises poor suburban areas with schools and large female residential populations where subway stations are located, and where a large proportion of the population express fears about crime sufficient to make them avoid going out in the neighbourhood. So, in a local context, fear of crime, female residential population and the presence of educational establishments are found to be important risk factors. Finally, it is interesting to note those risk factors at the *basområde* level that were not significant in either form of the whole map model or the local regression model, namely the presence of industrial or forested areas.

Finally, we turn to policy issues. Area-focused crime prevention strategies, particularly those designed for large urban areas, should draw on findings from both individual- and aggregate-level research. Although focusing on individual incidents will provide detailed insights into particular cases or groups of cases, small area research is better placed to make inferences about the environments within which particular crimes occur and the importance of local context. Local analyses can be used to tailor crime prevention measures to the locally significant risk factors. This study (combining the ‘whole map’ effects of risk factors on the risk of rape and the local characteristics of areas at different risk levels using profile regression) clearly shows the need to use a two-stage analysis approach when the goal is to provide evidence for the prevention of sexual violence in outdoor environments.

In the case of rapes, one way forward is to reflect upon the roles that transport nodes and the maintenance of public spaces around them play in the incidence of outdoor rapes (Ceccato, 2013, 2017). The transportation system is one of the most important situational elements for outdoor rapes in Stockholm. These findings confirm trends in perceived safety by passengers. Safety surveys in Stockholm show that individuals often feel safe in the carriage, the bus and in the station’s environment, but their perceived safety levels decrease as they walk from the station to home, and vice versa (City of Stockholm, 2011). This means that authorities should tackle safety by adopting a door-to-door strategy or the so-called ‘whole journey approach’, although, as suggested above, it is likely that strategies will need to be tailored to local circumstances.

Some cases of rape happen just after women leave the bus stop or subway station. Because the subway system closes down overnight on weekdays, special attention should be given to buses, as they may become the only way to return home using public transportation. For the offender, transportation environments function both as a place where women can be identified and as a means to escape from the crime scene. The design, location and management of public facilities at entrances or close to stations and

transport nodes may be enough to discourage offenders and diminish opportunities for rape in the immediate surroundings of the stations. In more peripheral stations, adding CCTV cameras (particularly at exits and entrances) can support the work of ‘place managers’, ‘handlers’ and ‘guardians’, such as subway personnel, guards, safety hosts or parking lot attendants (Felson, 1995). The role of place managers and support systems and their potential to exercise guardianship need to be examined in relation to the way public spaces are used by different communities in the city. ICT technology (particularly mobile phones and social media) has a role to play in reducing sexual attacks (For more recommendations, see Ceccato et al. 2017).

Phumzile Mlambo-Ngucka, Executive Director of UN Women, wrote: ‘we have to work with communities to change harmful social norms and attitudes, and social institutions that discriminate and tolerate violence against women’ (UN Women, 2014). What this research draws attention to are the types of environmental or contextual circumstances that increase the risk of a rape occurring in an area. This study has contributed to this knowledge base by assessing the importance of the contextual circumstances surrounding the crime of rape. Interventions are needed that include but also go beyond ‘quick fixes’ of the local environment at the scale of specific places (better lighting, greater physical openness of urban parks, perhaps closure of certain paths and passage-ways). Preventive measures to try to reduce rape also need to take account of specific geographical-ecological contexts. That means reflecting not only on how to address the ‘harmful social norms and attitudes’ raised by Phumzile Mlambo-Ngucka but also on environmental design issues at different scales and identified for specific places. This may involve adjustments in the way cities are built and policed, but also changes in urban functions that create micro and meso spaces that allow a sense of ownership and control access. These findings call for a whole journey perspective to women’s safety. Since many rapes take place while women are on the move (on the way to or from home; to or from work or leisure), understanding the dynamics of rape calls for (data permitting) a space–time analysis of women’s levels of risk across the urban landscape.

## **Acknowledgements**

We would like to thank several anonymous referees for their helpful comments and advice during the evaluation process. Many thanks go also to colleagues in the Police in Stockholm county and Stockholm municipality for providing the data for this study.

## **Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The Swedish Research Council FORMAS financed this research (Grant 2009-479).

## **Notes**

1. The standard regression model, by contrast, combines the data model and the process model whereas parameters are constants estimated from the data. The Bayesian form of the standard regression model also includes models for the unknown parameters.
2. Modelling correlated Poisson counts through the data likelihood raises severe computational challenges. Modelling spatial dependence in the process model, so that the Poisson counts are

- treated as independent given the process, is one of the major advantages of hierarchical modelling over standard forms of regression modelling when dealing with spatial count data.
3. The term ‘cluster’ in profile regression refers to groups of areas that are classified together. Spatial contiguity is not a criterion for this classification.
  4. The relative risk measures indicate how much the expected number of rape cases would increase/decrease if a particular risk factor were to be increased by one unit and all other risk factors remained constant. This quantity,  $\exp(\beta_i)$ , is not an odds ratio (as in a logistic regression) since the Poisson regression model here deals directly with event occurrence, not probability.
  5. This is only weak evidence of an income effect because the upper bound of the 95% CI is on the borderline of 1, the value that indicates no effect.

## References

- Amir A (1971) *Patterns in Forcible Rape*. Chicago: University of Chicago Press.
- Beauregard E, Proulx J and Rossmo DK (2005) Spatial patterns of sex offenders: Theoretical, empirical and practical issues. *Aggression and Violent Behaviour* 10: 579–603.
- Belknap J (1987) Routine activity theory and the risk of rape: Analyzing ten years of National Crime Survey Data. *Criminal Justice Policy Review* 2: 337–356.
- Bernasco W and Luykx F (2003) Effects of attractiveness, opportunity and accessibility to burglars on residential burglary rates of urban neighbourhoods. *Criminology* 41(3): 981–1002.
- Besag J, York J and Mollié A (1991) Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 43(1): 1–20.
- BRÅ [The Swedish National Council of Crime Prevention] (2012) *Kodning av brott: anvisningar och regler*. Version 11.0, Stockholm. URL (accessed 29 March 2018): [http://www.bra.se/download/18.22a7170813a0d141d21800050764/Brottskodningslistan-2012+v11\\_3.pdf](http://www.bra.se/download/18.22a7170813a0d141d21800050764/Brottskodningslistan-2012+v11_3.pdf)
- BRÅ [The Swedish National Council of Crime Prevention] (2014) Rape and sexual offences. URL (accessed 29 March 2018): <http://www.bra.se/bra/bra-in-english/home/crime-and-statistics/rape-and-sex-offences.html>
- Brantingham P and Brantingham P (1984) *Patterns in Crime*. New York: Macmillan.
- Brooks SP and Gelman A (1997) General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7: 434–455.
- Bursik RJ and Grasmick HG (1993) *Neighbourhoods and Crime*. New York: Lexington.
- Canter D and Larkin P (1993) The environmental range of serial rapists. *Journal of Environmental Psychology* 13: 63–69.
- Ceccato V (2013) *Moving Safely: Crime and Perceived Safety in Stockholm’s Subway Stations*. Plymouth: Lexington.
- Ceccato V (2014) The nature of rape places. *Journal of Environmental Psychology* 40: 97–107.
- Ceccato V (2017) Women’s victimisation and safety in transit environments. *Crime Prevention and Community Safety* 19: 163–167. doi: <https://doi.org/10.1057/s41300-017-0024-5>
- Ceccato V, Wiebe DJ, Eshraghi B and Vrotsou K (2017) Women’s mobility and the situational conditions of rape: Cases reported to hospitals. *Journal of Interpersonal Violence*. doi: <https://doi.org/10.1177/0886260517699950>
- City of Stockholm (2011) *Trygg i Stockholm? 2011 En stadsövergripande trygghetsmätning*. Stockholm: City of Stockholm, Social Development Unit.
- Congdon P (2013) A model for spatially varying crime rates in English districts: The effects of social capital, fragmentation, deprivation and urbanicity. *International Journal of Criminology and Sociology* 2: 138–152.



- Da Costa and Ceccato V (2015) Assessing the adoption of household safety protection (HSP) in Stockholm, Sweden. *Crime Science* 4(1): 1–12.
- Dagens Nyheter (2013) Kvinna överfallen och våldtagen i centrala Stockholm [Woman attacked and raped in central Stockholm]. 23 January 2013. URL (accessed 29 March 2018): <http://www.dn.se/sthlm/kvinna-overfallen-och-valdtagen-i-centrala-stockholm>
- Dagens Nyheter (2016) Ung man gripen for valdtakt utanfor Stockholm [A young man arrested for rape in Stockholm outskirts]. 6 May 2016. URL (accessed 29 March 2018): <http://www.dn.se/nyheter/varlden/ung-man-gripen-for-valdtakt-utanfor-stockholm>
- DCJS [Division of Criminal Justice Services] (2014) Predatory sexual assault. URL (accessed 29 March 2018): <http://www.criminaljustice.ny.gov/crimnet/ojsa/stats.htm>
- Felson M (1995) Those who discourage crime. In: Eck JE and Weisburd D (eds) *Crime and Place*. Monsey, NY: Criminal Justice Press, 53–66.
- Foster S and Giles-Corti B (2008) The built environment, neighborhood crime and constrained physical activity: An exploration of inconsistent findings. *Preventive Medicine* 47: 241–251.
- Fotheringham AS, Brunson C and Charlton ME (2002) *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: Wiley.
- Gelfand AE, Kim H, Sirmans CF and Banerjee S (2003) Spatial modelling with spatially varying coefficient processes. *Journal of the American Statistical Association* 98: 387–396.
- Gelman A (2006) Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 1(3): 515–533
- GeoBUGS (2004) GeoBUGS User Manual Version 1.2.
- Gray E, Jackson J and Farrall S (2011) Feelings and functions in the fear of crime: Applying a new approach to victimisation insecurity. *British Journal of Criminology* 51: 75–94.
- Hechter M and Kanazawa S (1997) Sociological rational choice theory. *Annual Review of Sociology* 23: 191–214.
- Hewitt A and Beaugard E (2014) Sexual crime and place: The impact of the environmental context on sexual assault outcomes. *Journal of Criminal Justice* 42(5): 375–383.
- Kornhauser R (1978) *Social Sources of Delinquency*. Chicago, IL: University of Chicago Press.
- Lebeau JL (1985) Some problems with measuring and describing rape presented by the serial offender. *Justice Quarterly* 2(3): 385–398.
- Lewakowski B (2012) Half locked? Assessing the distribution of household safety protection in Stockholm. Master's thesis, Royal Institute of Technology, Stockholm.
- Li G, Ceccato V and Haining R (n.d.) On understanding the ecological characteristics of rape occurrence: A Bayesian profile regression approach. In revision; the manuscript is available upon request.
- Li G, Haining R, Richardson S and Best N (2013) Evaluating the no cold calling zones in Peterborough, England: Application of a novel statistical method for evaluating neighbourhood policing policies. *Environment and Planning A* 45(8): 2012–2026.
- Liverani S, Hastie DI, Azizi L, Papathomas M and Richardson S (2015) PRemiuM: An R package for profile regression mixture models using Dirichlet processes. *Journal of Statistical Software* 64(7): 1–30.
- Lovett J and Kelly L (2009) *Different Systems, Similar Outcomes? Tracking Attrition in Reported Rape Cases across Europe*. Child and Women Abuse Studies Unit, London Metropolitan University, London.
- Lunn D, Thomas A, Best N and Spiegelhalter D (2000) WinBUGS – a Bayesian modelling framework: Concept, structure, and extensibility. *Statistics and Computing* 10: 325–337.
- Matthews S, Yang T, Hayslett-McCall K and Ruback R (2010) Built environment and property crime in Seattle, 1998–2000: A Bayesian analysis. *Environment and Planning A* 42(6): 1403–1420.



- Maume DJ (1989) Inequality and metropolitan rape rates: A routine activity approach. *Justice Quarterly* 6: 513–527.
- McCullagh P and Nelder J (1989) *Generalized Linear Models*, 2nd edn. New York: Chapman & Hall.
- Ministry of Justice (1999) *The Swedish Penal Code*. Stockholm: Ministry of Justice. URL (accessed 29 March 2018): <http://www.regeringen.se/rattsdokument/departementsserien-och-promemorior/1999/01/ds-199936/>
- Molitor J, Papathomas M, Jerrett M and Richardson S (2010) Bayesian profile regression with an application to the National Survey Of Children’s Health. *Biostatistics* 11: 484–498.
- Newman O (1972) *Defensible Space – Crime Prevention Through Urban Design*. New York: Collier Books.
- NYSCASA [New York State Coalition Against Sexual Assault] (2007) *A Summary of New York State Penal Code 130 Sex Offenses as of 2008*. Albany, NY: NYSCASA.
- Osgood D (2000) Poisson-based regression analysis of aggregate crime rates. *Journal of Quantitative Criminology* 16: 21–43.
- Pyle G (1974) *The Spatial Dynamic of Crime*. Chicago: Department of Geography, University of Chicago.
- RAINN [Rape, Abuse & Incest National Network] (2014) Was I raped? URL (accessed 29 March 2018): <http://www.rainn.org/>
- Rhodes W and Conly C (1981) Crime and mobility: An empirical study. In: Brantingham PJ and Brantingham PL (eds) *Environmental Criminology*. Beverly Hills, CA: Sage, 167–188.
- Rosay A and Langworthy R (2003) *Descriptive Analysis of Sexual Assaults in Anchorage, Alaska*. Anchorage: Alaska Justice Statistical Analysis Center, Justice Center, University of Alaska.
- Sampson RJ and Groves WB (1989) Community structures and crime: Testing social disorganisation theory. *American Journal of Sociology* 97: 774 – 802.
- Shaw CR and McKay HD (1942) *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.
- Sherman L, Gartin P and Buerger M (2006) *Hotspots of Predatory Crime: Routine Activities and the Criminology of Place*. New York: Wiley.
- Spiegelhalter DJ, Abrams KR and Myles JP (2004) *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. New York: Wiley.
- Statistik Stockholm (2012) Områdesfakta. URL (accessed 29 March 2018): <http://statistik.stockholm.se/omradesfakta>
- Tewksbury R and Mustaine EE (2006) Where to find sex offenders: An examination of residential locations and neighbourhood conditions. *Criminal Justice Studies* 19: 61–75.
- Uittenbogaard AC and Ceccato V (2012) Space–time clusters of crime in Stockholm. *Review of European Studies* 4: 148–156.
- UN Women (2014) URL (accessed 29 March 2018): <http://www.unwomen.org/en/news/stories/2014/11/un-women-and-new-york-city-sign-agreement-to-enhance-safety-and-empowerment-of-women>
- Wakefield J (2013) *Bayesian and Frequentist Regression Methods*. New York: Springer.
- Weisburd D, Goff E and Yang S (2012). *The Criminology Of Place: Street Segments and Our Understanding of the Crime Problem*. Oxford: Oxford University Press.
- Wheeler D and Tiefelsdorf M (2005) Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems* 7: 161–187.
- Zelinka A and Brennan D (2001) *Safescape: Creating Safer, More Liveable Communities Through Planning and Design*. Washington, DC: Planners Press.

## Appendix A

Table 2.

Data type	Description	Variable type	Source	Year	Transformation for normality <sup>a</sup>
Police-recorded data (according to Chapter 6, §1)	Code 0660: Rape of women 18 years or older, outdoors	Counts	Stockholm Police Headquarters	2008–2009	
	Code 0648: Attempted rape of women 18 years or older, outdoors				
	Code 0656: Rape, of girls 15–17 years old, outdoors				
	Code 0644: Attempted rape of girls 15–17 years old outdoors				
	Number of street robberies	Count/ categorical			
Demographic, socio-economic and land use data for <i>basområde</i>	Population turnover	Continuous	Stockholm city	2008	Log <sup>b</sup>
	Young female population	Continuous	Stockholm city	2008	Square root Log <sup>b</sup>
	Average annual income (in 10,000 Kronor)	Continuous	Stockholm city	2009	
	Presence of industrial land use in area	Binary	Stockholm city	2008	

(Continued)

**Table 2. (Continued)**

Data type	Description	Variable type	Source	Year	Transformation for normality <sup>a</sup>	
Geographical data	Proportion who avoid going out in the neighbourhood	Continuous	Stockholm Safety Survey	2010	None	
	Presence of subway station in area	Binary	SL – Stockholm Public Transport company	2009		
	Number of alcohol outlets in an area	Count/ categorical	Eniro	2010		
	Presence of forested area	Binary	Stockholm city – Stadsbyggnadskontoret	2008		
	Presence of industrial area	Binary	Stockholm city – Stadsbyggnadskontoret			
	Presence of school(s)	Binary	Stockholm city – Stadsbyggnadskontoret			
	City centre	Binary	Own calculation	2013		
	Small area units ( <i>basområde</i> ) for Stockholm city	Topological	Stockholm city – Stadsbyggnadskontoret	2008		

Notes: <sup>a</sup> Under the profile regression model, each continuous risk factor is required to follow approximately a normal distribution since continuous risk factors are jointly modelled using a mixture of multivariate normal distributions.

<sup>b</sup> There are 22 and 44 *basområdes* with 0 values for average income and population turnover and the log transform was not applied to these 0 values.

## Appendix B

```
##### WinBUGS codes to implement the Poisson regression model (Model 1 in Table 1)
model {
  for (i in 1:N) {
    O[i] ~ dpois(mu[i])
    log(mu[i]) <- log_mu[i]
    log_mu[i] ~ dnorm(mu_log_mu[i], tau_v)
    mu_log_mu[i] <- alpha + beta[1]*Women[i] + beta[2]*Education[i] + beta[3]*PopTurnover[i] + beta[4]*Income[i] + beta[5]*FearOfCrime[i]
      + beta[6]*Industry[i] + beta[7]*Forest[i] + beta[8]*CityCentre[i] + beta[9]*Subway[i]
      + beta_alcohol[Alcohol[i]] # incorporating number of alcohol outlets as a categorical covariate (with 4 levels and level 1 as reference)
      + beta_robbery[Robbery[i]] # incorporating number of street robbery cases as a categorical covariate (with 4 levels and level 1 as reference)
      + u[i]
  }

##### recovering the spatially-unstructured random effects v_i
  v[i] <- log_mu[i] - mu_log_mu[i]
}

##### the improper uniform prior on the whole real line for the intercept
alpha ~ dflat()

##### vague priors for the regression coefficients for continuous and binary covariates
for (i in 1:9) {beta[i] ~ dnorm(0,0.0001)}

##### priors for the regression coefficients for the two categorical covariates
beta_alcohol[1] <- 0 # set the reference category
beta_robbery[1] <- 0 # set the reference category
for (k in 2:4) {
  beta_alcohol[k] ~ dnorm(0,0.0001)
  beta_robbery[k] ~ dnorm(0,0.0001)
}
```

```

#### intrinsic conditional autoregressive model for u_i
u[1:N] ~ car.normal(adj[],weights[],num[],tau_u)

#### priors for the random effect standard deviations
tau_u <- pow(sigma_u,-2)
sigma_u ~ dnorm(0,1)I(0,)
tau_v <- pow(sigma_v,-2)
sigma_v ~ dnorm(0,1)I(0,)
}

```

Estimation of parameters was performed in WinBUGS using MCMC algorithms. Two MCMC chains with different starting values were run for a total of 30,000 iterations. The first 10,000 iterations were discarded as burn-in and every 20th iteration from the remaining 20,000 iterations was saved for inference. Thus 2000 iterations in total, from both chains, were used for inference. Both the Brooks-Gelman-Rubin (BGR; Brooks and Gelman, 1997) statistic and the history plots show convergence of the two chains, and visual inspection of the parameter history plots demonstrates good mixing.