J Quant Criminol (2010) 26:113–138
DOI 10.1007/s10940-009-9086-6

ORIGINAL PAPER

Modeling Micro-Level Crime Location Choice: Application of the Discrete Choice Framework to Crime at Places

Wim Bernasco

Published online: 22 December 2009

Abstract Discrete choice recently emerged as a new framework for analyzing criminal location decisions, but has thus far only been used to study the choice amongst large areas like census tracts. Because offenders also make target selection decisions at much lower levels of spatial aggregation, the present study analyzes the location choices of offenders at detailed spatial resolutions: the average unit of analysis is an area of only 18 residential units and 40 residents. This article reviews the discrete choice and spatial choice literature, justifies the use of geographic units this small, and argues that because small spatial units depend strongly on their environment, models are needed that take into account spatial interdependence. To illustrate these points, burglary location choice data from the Netherlands are analyzed with discrete choice models, including the spatial competition model.

Keywords Discrete choice · Spatial competition · Sampling-from-alternatives · Burglary

Discrete choice has recently emerged as a new framework for analyzing crime location decisions. The discrete choice framework was originally developed in economics as a model of individual choice behavior based on the principle of random utility maximization. Since the original derivation of the 'workhorse' of the discrete choice framework, the multinomial logit model (McFadden 1974), various more flexible generalizations of this model have been developed.

Since its initial application to travel demand theory, the discrete choice framework has been applied in a number of disciplines and subject fields to analyze a great variety of choice situations, including location choice (spatial choice). Some examples of applications to location choice are the consumer's choice of a shopping center (Bhat and Zhao 2002; Fotheringham 1985), the angler's choice of a fishing site (Shaw and Ozog 1999;

W. Bernasco (✉)
Netherlands Institute for the Study of Crime and Law Enforcement (NSCR),
P.O. Box 71304, 1008 BH Amsterdam, The Netherlands
e-mail: wbernasco@nscr.nl

Train 1998), and the migrant's choice of a new residential area (Boots and Kanaroglou 1988; Duncombe et al. 2001).

The application of the discrete choice framework to crime location choice—the offender's choice of where to commit an offense—is relatively new. It consists of only five published studies (Bernasco 2006, forthcoming; Bernasco and Block 2009; Bernasco and Nieuwbeerta 2005; Clare et al. 2009).

These five studies all explore offenders' choices among relatively large geographical areas—census tracts, neighborhoods or suburbs with populations of 3,500 or more. Such units of analysis are much larger and more heterogeneous than the areas usually referred to in contemporary theoretical notions of offender behavior and target attractiveness, and also much larger than what is useful from the perspective of crime prevention (Weisburd et al. 2009a).

When push comes to shove, the concrete act of burglary involves the choice of just a single residential unit to break into. But before deciding to target that specific home, a prospective burglar has to implicitly or explicitly pass along several branches of a spatially structured decision tree: Which neighborhood in the city? Which street in the neighborhood? Which block on the street? Consequently, burglary location choice can be investigated at various levels of spatial aggregation. Because the abovementioned applications of the discrete choice framework to burglary location choice have explored choices amongst large areas, this article applies the framework at a much more detailed spatial resolution. The empirical analysis distinguishes small residential areas approximately the size of a football field, containing only 18 residential units and 40 residents on average. While these units are somewhat larger than a single residential unit, their geographical scale matches the scale at which many theories of crime at places are articulated, and it may also appeal to crime prevention practitioners.

The focus on small spatial units requires consideration of the *aspatial nature* of the multinomial logit model, and the utilization of the spatial competition model. The spatial competition model is not saddled with the *independence of irrelevant alternatives* (IIA) property, a property that has been viewed as the Achilles heel of the multinomial logit model because its implications can be counterintuitive.

The remainder of this article consists of seven sections. The first section provides a justification for the use of small spatial units in the study of crime in general and burglary in particular, and argues that the focus on small spatial units requires that the influence of the spatial context be taken into account. The second section reviews the literature on the discrete choice framework. The third section specifically discusses discrete modeling of spatial choice. The fourth section briefly reviews the five abovementioned studies, as well as other ways to model the journey to crime. The fifth section presents and analyzes micro data on burglary location choice of 1,871 burglaries in the Greater The Hague Area. Special attention is paid to the technique of *sampling of alternatives* to guarantee the computational tractability of discrete choice models with large choice sets. The sixth section presents the findings, and the seventh section summarizes the advantages of the discrete choice model, addresses the theoretical challenge of how to represent the spatial choice set of burglars, and discusses challenges for future exploration.

Micro Places in the Geography of Crime

In this section, a justification is provided for using small spatial units of analysis in the application of the discrete choice model to the analysis of crime location choice in general and to burglary location choice in particular. Subsequently, it is argued that the use of small spatial units implies a potentially greater role for spatial interaction between observational units, and thus requires the use of statistical methods that are able to capture this spatial interaction.

Why Small Spatial Units of Analysis?

In the abovementioned five applications of the discrete choice framework to crime location choice, relatively large spatial units were used, census tracts, suburbs, and neighborhoods with average residential populations ranging between 3,500 and 5,000 residents. These areas are much larger and much more heterogeneous than the areas usually referred to in contemporary theoretical notions of offender behavior and target attractiveness, and also much larger than what is useful from the perspective of crime prevention.

Many contemporary scholars in geographic criminology advocate the use of much smaller spatial units of analysis (Groff et al. 2009; Oberwittler and Wikström 2009; Smith et al. 2000) and prior research has indeed identified census blocks, street segments (Weisburd et al. 2004), street blocks (Kurtz et al. 1998; Taylor 1997), street corners (McCord and Ratcliffe 2007) and even single parcels (Kinney et al. 2008) or addresses (Groff and La Vigne 2001; Sherman et al. 1989) as meaningful spatial units of analysis. But why precisely would we need smaller spatial units? There are various reasons that underlie the plea for small units.

The first reason to favor small spatial units is the generic principle in all sciences that to understand the characteristics and behavior of an object of inquiry, it is helpful to understand the characteristics and the behavior of its constituting elements, as well as the nature of the relations between these constituting elements.

Another reason to favor small units is that the relevant boundaries between units may not be known to the researcher. The definition of spatial objects, such as neighborhoods, is often a result of arbitrary administrative decisions, but may not correspond to the perceptions of the area by those who use it (Coulton et al. 2001). In the absence of well-defined boundaries between spatial units, the measurement of small entities is to be preferred, even if the mental maps of the users are less fine-grained. In that case the small entities are just replicated observations, whereas all relevant differences between small entities will remain unobserved if the entities are aggregated before measurement.

An empirical justification of the use of small spatial units of analysis is the observation that larger spatial units are often heterogeneous with respect to crime itself and many related variables. One study showed that only three and a half percent of the addresses in Minneapolis generated half of all calls for service to the police (Sherman et al. 1989). In another study (Weisburd et al. 2004), it was found that less than 5% of the street segments in Seattle accounted for half of all crime incidents. In a recent ethnographic study (St. Jean 2007), it was noted that even within a single high-crime Chicago police beat comprising 3386 residents, crimes occurred much more frequently on some neighborhood blocks then on others that remained free of crime.

These skewed distributions imply that as the spatial unit of analysis decreases, the *variation within* units of analysis decreases, which results in more precise measurement, and both the *number of units* and the *variation between* the units of analysis increase, which results in an improved capability to detect differences (statistical power).

Ultimately, the main criterion for choosing a particular unit of analysis is the theory to be tested. The unit of analysis should match the theory (Weisburd et al. 2009b). The ecological fallacy (Robinson 1950) is the reasoning error by which conclusions about small units are made on the basis of empirical findings at higher levels of aggregation. The best way to prevent this fallacy is to measure and analyze the data at units of analysis that match the theory.

Many theoretical notions that relate to spatial crime patterns, including propositions about offender behavior, informal social control and the nature of crime settings, apply to units of analysis that are arguably much smaller than neighborhoods, census tracts or block groups. Social disorganization theory, for example, posits that lack of social control and lack of cohesion in an area make it difficult for residents to take action against crime in their area. Social control and cohesion involve processes (e.g., knowing each other, keeping an eye on each other's properties or children) that take place between close neighbors, but not between people who live many blocks apart. For this reason, Taylor (1997) argued that the street block be considered the appropriate spatial level for testing social disorganization theory. Street blocks fit many of the criteria of what ecological psychologists refer to as 'behavior settings': small scale natural units of interaction between people and objects that are characterized by predictable and habitual patterns of behavior and are contained within a place that is often physically bounded (Barker 1968; Wicker 1987). An elegant definitional solution to how small or large a place must be to be considered a setting (is the setting a school, or rather a classroom?) is Wikström's definition of a setting as "the social and physical environment (objects, persons, and events) that the individual, at a particular moment in time, can access with his senses (e.g., what he can see, hear and feel)" (Wikström 2006: 86–87).

Spatial Units of Analysis in the Study of Burglary

With respect to burglary, additional and more specific arguments for the use of small spatial units apply. First of all, the living quarters of a single household, family, or extended family, i.e., the place that they call their home and which is here referred to as a 'residential unit', forms a natural smallest unit of analysis. A residential unit is demarcated by symbolic and physical barriers that restrict access to others than those who live in it (and usually spend the night in it). All over the world, whether applied to apartments, houses, tents or castles, jurisdictions define residential burglary (or domestic burglary) as the forced entry into a private place 'where residents sleep', with the intent to steal and without the permission of the owner or resident (Bernasco 2009a).

Because legal definitions of burglary refer to buildings or parts of buildings as the objects of burglary, and because the boundaries of these objects are usually clearly indicated by walls, windows, locked doors, fences and signs, the residential unit has since long and without much dispute been viewed as the natural unit of analysis in research on burglary. The relative lack of ambiguity in the definition of a natural spatial unit of analysis in burglary is strengthened by the fact that residential units are normally not mobile, expect for ships, houseboats and mobile homes. Thus, unlike crimes with mobile targets, such as assault or car theft, in residential burglary the characteristics of the target coincide with the characteristics of the spatial unit of analysis.

Although the residential unit is a natural unit of analysis in burglary research, some theories suggest that the occurrence of burglary is influenced by mechanisms at higher levels of aggregation than the residential unit itself. An example is social disorganization theory, as discussed above. For this reason, most contemporary empirical research on burglary victimization uses multilevel modeling to estimate contextual effects of characteristics of the environment (e.g., street block, neighborhood) in addition to effects of characteristics of residential units and residents (e.g., Markowitz et al. 2001; Wilcox et al. 2007; Zhang et al. 2007).

Burglary target selection has been described as a spatially structured hierarchical process (Brantingham and Brantingham 1978; Brown and Altman 1981; Cornish and Clarke 1986), in which the offender first selects a neighborhood or equivalent geographic area, and subsequently targets a specific street block in the chosen area, and a specific house in the chosen street block. In fact, once the burglar is inside the house, the rooms to be searched or the items to be stolen can be viewed as the actual targets. Because of this concept of spatially hierarchical decision-making, all these spatial levels (i.e., areas, street blocks and houses) can be referred to as burglary targets, and be studied at these various levels of spatial aggregation.

In support of the idea of spatially structured hierarchical decision-making, accounts of burglars indicate that their considerations and evaluations do apply to various levels in a spatial hierarchy. They distinguish between large areas such as neighborhoods (Rengert and Wasilchick 2000) and evaluate them on the basis of properties of affluence, ethnic composition and other general characteristics of the resident population. They are also able to distinguish between targets at disaggregated levels of geography, such as face blocks and individual properties, and assess their attractiveness as burglary targets. In some studies offenders have been shown photographs of individual houses asking them to rate these houses in terms of attractiveness as burglary targets (Logie et al. 1992; Taylor and Nee 1988; Wright et al. 1995). Taylor and Nee (1988), for example, demonstrated that when burglars were shown photographs of houses, they assessed the attractiveness of these houses as burglary targets on the basis of home layout, visible signs of wealth, occupancy, presence of escape routes and degree of security.

Spatial Interaction Between Small Spatial Units

A complication in the analysis of spatial units is the presence of spatial interaction. Spatial interaction is present when events or situations in a spatial unit affect events or situations in a nearby or adjacent spatial unit. The smaller the spatial unit of analysis that is chosen for analyzing a phenomenon, the more plausible it is that the observed unit is affected by its spatial environment. The reason is that the environment of small spatial units is larger, relative to the size of the unit of analysis, than the environment of large spatial units.

Figure 1 illustrates this point graphically. It displays three partial maps of a city street grid, highly stylized for ease of exposition, where each small square represents a block. Consider a situation in which there is spatial interaction that spans the length of one block, but not beyond. Thus, an event in a block affects events in that block and affects events in all adjacent blocks, but not in blocks farther away.

In the left map, the unit of analysis chosen for the analysis is a single block. The ring of blocks adjacent to it consists of eight blocks. If we assume that spatial interaction spans the length of one block, the environment from which the block in the center receives influence consists of eight blocks, which equals eight times the size of the unit of analysis.

In the middle map, the unit of analysis is four blocks. Assuming again that spatial interaction spans the length of a single block, the environment from which the 4-block unit in the center receives influence consists of 12 blocks, which equals three times the size of the unit of analysis.

In map on the right size, the unit of analysis comprises 9 blocks, and the ring of influence around it consists of 16 blocks, which amounts to only 1.8 times the size of the unit of analysis. Figure 1 demonstrates that as the spatial unit of analysis decreases, the

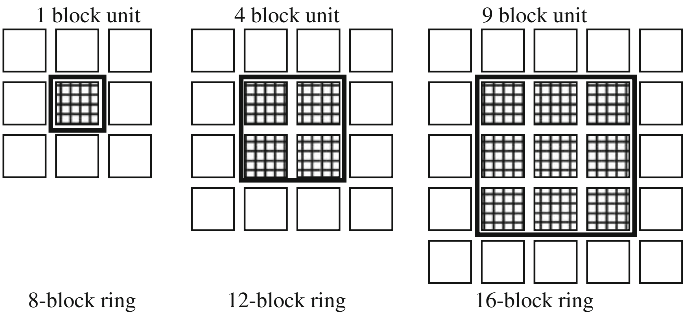

Fig. 1 As unit size increases from 1 to 4 to 9 blocks, size of block ring around unit relative to unit size decreases from 8 to 3 to 1.8

importance of the local environment increases. As a consequence, in the explanation of social phenomena, it generally holds that the smaller the unit of analysis, the more urgent is the need to consider the presence of spatial interaction.

While spatial effects are increasingly recognized and studied in the analysis of crime and criminal justice issues (Andresen 2006; Baller et al. 2001; Deane et al. 2008; Kubrin 2003; Mears and Bhati 2006; Morenoff et al. 2001), they have thus far not been used in applications of the discrete choice framework to crime location choice. The discrete choice framework proper is aspatial, because the spatial relations between the alternatives are not specified. Therefore, in modeling an offender's evaluation of potential crime locations, these locations have been analyzed in isolation from their spatial environment. To correct this situation, the analysis in this article will apply the spatial competition model (Fotheringham et al. 2000; Pellegrini and Fotheringham 2002), a model that is able to estimate the extent to which the choice of a spatial alternative is influenced by the presence and the size of nearby alternatives.

Discrete Choice Framework

The discrete choice framework is a set of assumptions and methods to model a decision-maker's choice among a set of alternatives that are mutually exclusive and collectively exhaustive. Most of the assumptions are based on random utility maximization (RUM) theory (McFadden 1973). In essence, RUM theory is the micro-economic theory of consumer behavior with a random component added to the utility function. By making certain assumptions about the distribution of this random component, the theory is directly linked to a statistical model that allows probabilistic statements to be formulated and tested.

The discrete choice framework was developed in the 1970s by McFadden and others working in the field of travel demand, and the first applications of discrete choice were in the study of travel mode choice (i.e., the choice between train, bus, car, or airplane). Later the model was also applied to the choice of a travel routes and travel destinations (Ben-Akiva and Lerman 1985).

The discrete choice framework consists of a set of assumptions regarding four elements of a choice situation (Ben-Akiva and Bierlaire 1999):

(1) Decision-makers. The decision-maker is the person or agent that makes a choice.

(2) Alternatives. The decision-maker must choose one alternative from the choice set, i.e., the set of available alternatives that are mutually exclusive and collectively exhaustive.
(3) Attributes. Alternatives have attributes that make them attractive to the decision-maker. The decision-maker evaluates the attractiveness of all alternatives.
(4) Decision rule. In the random utility maximization model, the decision-maker chooses the alternative that maximizes her (expected) utility (net gain, profits, satisfaction).

Most applications of discrete choice models are based on the premise of 'revealed preference', according to which the preference of the decision-maker is *measured* by the behavioral outcome of his or her choices. An alternative is the 'stated preference' approach, in which the preference is measured by psychometric methods, such as questionnaire items (Hensher and Bradley 1993).

In discussing discrete choice modeling, I will mostly follow the notation of Train (2009) as well as his decision to refer to the decision-maker as a 'she' and to the researcher as a 'he'.

A decision-maker, labeled n, must make a choice among J alternatives. Decision-maker n obtains a level of utility (profits, satisfaction) U_{ni} from alternative i if that alterative is chosen. The principle of utility maximization asserts that the decision-maker decides in favor of the alternative i if and only if she expects to derive more utility from alternative i than from any other available alternative. Thus, if she decides in favor of alternative i, she must expect to derive less utility from each of the other alternatives.

$$U_{ni} > U_{nj} \forall j \neq i. \tag{1}$$

The utilities are known by the decision-maker, but not by the researcher. The researcher only observes the J alternatives, some attributes a_{ni} of the alternatives, some attributes d_n of the decision-maker, and he can specify a function V, often called *representative utility* or *systematic utility*, that links these observed attributes to the decision maker's utility:

$$V_{ni} = V(a_{ni}, d_n) \forall i \tag{2}$$

The researcher incompletely observes utility, so that generally $U_{nj} \neq V_{nj}$. The utility can be written as the sum of representative utility V_{nj} and a term ε_{nj} that captures the factors that determine utility but are not observed by the researcher, and that is treated as random.

$$U_{ni} = V_{ni} + \varepsilon_{ni} \tag{3}$$

The probability that decision-maker n chooses alternative i is the probability that the utility associated with choosing i is greater than the utility associated with any other alternative in the choice set:

$$\begin{aligned} P_{ni} &= \Pr(U_{ni} > U_{nj} \forall j \neq i) \\ P_{ni} &= \Pr(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \forall j \neq i) \\ P_{ni} &= \Pr(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i) \end{aligned} \tag{4}$$

This is the most general formulation of the discrete choice model, and any specific choice model that is consistent with random utility maximization can be derived from specific assumptions on the joint distribution of the unobserved utility term ε_{ni}.

Multinomial Logit Model

If the unobserved random utility components ε_{ni} are independent and identically distributed according to an extreme value distribution (also referred to as Gumbel distribution), the *multinomial logit model* can be derived, which was originally labeled the *conditional logit model* (McFadden 1974) and is also simply referred to as 'logit model' in the discrete choice literature. In the multinomial logit model, the choice probability P_{ni}, the probability that decision-maker n chooses alternative i, is given by:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_{j=1}^{J} e^{V_{nj}}} \tag{5}$$

For computational convenience, and because any function can be closely approximated by a linear function, representative utility V_{ni} is usually assumed to be linear in the parameters.

$$P_{ni} = \frac{e^{\beta' x_{ni}}}{\sum_{j=1}^{J} e^{\beta' x_{nj}}} \tag{6}$$

In Eq. (6), β' is a vector of parameters that can be estimated when the actual choices made have been observed, and x_{ni} is a vector that describes observed attributes that vary across alternatives. In model (6), effects of attributes of the decision-maker, which do not vary across alternatives for the decision-maker, for example age, gender or income, cannot be estimated. This is easily demonstrated by inserting a variable z_n in (6) that varies only across decision-makers.

$$P_{ni} = \frac{e^{\beta_1 x_{ni} + \beta_2 z_n}}{\sum_{j=1}^{J} e^{\beta_1 x_{nj} + \beta_2 z_n}} = \frac{e^{\beta_1 x_{ni}} e^{\beta_2 z_n}}{\sum_{j=1}^{J} e^{\beta_1 x_{nj}} e^{\beta_2 z_n}} = \frac{e^{\beta_1 x_{ni}}}{\sum_{j=1}^{J} e^{\beta_1 x_{nj}}} \tag{7}$$

Variable z_n and parameters β_2 cancel out of the equation. The characteristic of the decision-maker cannot affect which alternative is chosen because it does not vary across the alternatives. Still, it can be interacted with characteristics that vary across alternatives, to test the hypothesis that effects of attributes of alternatives vary across decision-makers (e.g., that the cost of an alternative is less restrictive for affluent decision-makers than for others).

Independence of Irrelevant Alternatives

One of the most widely criticized aspects of the multinomial logit model is the independence of irrelevant alternatives (IIA) property. It has been viewed as the Achilles heel of the multinomial logit model, and the implausibility of the implications of this property has been the major impetus for the development of RUM consistent discrete choice models without this property.

In a general context, the IIA is an axiom that states that if alternative A is preferred to B if the choice set is (A, B), then the introduction of another alternative C in the choice set must not make B preferable to A. In the discrete choice framework it holds that for a specific decision-maker the ratio of the choice probabilities of any two alternatives is entirely unaffected by the representative utilities (V_{ni}) of any other alternatives (Ben-Akiva and Lerman 1985). This can be easily seen by substituting Eq. (5) into the ratio of P_{ni} and P_{nk}. Because both denominators cancel,

$$\frac{P_{ni}}{P_{nk}} = \frac{e^{V_{ni}}}{e^{V_{nk}}} = e^{V_{ni} - V_{nk}} \tag{8}$$

Thus this ratio, the relative odds of choosing i over k, does not depend on the availability or on the attributes of other alternatives, which because of this property can be seen as *irrelevant* alternatives.

As a choice axiom, the IIA property has intuitive appeal. For example, one normally prefers coffee over tea, or vice versa, independently of the other drinks available. An example used to illustrate how the IIA can lead to a paradox is the red-bus/blue-bus example (McFadden 1974). Consider a commuter's choice between two modes of transportation, which initially are 'car' with probability P_c and 'red bus' with probability P_r. Suppose that commuter chooses between these two options with equal probability so that $P_c/P_r = 1$. Now suppose a third mode, a blue bus with probability P_b, is added. According to the IIA property, the ratio of the probabilities P_c/P_r is preserved. Because P_c, P_r and P_b sum to 1, if the blue bus has a non-zero choice probability, both P_c and P_r must decrease by the same amount. Intuitively, this is implausible because it does not take account of the fact that red buses and blue buses are very similar, and are likely to be perfect substitutes. More formally, the reason for this implausibility lies in the assumption that the unobserved utility components ε_{ni} in Eq. (3) are assumed to be independent (Ben-Akiva and Lerman 1985). In the case of red buses and blue buses this assumption is implausible because these two alternatives share all the unobserved characteristics of buses (i.e., characteristics not included in representative utility V_{ni}). Thus, the extent to which the IIA property is plausible cannot be judged in the abstract, but depends crucially on whether the systematic component of utility in the model captures all relevant choice criteria.

As will be discussed shortly in the section on spatial choice, the IIA property may be overly restrictive, like it is in the red bus/blue bus example, in applications of the multinomial logit model to spatial choice, where nearby places not only are close substitutes because there are situated in proximity of each other but also because as a rule nearby places have more characteristics in common than distant places do.

Before addressing spatial choice, I will briefly discuss a number of alternative discrete choice models that have been developed to alleviate the restrictive IIA property by allowing more flexible patterns of correlations amongst the unobserved random elements of utility. These include a family of generalized extreme value (GEV) models, the multinomial probit model and the mixed logit model, which have all been discussed in detail elsewhere (Train 2009).

Other Discrete Choice Models

Generalized extreme value (GEV) models are a class of models that allow a variety of substitution patterns between the unobserved random utility components ε_{nj} of the standard multinomial logit model.

The most widely used member of the GEV family is the nested logit model. In the nested logit model, the choice set is partitioned into subsets, called *nests*. Each nest may in turn also be partitioned into sub-nests. This model is typically interpreted as representing a hierarchical choice structure. For example, a decision-maker decides on a type of restaurant, and conditional on the chosen type decides on the specific type of food to order. In the nested logit model, the unobserved utility components ε_{nj} are from an extreme value distribution, but they are not completely independent. The ε_{nj} are allowed to be correlated within nests, but are still assumed to be uncorrelated between nests.

In the application to location choice, the nested logit model has been proposed as a model of hierarchical spatial choice. Thus, although mathematically the single defining characteristic of the model is its structure of interdependencies of the unobserved error component, the model is interpreted in terms of sequential hierarchical decision-making: in the first phase, the decision-making chooses a nest from the set of available nests, while in the second phase she chooses a final location from the nest chosen in the first phase. Just like the multinomial logit model it does not assume that all decision-makers have the same choice set, and in addition the nested logit model does not assume that the nesting structure is the same for all decision-makers that do have the same choice set. Two alternatives may be in the same nest for one decision-maker, but in another nest for other decision-makers.

Other generalized extreme value models include the generalized nested logit model (in which each alternative can be a member of more than one nest), the paired combinatorial logit model (in which each pair of alternatives is considered to be a nest), and the heteroskedastic logit model (in which the variance of the unobserved utility components ε_{nj} is allowed to differ across alternatives).

The multinomial probit model is even more flexible than the GEV family of models, as is not only allows substitution patterns between alternatives, but also can handle random taste variation (i.e., the situation where the β parameters in Eq. (7) vary across decision-makers) and panel data with temporally correlated errors. It is based on the assumption that unobserved utility components follow a Gaussian distribution. Like the multinomial probit model, the mixed logit model allows random taste variables, unrestricted substitution and correlation of the unobserved utility component over time, but is not restricted to normal distributions and is computationally simple.

A discussion of these and other models can be found in Train (2009). The point to be taken from this discussion is that the main purpose of the application of these models to choice problems has been the need to alleviate the IIA property, as it is seen as much too restrictive, and that this is realized through alternative assumptions regarding the distribution of the unobserved utility component.

The models described above have all been developed for use with 'revealed preference' data, where the actual choice of an alternative by the decision-maker is observed. Multiple-indicator-multiple cause (MIMC) models integrate the discrete choice framework with structural equations modeling, by including latent variables, typically in the form of 'stated preference' data and psychometric data on beliefs and attitudes (McFadden 2001).

Spatial Choice

Spatial choice is the choice of a destination from a choice set of spatial alternatives. For example, a consumer can choose a shopping centre from a set of shopping centers at various locations in the city where she lives, or a family can decide on where to move to, by choosing from a choice set of towns and cities in the region of their destination. As opportunities for crime are omnipresent, a criminal can choose from a wide variety of potential crime locations.

Within the discrete choice framework, spatial choice has two special features that other types of choice do not have. First, there is often uncertainty about the elements of the choice set because essentially space is not a discrete but a continuous variable. While space can easily be partitioned into discrete units, it is not obvious what the size of these units should be, and even with the size fixed there is an infinite number of ways in which space can be partitioned. The problem that analytical results depend on how continuous space is partitioned into units, is known in geography as the modifiable areal unit problem (MAUP; Openshaw 1984). In practice, in spatial choice the choice sets that are analyzed are usually much larger than in other choice problems, although very large choice sets are also encountered when analyzing travel routing decisions (Frejinger et al. 2009).

The second feature that distinguishes spatial choice from other types of choices is that spatial alternatives are part of a spatial structure, and are almost by definition very unlikely to be independent of each other. The first law of geography asserts that "everything is related to everything else, but near things are more related than distant things" (Tobler 1970: 236), and thus nearby places are more similar to each other than distance places, a relation that is known as *spatial autocorrelation* (Dubin 1998; Getis 2007; Legendre 1993). The empirical ubiquity of spatial autocorrelation in almost any physical or social characteristic, implies that in spatial choice, nearby spatial alternatives cannot be assumed to be independent, and may often be near substitutes when they are the object of choice. Thus, the independence of the unobserved utility components is often implausible in spatial choice, as is the independence of irrelevant alternatives (IIA) property that follows from it.

As an example, in the left panel (a) of Fig. 2, consider the choice of decision-maker *n* between the identical alternatives 1 and 4. As they are identical and situated at the same distance from the decision-maker, she should be indifferent between alternative 1 and 4. In practice, it may be plausible to assume that alternative 1 is preferred as it is located in the proximity of other identical destinations. On the other hand, alternative 4 may be preferred because it is more visible and distinct because of its isolated position. Thus, spatial choice situations suggest that spatial clustering of alternatives could affect preferences and thus violate the IIA property.

Similar variations could occur if not the locations of the alternatives but their size or attractiveness levels are clustered. For example, in the right panel (b) of Fig. 2, the spatial distribution is completely uniform (equally spread) and each alternative is located at the same distance from the decision-maker and at the same distance from all other alternatives. Nevertheless, alternative 1 is situated between two large or attractive alternatives (2 and 3), while alternative 4 is surrounded by regularly sized alternatives of average attractiveness. Again, on the basis of the characteristics of the alternatives and their locations in the spatial system, the decision-maker should be indifferent between alternatives 1 and 4. In practice, the preference for an alternative may be either increased or decreased by the proximity to large or attractive targets.

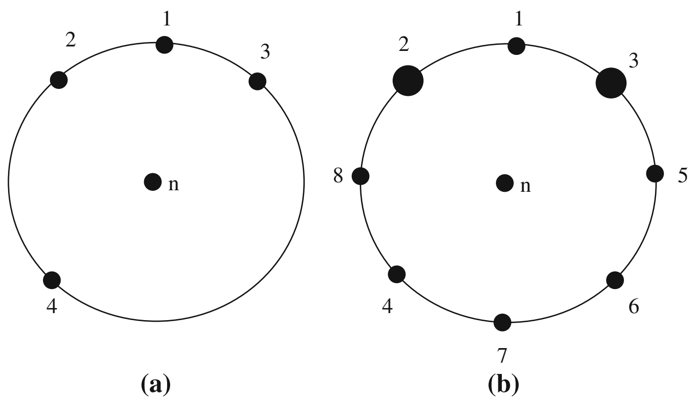

Fig. 2 Competing destinations in spatial choice

Nested Logit Model as a Spatial Model

A partial solution to alleviate the violation of independence due to spatial autocorrelation is the use of the nested logit model that was discussed above. By grouping nearby locations into the same nest, spatial autocorrelation within nests is accounted for (an example is Chattopadhyay 2000). This solution is very similar to the way that spatial autocorrelation is accounted for in hierarchical linear (multilevel) models of geographic data (Elffers 2003). The disadvantage of the nested logit model as a solution to the problem of interdependence of spatial choice alternatives is twofold.

The first disadvantage is that the nested logit model insufficiently captures spatial interdependencies between alternatives. It assumes that all spatial interdependence within a particular nest is equal. However, spatial autocorrelation would also hold between units within the nest, where its strength is expected to depend on within-nest distance. Further, the nested logit formulation ignores spatial interdependence between the nests. Thus, adjacent nests are still assumed to be independent, including near alternatives in different nests that are coincidentally separated by the geographical boundary between their nests.

The second disadvantage of the nested logit model is that it requires the analyst to know how the decision-maker groups the spatial alternatives, and also to know which alternatives are in each individual decision-maker's choice set.

As an alternative to the nested logit model Fotheringham (1983a, b) proposed a *competing destinations model*, a model that has found support in the location choice literature (Fik 1988; Fotheringham et al. 2001; Gitlesen and Thorsen 2000; Hu and Pooler 2002; Hunt et al. 2004; Pooler 1997).

Competing Destinations Model

Two justifications for the competing destinations model have been proposed, each referring to one of the two abovementioned disadvantages of the nested logit model. One justification refers to the incomplete pattern of spatial interdependence that is assumed in the nested logit formulation, the other emphasizes that the competing destinations model allows for a decision-maker who follows a hierarchical spatial choice strategy where the structure of the choice set is assumed to be unknown by the analyst, but to depend on the location of other alternatives.

The spatial competition model (also see Fotheringham et al. 2000; Pellegrini and Fotheringham 2002) allows for spatial decision-making to be spatially hierarchical, and is not saddled with the IIA property nor with the disadvantages of the nested logit model. In the model, a decision-maker n does not necessarily evaluate all choice alternatives, but a subset of alternatives, the restricted choice set M_n. This is expressed by replacing Eq. (6) with

$$P_{ni} = \frac{e^{\beta' x_{ni}} \ell_n(i \in M_n)}{\sum_{j=1}^{J} e^{\beta' x_{nj}} \ell_n(j \in M_n)} \tag{9}$$

where $\ell_n(i \in M_n)$ is the likelihood (the probability divided by a constant with respect to i) that alternative i is in the restricted choice set of the decision-maker n, and is thus evaluated by her.

There are various approaches on how to define $\ell_n(i \in M)$. Fotheringham (1983a, b) suggests that the likelihood that an alternative is in the restricted choice set depends on its dissimilarity to all other alternatives, such that the *accessibility* A_k of an alternative k to all other alternatives is a function of the 'size' S of these alternatives and their distances D from alternative k:

$$A_k = \sum_{j \neq k} \frac{S_j}{D_{jk}} \tag{10}$$

The justification for the inclusion of this competition variable, which can be seen as a measure of the local concentration or clustering of alternatives,[1] is the finding that individuals tend to underestimate the size of large objects (Fotheringham et al. 2000). Consequently, the size of spatial clusters of alternatives would be underestimated and the likelihood of an elementary alternative within such a cluster to be within the restricted choice set would be reduced. In other words, the alternative would be subject to more competition than other, more isolated alternatives. The reverse might as well be true. In that case an alternative becomes more likely to be in the restricted choice set if it is located in the proximity of competing destinations. In the retail literature this is called an 'agglomeration effect': firms may increase their sales by locating in the proximity of competitors. Therefore, the accessibility term is also parameterized, leading to the following competing destinations model:

$$P_{ni} = \frac{e^{\beta' x_{ni}} A_i^{\delta}}{\sum_{j=1}^{J} e^{\beta' x_{nj}} A_j^{\delta}} = \frac{e^{\beta' x_{ni} + \delta \ln A_i}}{\sum_{j=1}^{J} e^{\beta' x_{nj} + \delta \ln A_j}} \tag{11}$$

which allows for agglomeration effects ($\delta > 0$) as well as competition effects ($\delta < 0$).

Consider again, in the left panel (a) of Fig. 2, the probabilities of alternatives 1 and 4. Because alternative 1 is located nearby alternatives 2 and 3 and alternative 4 is not, $A_1 > A_4$ in Eq. (11), so that in the case of an agglomeration effect ($\delta > 0$) alternative 1 is more likely to be chosen than alternative 4, while in the case of a competition effect ($\delta < 0$) alternative 4 is more likely to be chosen than alternative 1. A similar reasoning applies in the right panel of (b) of Fig. 2, where $A_1 > A_4$ because the two larger alternatives 2 and 3 are closer to alternative 1 than to alternative 4.

The competing destinations model is not saddled with the IIA property, because the ratio of two choice probabilities P_i and P_k does depend on the location of each of them relative to all alternatives in the spatial structure that are captured in A_i and A_k, respectively. When another alternative becomes available, the ratio of the choice probabilities can change.

Note that the IIA property is not removed from the competing destinations model by way of assuming a more flexible structure of the unobserved utility term, but by bringing this source of interdependence between spatial alternatives—proximity to other alternatives—into the systematic utility V_{ni} and giving it a substantive interpretation.

Crime Location Choice Studies

While the discrete choice model has only recently been applied to the choice of a crime location, patterns of criminal movement had been studied before. Most research has only

[1] Note that the accessibility of an alternative is defined in terms of its distance to other alternatives, not in terms of its distance to an origin location.

focused on the *distance* between the offender's home and the crime location (e.g., Snook et al. 2005; Wiles and Costello 2000). It has been replicated across countries, across offender and across offense types that offenders tend to travel short distances, and that the number of crimes committed decreases as the distance to the home location increases (*distance decay*).

The spatial interaction model that preceded the discrete choice framework in geography has also been used to study aggregated 'flows of crime' in the urban environment. An early application by Smith (1976) used a variety of spatial interaction models to study crime flow patterns in Rochester, NY. Kleemans (1996) used the model to study burglary flow patterns in Enschede, the Netherlands. More recently, spatial interaction models were used to specifically test whether the presence of ethnic and social (Elffers et al. 2008; Reynald et al. 2008) or physical (Peeters and Elffers forthcoming) barriers reduce numbers of crime trips between origins and destinations.

The first application of the multinomial logit model to crime location choice analyzed the choice of a residential burglary destination neighborhood in The Hague, the Netherlands (Bernasco and Nieuwbeerta 2005). In the city 89 neighborhoods were distinguished with an average population of 4,944 residents. The authors found that burglars were generally attracted to neighborhoods nearby their home (and that this attraction was stronger for minors than for adults), to ethnically mixed neighborhoods (and that this attraction was stronger for non-native burglars than for native burglars) and to neighborhoods with high percentages of single-family dwellings. The affluence of the neighborhood, its residential mobility, and its proximity to the city center did not have significant effects.

In contrast to Bernasco and Nieuwbeerta, who selected solitary burglars, a follow-up study (Bernasco 2006) analyzed that same data with co-offending burglar groups included. Co-offending patterns introduce complexities in the multinomial logit model, because multiple co-offenders must naturally be treated as a single decision-maker, but have individual characteristics (place of residence, sex, age, ethnic origin) that complicate the analysis. Moreover, the criteria that enter the decision-making of a group may differ from those of solitary offenders. Bernasco did not find evidence for the hypothesis that the spatial choices of co-offending burglars groups differ from those of solitary burglars.

While the first two studies on burglary used the Euclidian distance between the burglars' homes and their potential destination areas as a measure of the effort needed to travel from the origin to the destination, a study of burglars in Perth, Australia focused on the role of physical barriers (such as rivers, and major road ways) that are expected to impede travel, and the role of connectors and connectors (such as train lines) that are expected to facilitate it (Clare et al. 2009). Thus, this study used much more sophisticated measures of travel time than the previous two discrete choice burglary studies.

In Chicago, the multinomial logit model was used to analyze the location choice of street robbers (Bernasco and Block 2009). The authors defined the 865 census tracts as the choice set, and showed that in addition to being attracted to tracts nearby their residence, robbers are also attracted to tracts characterized by the presence of illegal markets as well as legal retail activities, tracts with a high school and with low levels of collective efficacy. They also showed the presence of 'ethnic barriers' in addition to simple distance: African-American, Hispanic and White offenders preferred to offend in tracts where the majority of the population was of their own racial or ethnic origin.

The fifth application of the multinomial logit model is a study on the role of past residence in crime location choices of offenders that were charged with robbery, burglary, theft from vehicle or assault (Bernasco forthcoming). The author showed that offenders are not only attracted to target areas close to their current homes, but also to the environment of their past homes.

Data and Method

Data for the present analyses were provided by the Haaglanden (greater The Hague area) police force. The police force services a geographical area of 400 sq. km that houses a population of one million residents and 460,000 residential units. It comprises the city of The Hague, several smaller cities including Zoetermeer and Delft, and a number of smaller towns and villages.

From the police files all cases registered as residential burglary that took place in the period 2002–2007 were selected if at least one perpetrator was arrested for the incident (thus, uncleared cases are excluded), if the home address of the perpetrator and address of the burglary were both known and located inside the study area. Although 90% of the burglaries in the Netherlands are reported to the police, the burglary clearance rate in the Greater The Hague area was about 8% in the period covered. Thus, the burglaries included in the analysis amount to a small minority of all residential burglaries reported to the police. Admittedly, it is possible that the cleared cases are a selective sample of all reported cases, because the investigative process produces bias. For example, burglars caught in the act are overrepresented, as are burglaries witnessed by victims or bystanders (Coupe and Girling 2001; Coupe and Griffiths 1996). However, because the same research shows that is fairly difficult to actually predict which burglars are caught in the act or have their actions witnessed, there is sufficient randomness involved in the process of clearance to view the sample of cleared burglaries as a random sample from which tentative generalizations can be drawn. As a cautionary note it should be mentioned, because the distance between the burglar's home and the target address is an important explanatory variable, that it seems likely that offenders with previous burglary convictions are more likely to be arrested, especially if they are locals that live nearby, as they are the 'usual suspects' that the police may profile in trying to solve a burglary case. This implies some risk of overestimating the effect of distance.

The addresses of the burglars as well as the addresses of the burgled properties were geocoded using the Dutch postal code associated with the addresses. This postal code is the spatial unit of choice used in the subsequent analysis. Throughout the Netherlands there are about 435,000 postal code areas. On average, a residential postal code area in the region of the Netherlands studied here is roughly the size of a football field and contains 18 residential properties and 40 residents.

The postal code system was created with the facilitation of postal delivery services in mind. Therefore, a single postal code is nearly always on the same street, applies to adjacent properties, and is not subdivided by physical barriers that impede pedestrian or car transportation. Although the precise forms and sizes of postal code areas vary because cities, towns and villages in the Netherlands are not always built along grid-like patterns, the majority of postal codes refer to either one or both of the two block faces on both sides of a street between two intersections, which has in other countries been referred to as a street segment, street block or face block (Weisburd et al. 2004). An example of what would be a typical arrangement of postal codes is given in Fig. 3.

Typically, in small streets with little traffic both the odd and the even numbered sides of the street have the same postal code, while in larger streets with more traffic the even and the odd sides are in a different postal code. Thus, the boundaries between postal codes

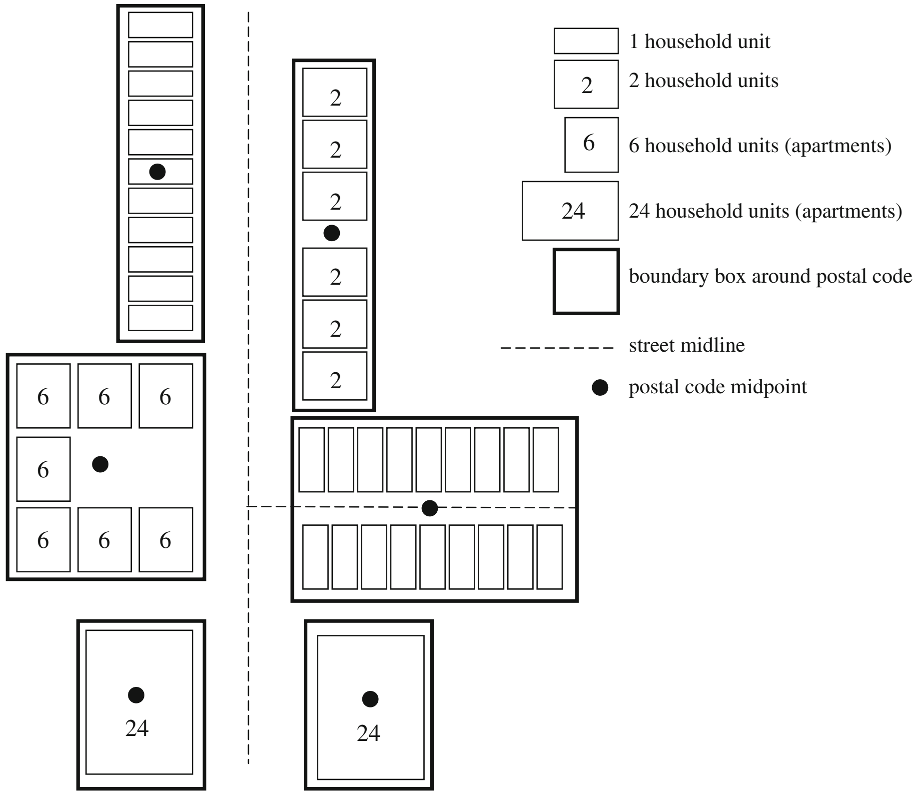

Fig. 3 Example spatial structure of postal codes

mostly align with the boundaries that impede interaction between residents, and typically all properties in the same postal codes can be observed from a single point. Nearly always a single postal code is reserved for residential units that share the same joint entrance in multi-unit dwellings (apartment blocks, blocks of flats). In sum, the qualifications of a behavior setting (Barker 1968; Wicker 1987) or setting (Wikström 2006) seem to fit the postal code area unit well.

The postal codes in the area comprise 18 addresses on average. The geographic position of a postal code is calculated as the geographic mean of these addresses. Figure 3 provides an example of a typical arrangement of properties in postal codes. In the figure, a bounding box is drawn around properties within the same postal code for the purpose of this exposition, but in geographic terms, the postal code is just a set of points of which the midpoint is known. It is not a lattice or, in terms of geographic information systems, a polygon.

As mentioned above, the study area (Greater The Hague Area) measures 400 sq. km, and houses a population of 1,000,000 residents in 460,000 residential units. This area comprises 26,214 postal code areas. For the present study on residential burglary, only the 23,984 postal codes areas (91.5%) were used that contain at least 3 residential units and at least 3 residents (the data provided are rounded to quintuples).

The available data on the attributes of postal codes is limited and refers mostly to the location, the population (number of residents, age, ethnic composition), and the residential units (number of properties and average value) in the area. The data are from a free public access data file of postal codes in the year 2004 constructed and provided by Statistics Netherlands. Table 1 describes the means standard deviations, minimum and maximum of the characteristics of the postal code areas used in this study.

We analyze the burglary location choice in 1,311 burglaries that involved 1,023 offenders. Of the 1,023 burglars, 755 (74%) participated in only one burglary in the period 2002–2007 and 268 (26%) took part in multiple burglaries (they were repeat offenders), 272 (27%) were juveniles (aged 12–17) and 751 were adults, 115 (11%) were female and 908 (89%) were male, 673 (66%) were born in the Netherlands and 350 (34% were born abroad). Altogether, there were 1,871 journey-to-burglary trips, i.e., combinations of the postal code area where an offender lived and the postal code area where the offence took place.

The distance between the origin and destination is a central issue in the journey to crime. Of the 1,871 journeys-to-crime analyzed (a single burglary by multiple co-offending burglars implies multiple journeys-to-crime), 58 (3.10%) were committed within the postal code area where the offender lived. Given that the average postal code area contains only 18 properties, these burglars almost literally victimize their neighbors. Figure 4 presents the distribution of the distance of the journey-to-crime, which displays a clear distance-decay function.

The spatial competition variable was constructed according to Eq. (10) with the number of residential units in the postal code areas as the size variable S.

Sampling from Alternatives

This section addresses the practical challenge of estimating the multinomial logit model in a situation where the choice set—the enumeration of all choice alternatives available to the decision-maker—is very large, in this case 23,984 residential postal code areas. While for many types of regression analysis—for example ordinary least squares, logistic, or negative-binomial regression—samples of this size and larger do not pose any problem on modern desktop computers, they do for multinomial logit regression analysis. The reason for the difference is that in other regression analyses of a large number of small spatial entities, the data matrix for estimation contains J rows, where J is the number of areas that are distinguished. Even the use of many areal units, such as the thousands of street

Table 1 Characteristics of greater The Hague postal code areas ($N = 23{,}984$)

Variable	Average	SD	Min	Max	Valid N
Residents (#)	39.97	24.47	5	545	23,984
Residential units (#)	18.49	10.03	5	205	23,984
Age 0–15 years (%)	15.99	10.85	0	62	23,398
Age 15–25 years (%)	11.55	9.15	0	100	23,398
Age 25–45 years (%)	30.91	15.29	0	100	23,398
Age 45–65 years (%)	25.55	12.68	0	90	23,398
Age 65+ (%)	16.00	18.85	0	100	23,398
Non-native population (%)*	18.60	20.24	2.5	60	23,398
Mean value property (€ 1,000)	133.46	104.47	14	1,683	23,134

* Data include percentages non-native population in five categories. To calculate overall percentage, midpoints of the categories (0–5, 5–10, 10–20, 20–40, 40–100) were used

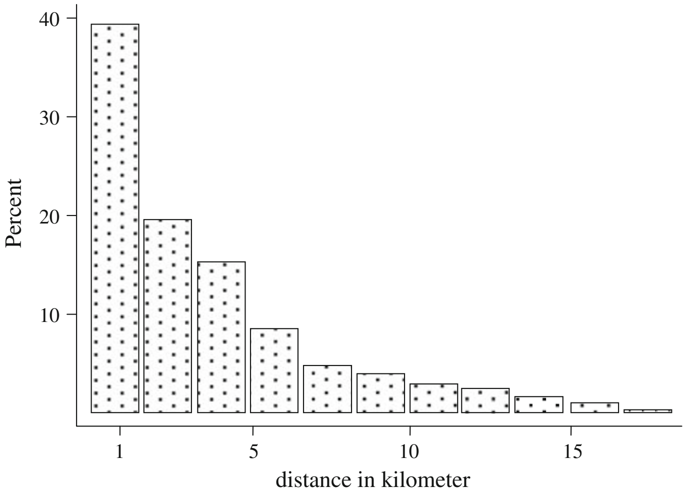

Fig. 4 Distance decay ($N = 1{,}871$ journeys to burglary)

segments in Seattle (Weisburd et al. 2004), does not seriously challenge the software. But estimation of the multinomial logit model requires access to the characteristics of all individual decision-makers multiplied by all potential alternatives. This can be easily seen by considering the log-likelihood of the multinomial logit model, which is

$$\ell = \sum_{n=1}^{N} \sum_{i=1, i \neq j}^{J} y_{in} \left(\beta' x_{in} - \ln \sum_{j=1}^{J} e^{\beta' x_{nj}} \right) \tag{12}$$

where y_{in} is the observed choice, such that $y_{in} = 1$ if decision-maker n chooses alternative i and $y_{in} = 0$ if she chooses another alternative. As the likelihood function shows, maximum likelihood estimation requires summation across decision-makers and across alternatives, which is prohibitively time-consuming or even impossible in situations where both N and J are large. Because the data apply to $n = 1{,}871$ burglaries each of which might have been committed in any of the $J = 23{,}984$ postal code areas with a residential function, the likelihood function must be calculated over 44,874,064 combinations. Because each combination requires at least 40 bytes to store the variables in the estimation as well as identifiers for decision-makers and alternatives, the data for estimation only require already 1.67 gigabyte of computer memory, which is exclusive of any buffer memory that the estimation algorithms need to store intermediate results. The software and hardware I used[2] did not allow estimation of the model on the full dataset.

A solution that involves taking a sample of alternatives from the choice set has been suggested by McFadden (1978), and is used here to overcome the computational burden. McFadden showed that if, for each decision-maker, a sample of elements of the choice set is taken that includes the chosen alternative and one or more randomly chosen other alternatives, the estimation of the multinomial logit model on this reduced choice set will yield consistent estimates of the model parameters. In other words: it is not necessary to use all data in the estimation, a random sample is sufficient provided that the chosen

[2] I used the 'clogit' module in Stata/SE version 10 on a 32-bits Windows XP computer with 2 Gb RAM, of which 800 Mb of contagious memory was available for the estimation problem.

alternative is included. There are various cases in the literature that have followed this approach (Chattopadhyay 2000; Duncombe et al. 2001; Frejinger et al. 2009; Li et al. 2005; Nerella and Bhat 2004). In the analyses presented here, in addition to the chosen alternative, a random sample of 1,499 alternatives was selected from the full choice set of 23,984.

Ben-Akiva and Lerman (1985) suggest that intuitively it seems logical to have the sampling procedure not generate a simple random sample of all non-chosen alternatives, but let the sampling procedure for any given decision-maker give greater weight to elements in the choice set that are a priori known to have a high probability of being chosen by the decision-maker, and give lower weight to unlikely alternatives. Sampling procedures with this property are referred to as 'importance sampling'. In order to generate sampling weights, they require at least some prior knowledge of the behavior of decision-makers. In the case of burglary location choice, where from the literature it is known that distance from home is an important factor, one could set up a sample that gives relatively high inclusion probabilities to nearby houses and low inclusion probabilities to distant targets. It was decided not to follow this approach because importance sampling has optimal properties only for the application of a model to predict choices, but for model estimation it is only an intuitively but reasonable sampling strategy (Ben-Akiva and Lerman 1985: 265).

Unfortunately, the proof that a random sample from the alternatives yields consistent parameters, depends on the IIA property (also see Nerella and Bhat 2004). Thus, when the model to be estimated lacks the IIA property, there is no proof that sampling from alternatives—either random sampling or importance sampling—yields consistent estimates. Because there simply is no other way to estimate this model with a large choice set, it will be assumed that the sampling from alternatives in model E, that lacks the IIA property, will also in this case lead to consistent estimates with the sampling from alternatives procedure.

Findings

Following a common modeling strategy, the analysis commences by specifying a simple model with only a few variables, and is gradually extended by including more groups of explanatory variables and interaction terms. Each model will involve the number of residential properties in the area as an explanatory variable. For residential burglary, this is precisely the number of potential opportunities in the area and therefore an adequate general measure of burglary opportunity.

Because in the literature on location choice the effect of distance is dominant, model A in Table 2 includes only distance as an additional explanatory variable beyond the number of properties. The shape of the distance decay curve displayed in Fig. 1 suggests that the distance decay curve more or less follows a negative exponential function of distance. To correct for the nonlinearity and to transform the original distance metric into a measure of 'proximity', the reverse transformation was made, and the negative of the logarithm of the distance was entered in the equation, hypothesizing a linear positive effect of proximity. The results, displayed in Table 2 under column label 'model A', indicate that both the number of residential properties and proximity increase the likelihood of an area to be selected from burglary. As the number of properties increases by 10, a postal code's odds of being chosen increases by a factor 1.14 (by 14%). For proximity (i.e., negative log distance) the effect is also positive, indicating offenders are more likely to choose near than distant areas. Both effects are highly significant.

Table 2 Multinomial logit model estimates

	Model A		Model B		Model C		Model D		Model E	
Variable	e^{β}	z	e^{β}	z	e^{β}	z	e^{β}	z	e^{β}	z
nr of properties (x 10)	1.14*	7.4	1.14*	7.45	1.14*	7.2	1.16*	8.3	1.54*	10.1
proximity -ln(km)	3.27*	57.4								
juvenile offenders			3.91*	49.9	3.28*	51.1	3.31*	50.6	3.31*	51.3
adult offenders			2.77*	34.1	2.53*	36.8	2.56*	37.0	2.64*	38.8
population nonnative (x 10%)										
non-native offenders					1.01	0.2	1.01	0.2	1.04	1.5
native offenders					.92*	-3.8	.93*	-3.4	.97	-1.6
population ages 15-25 (x 10%)							1.01*	3.2	1.01*	3.2
value properties (x €1000)							3.09*	6.1	2.36*	4.2
Spatial competition factor									0.99*	-6.6

$N = 1{,}871$ burglaries. Choice set is the postal code area of the burglary and a random sample of 1,499 out of 23,984 other residential postal code areas

Solid lines around a pair of coefficients indicate they are statistically different, $p < 0.01$

Dashed lines around a pair of coefficients indicate they are statistically different, $p < 0.05$

* $p < 0.01$

In model B, the same model is estimated, this time allowing the effect of proximity to vary between juvenile offenders (up to age 21) and adults (age 22 and beyond) as the mobility of juveniles is often more constrained because they are subjected to more supervision, maybe have less discretionary time available and because they have access to less modes of travel (Bernasco and Block 2009; Bernasco and Nieuwbeerta 2005). While in model A the effect was constrained to be the same for all offenders, the outcomes of model B show that proximity is important for both juvenile and adult offenders (both effects, 3.91 and 2.77, are significant), but also that it is a more salient choice criterion for juveniles than for adults (the significance of the difference between the effects of 3.91 and 2.77 is indicated in the table by drawing a border around them).

Model C introduces another attribute of potential target areas that may vary across offender types: the percentage of the resident population of nonnative (that is, non-Dutch) origin. As has been argued elsewhere (Bernasco and Block 2009; Bernasco and Nieuwbeerta 2005; Reynald et al. 2008), the ethnic background of offenders themselves may influence their preference for areas with mixed ethnic or racial composition, mainly because by skin color they are easily identified as non-natives by local residents. Interestingly, the results differ from those that Bernasco and Nieuwbeerta report on the basis of aggregated neighborhood level data from the city of The Hague (which are, however, a subset of the data analyzed here). While they found that all burglars were attracted to ethnically mixed neighborhoods, but nonnative burglars more so than native burglars, the present findings suggest that native Dutch burglars actually avoid ethnically mixed postal code areas (as for them the likelihood of choosing an area decreases by a factor 0.92 as the percentage non-natives in the population increases by 10%), while burglars with a non-native background appear indifferent to the ethnic composition of the target area. Thus, ethnic composition has a different effect for native and non-native burglars. The difference between native and non-native burglars is significant ($p < 0.05$). The observed difference with prior findings may be due to the smaller units of analysis used in the present analysis, or to the fact that the present analysis uses burglary data from a larger area (Greater The Hague area) than the prior work (city of The Hague).

In model D two more postal code area characteristics are added to the equation: the percentage 15–25 years old in the population and the average value of the properties in the postal code area. Both attributes, but the real estate values in particular, positively affect the attractiveness of the area to burglars.

In the final model, model E, the spatial competition factor is included. The estimated value of 0.99 is not readily interpretable in substantive terms as it represents the spatially weighted sum of nearby residential properties, but the finding that it is negative (odds ratio below 1) and significant means that spatial competition rather than spatial agglomeration is present. Thus, a postal code area that is located in the proximity of many other residential units is less likely to be targeted than an area that is relatively isolated. The significance of this term also indicates that the IIA assumption may not be justified.

Discussion

Observing that prior studies of offender mobility and crime location choice have used relatively large spatial units of analysis, and recognizing that the choice processes involved in crime target selection also operate at considerably smaller geographic scales, this paper demonstrates how the conceptual and analytical tools of the discrete choice framework (notably the multinomial logit model, the spatial competition model and the sampling from alternatives procedure) can be used in situations where the spatial units of analysis are very small (roughly the size of a football field). The analysis of burglary location choices in the Greater The Hague area, in which small spatial entities with an average of 18 properties and 40 residents are distinguished, illustrates these points.

The remainder of this discussion will first highlight the advantage of the discrete choice framework compared to other methods of studying the geography of burglary, and elaborate its theoretical basis. Next, the issue of the proper choice set in burglary location choices is addressed. It is argued that burglars are likely to follow a spatially structured hierarchical target selection process. Finally, opportunities for future inquiry are discussed.

Why Discrete Spatial Choice Modeling in the Geography of Burglary?

There exist hundreds of studies that explore spatial aspects of burglary. Some deal with the identification of hot spots (e.g., Bennett 1995), some test the distances that burglars travel from home to the burgled residence (e.g., Snook 2004), some investigate space-time clustering in burglary (e.g., Bowers and Johnson 2005), and some model the distribution of burglaries using (spatial) regression models (e.g., Ceccato et al. 2002) or hierarchical linear (multilevel) models (e.g., Wilcox et al. 2007). What is the advantage of the discrete choice framework for modeling burglary location choice?

A major innovation of the discrete choice model for research on burglary is that it integrates offender variability into an analytical strategy that is exclusively target-based. Target-based analytical strategies use individual homes, or aggregates like street blocks or neighborhoods, as the unit of analysis, and formulate and test propositions on the relation between the characteristics of the homes—or the areas—and the burglary risk or the burglary rate. Although this analytical framework may in principle be regarded as a decision-model of the burglar (the characteristics significantly related to burglary risk are the choice criteria) it assumes that all burglars use the same criteria when choosing a burglary target, and because travel distance is necessarily excluded it also assumes that burglars' mobility is unlimited. While some have attempted to resolve the latter issue by including in the burglary rate regression equation a measure of the concentration of known burglars living nearby the target area (Bernasco and Luykx 2003), this does not solve the former issue (the assumption that all burglars use the same decision criteria). The discrete choice framework does. It makes explicit that characteristics of potential targets do not attract burglaries in some mysterious way, but through the decision-making of burglars. It is burglars who search and choose burglary targets, but not every burglar is equal. For example, the presented findings in this article confirm that where a burglar lives, is an important choice criterion, and Bernasco (forthcoming) recently demonstrated that offenders, including burglars, often even offend near areas of past residence.

According to the discrete choice perspective, burglars searching and targeting properties behave like *optimal foragers* (Bernasco 2009b; Felson 2006; Johnson et al. 2009). Optimal foraging theory asserts that when predatory animals select hunting areas and prey, they optimize rewards by weighting the nutrition value of a potential prey with the efforts and risks involved in finding, attacking and eating it. In the same vein, burglars may be assumed to maximize their revenues by selecting neighborhoods and dwellings that require little effort to reach and enter, that appear to contain valued items, and that give the impression that the likelihood of being disturbed or apprehended there is low.

This perspective from behavioral ecology is particularly attractive because it combines elements of rational choice theory—the assertion that burglars maximize rewards by purposefully selecting targets from a set of alternatives—with the notion that the actors may sometimes act impulsively and need themselves not be aware of the laws that drive their behavior.

The Spatial Choice Set in Burglary

A recurrent issue in the analysis of discrete choice problems is the question of the appropriate choice set (Thill 1992). Simply stated, that question is "If this situation is perceived by the subject as a choice, what precisely are the alternatives that s/he is aware of? Do choice sets differ between subjects?" Although this issue has been discussed for much simpler behavioral choices than location choice, it is an even more serious problem for location choice, especially if the number of potential spatial alternatives is more than a few. In the analysis reported here, a model was formulated where for each burglary it was assumed the burglar was to choose one postal code area out of no less than 23,984 postal code areas. Needless to say that in practice it would be impossible for any burglar to be aware of this number of areas and to be familiar with more than a few percent of them, let alone with the more than 430,000 residential units in the area!

It has been argued that in selecting a target burglars follow a spatially structured hierarchical sequence (Brantingham and Brantingham 1978; Brown and Altman 1981; Cornish and Clarke 1986), in which the offender first selects a neighborhood or equivalent geographic area, and subsequently targets a specific street block in the chosen area, and a specific house in the chosen street block. Such a search strategy would be sensible and plausible given the limited information processing capabilities of humans. It justifies spatial choice analysis at various levels of spatial aggregation, including the level of individual residential units.

Two problems of this model are that the researcher generally cannot know how the individual burglar partitions space (which sizes and which boundaries) and that each burglar partitions space in a different way. Because it is very difficult or even impossible to know how people cluster alternative destinations in space; researchers have instead modeled the likelihood of a destination being considered (i.e., being part of the choice set) as a function of destination attributes thought to affect spatial information processing (Fotheringham et al. 2000). The spatial competition model (Eq. (11), and model E in Table 2) used in this article is an example of such a model.

Opportunities for Future Inquiry

An issue with the spatial competition model is that as a model of spatial interaction it is of limited value, because it does not model how the attraction of a spatial unit of analysis for an offender is affected by characteristics of its environment, other than what is reflected in its 'accessibility', the spatially weighted 'size' measure. A more realistic conceptualization of spatial influence would at least be multidimensional, and allow the focal unit of analysis to be affected in different ways by, for example, the size of its neighbors, their affluence level and their age composition. Some of these spatially weighted variables may pull offenders to the focal area, while other variables may push offenders out. Future applications of crime location choice could attempt to model multiple spatial effects by including more than a single spatially weighted attribute in the model. An issue of concern is that the potentially large amount of spatial autocorrelation between variables and their spatial lags may lead to problems of multicollinearity in the discrete choice model.

There are numerous other issues in burglary location choice that deserve attention. These include the theoretical and empirical analysis of crime location choice in co-offending (Bernasco 2006), and the integration with temporal aspects of burglary location choice, because some places are excellent day-time targets but poor night-time targets (Coupe and Blake 2006). Another open question is how burglars learn form prior burglary experiences. Are successful burglaries followed by similar burglaries in the same place, and do unsuccessful burglaries lead burglars to adapt their spatial or temporal choices?

Although these issues will not be resolved in the current article, its analytical approach and the solutions obtained for analytical problems lay the groundwork for future applications of location choices models at the level of micro-places.

Acknowledgments Crime data were kindly made available by *politie Haaglanden* (Greater The Hague Police Force). I thank Henk Elffers and Gerben Bruinsma for the fruitful discussions that helped me realize this paper, and the participants of the Crime and Place Working Group, three anonymous reviewers of this journal and the editors of this special issue, for insightful comments on previous drafts.

References

Andresen MA (2006) Crime measures and the spatial analysis of criminal activity. Br J Criminol 46: 258–285

Baller RD, Anselin L, Messner SF, Deane G, Hawkins DF (2001) Structural covariates of US county homicide rates: incorporating spatial effects. Criminology 39:561–590

Barker RG (1968) Ecological psychology: concepts and methods for studying the environment of human behavior. Stanford University Press, Stanford

Ben-Akiva M, Bierlaire M (1999) Discrete choice methods and their applications to short term travel decisions. In: Hall RW (ed) Handbook of transportation science. Kluwer, Norwell, pp 5–34

Ben-Akiva ME, Lerman SR (1985) Discrete choice analysis: theory and applications to travel demand. MIT Press, Cambridge

Bennett T (1995) Identifying, explaining, and targeting burglary 'hot spots'. Eur J Crim Policy Res 3: 113–123

Bernasco W (2006) Co-offending and the choice of target areas in burglary. J Invest Psychol Offender Profiling 3:139–155
Bernasco W (2009a) Burglary. In: Tonry M (ed) The oxford handbook of crime and public policy. Oxford University Press, Oxford, pp 165–190
Bernasco W (2009b) Foraging strategies of Homo Criminalis: lessons from behavioral ecology. Crime Patterns Anal 2:5–16
Bernasco W (forthcoming) A sentimental journey to crime; effects of residential history on crime location choice. Criminology 48(2)
Bernasco W, Block R (2009) Where offenders choose to attack: a discrete choice model of robberies in Chicago. Criminology 47:93–130
Bernasco W, Luykx F (2003) Effects of attractiveness, opportunity and accessibility to burglars on residential burglary rates of urban neighborhoods. Criminology 41:981–1001
Bernasco W, Nieuwbeerta P (2005) How do residential burglars select target areas? A new approach to the analysis of criminal location choice. Br J Criminol 45:296–315
Bhat C, Zhao H (2002) The spatial analysis of activity stop generation. Transp Res Part B Methodol 36: 557–575
Boots BN, Kanaroglou PS (1988) Incorporating the effects of spatial structure in discrete choice models of migration*. J Reg Sci 28:495–510
Bowers KJ, Johnson SD (2005) Domestic burglary repeats and space-time clusters: the dimensions of risk. Eur J Criminol 2:67–92
Brantingham PJ, Brantingham PL (1978) A theoretical model of crime site selection. In: Krohn MD, Akers RL (eds) Crime, law and sanctions. Theoretical perspectives. Sage, Beverly Hills, pp 105–118
Brown BB, Altman I (1981) Territoriality and residential crime: a conceptual framework. In: Brantingham PJ, Brantingham PL (eds) Environmental criminology. Sage, Beverly Hills, pp 55–76
Ceccato V, Haining R, Signoretta P (2002) Exploring offence statistics in Stockholm city using spatial analysis tools. Ann Assoc Am Geogr 92:29–51
Chattopadhyay S (2000) The effectiveness of McFaddens's nested logit model in valuing amenity improvement. Reg Sci Urban Econ 30:23–43
Clare J, Fernandez J, Morgan F (2009) Formal evaluation of the impact of barriers and connectors on residential burglars' macro-level offending location choices. Aust New Zealand J Criminol 42:139–158
Cornish DB, Clarke RV (1986) Introduction. In: Cornish DB, Clarke RV (eds) The reasoning criminal: rational choice perspectives on offending. Springer, New York, pp 1–16
Coulton C, Korbin J, Chan T, Su M (2001) Mapping residents' perceptions of neighborhood boundaries: a methodological note. Am J Commun Psychol 29:371–383
Coupe T, Blake L (2006) Daylight and darkness targeting strategies and the risks of being seen at residential burglaries. Criminology 44:431–464
Coupe RT, Girling AJ (2001) Modelling police success in catching burglars in the act. Omega 29:19–27
Coupe T, Griffiths M (1996) Solving residential burglary (Crime Detection and Prevention series, No. 77). Home Office, Police Research Group, London
Deane G, Messner S, Stucky T, McGeever K, Kubrin C (2008) Not 'Islands, Entire of Themselves': exploring the spatial context of city-level robbery rates. J Quant Criminol 24:337–421
Dubin RA (1998) Spatial autocorrelation: a primer. J Hous Econ 7:304–327
Duncombe W, Robbins M, Wolf DA (2001) Retire to where? A discrete choice model of residential location. Int J Popul Geogr 7:281–293
Elffers H (2003) Analysing neighbourhood influence in criminology. Stat Neerl 57:347–367
Elffers H, Reynald D, Averdijk M, Bernasco W, Block R (2008) Modelling crime flow between neighbourhoods in terms of distance and of intervening opportunities. Crime Prev Commun Saf 10:85–96
Felson M (2006) Crime and nature. Sage, Thousand Oaks
Fik TJ (1988) Hierarchical interaction: the modeling of a competing central place system. Ann Reg Sci 22:48–69
Fotheringham AS (1983a) A new set of spatial interaction models: the theory of competing destinations. Environ Plann A 15:15–36
Fotheringham AS (1983b) Some theoretical aspects of destination choice and their relevance to production-constrained gravity models. Environ Plann 15:1121–1132
Fotheringham AS (1985) Spatial competition and agglomeration in urban modelling. Environ Plann A 17:213–230
Fotheringham AS, Brunsdon C, Charlton M (2000) Quantitative geography, perspectives on spatial data analysis. Sage, London

Fotheringham AS, Nakaya T, Yano K, Openshaw S, Ishikawa Y (2001) Hierarchical destination choice and spatial interaction modelling: a simulation experiment. Environ Plann A 33:901–920
Frejinger E, Bierlaire M, Ben-Akiva M (2009) Sampling of alternatives for route choice modeling. Transp Res Part B Methodol 43:984–994
Getis A (2007) Reflections on spatial autocorrelation. Reg Sci Urban Econ 37:491–496
Gitlesen JP, Thorsen I (2000) A competing destinations approach to modeling commuting flows: a theoretical interpretation and an empirical application of the model. Environ Plann A 32:2057–2074
Groff ER, La Vigne NG (2001) Mapping an opportunity surface of residential burglary. J Res Crime Delinq 38:257–278
Groff E, Weisburd D, Morris NA (2009) Where the action is at places: examining spatio-temporal patterns of juvenile crime at places using trajectory analysis and GIS. In: Weisburd D, Bernasco W, Bruinsma GJN (eds) Putting crime in its place: units of analysis in geographic criminology. Springer, New York, pp 61–86
Hensher DA, Bradley M (1993) Using stated response choice data to enrich revealed preference discrete choice models. Mark Lett 4:139–151
Hu P, Pooler J (2002) An empirical test of the competing destinations model. J Geogr Syst 4:301–323
Hunt LM, Boots B, Kanaroglou PS (2004) Spatial choice modelling: new opportunities to incorporate space into substitution patterns. Prog Hum Geogr 28:746–766
Johnson S, Summers L, Pease K (2009) Offender as forager? A direct test of the boost account of victimization. J Quant Criminol 25:181–200
Kinney JB, Brantingham PL, Wuschke K, Kirk MG, Brantingham PJ (2008) Crime attractors, generators and detractors: land use and urban crime opportunities. Built Environ 34:62–74
Kleemans ER (1996) Strategische misdaadanalyse en stedelijke criminaliteit. Een toepassing van de rationele keuzebenadering op stedelijke criminaliteitspatronen en het gedrag van daders, toegespitst op het delict woninginbraak. Universiteit Twente, Enschede, The Netherlands
Kubrin CE (2003) Structural covariates of homicide rates: does type of homicide matter? J Res Crime Delinq 40:139–170
Kurtz EM, Koons BA, Taylor RB (1998) Land use, physical deterioration, resident-based control, and calls for service on urban streetblocks. Justice Q 15:121–149
Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? Ecology 74:1659–1673
Li M-T, Chow L-F, Zhao F, Li S-C (2005) Geographically stratified importance sampling for the calibration of aggregated destination choice models for trip distribution. Transp Res Rec J Transp Res Board 1935:85–92
Logie R, Wright RT, Decker SH (1992) Recognition memory performance and residential burglary. Appl Cogn Psychol 6:109–123
Markowitz FE, Bellair PE, Liska AE, Liu JH (2001) Extending social disorganization theory: modeling the relationships between cohesion, disorder, and fear. Criminology 39:293–320
McCord ES, Ratcliffe JH (2007) A micro-spatial analysis of the demographic and criminogenic environment of drug markets in philadelphia. Aust New Zealand J Criminol 40:43–63
McFadden D (1973) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed) Frontiers in econometrics. Academic Press, New York, pp 105–142
McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed) Frontiers in econometrics. Academic Press, New York, pp 105–142
McFadden D (1978) Modeling the choice of residential location. In: Karlkvist A, Lundkvist L, Snikars F, Weibull J (eds) Spatial interaction theory and planning models. North-Holland Publ. Corp., Amsterdam, pp 75–96
McFadden D (2001) Disaggregate behavioral travel demand's RUM side: a 30-YEAR retrospective. In: Henscher DA (ed) Travel behavior research; the leading edge. Pergamon, Oxford, pp 17–63
Mears DP, Bhati AS (2006) No community is an Island: the effects of resource deprivation on urban violence in spatially and socially proximate communities. Criminology 44:509–548
Morenoff JD, Sampson RJ, Raudenbush SW (2001) Neighbourhood inequality, collective efficacy, and the spatial dynamics of urban violence. Criminology 29:517–559
Nerella S, Bhat C (2004) Numerical analysis of effect of sampling of alternatives in discrete choice models. Transp Res Rec J Transp Res Board 1894:11–19
Oberwittler D, Wikström P-OH (2009) Why small is better: advancing the study of the role of behavioral contexts in crime causation. In: Weisburd D, Bernasco W, Bruinsma GJN (eds) Putting crime in its place: units of analysis in geographic criminology. Springer, New York, pp 35–59
Openshaw S (1984) The modifiable areal unit problem. Geo Books, Norwich
Peeters MP, Elffers H (forthcoming) Do physical barriers affect urban crime trips? The effects of a highway, a railroad, a park or a canal on the flow of crime in The Hague. Crime Patterns Anal 3

Pellegrini PA, Fotheringham AS (2002) Modelling spatial choice: a review and synthesis in a migration context. Prog Hum Geogr 26:487–510
Pooler J (1997) Competition among destinations in spatial interaction models: a new point of view. Chin Geogr Sci 8:212–224
Rengert GF, Wasilchick J (2000) Suburban burglary: a tale of two suburbs. Charles C. Thomas, Springfield, IL
Reynald D, Averdijk M, Elffers H, Bernasco W (2008) Do social barriers affect urban crime trips? The effects of ethnic and economic neighbourhood compositions on the flow of crime in The Hague, The Netherlands. Built Environ 34:21–31
Robinson WS (1950) Ecological correlations and the behavior of individuals. Am Soc Rev 15:351–357
Shaw WD, Ozog MT (1999) Modeling overnight recreation trip choice: application of a repeated nested multinomial logit model. Environ Res Econ 13:397–414
Sherman L, Gartin PR, Buerger ME (1989) Hot spots of predatory crime: routine activities and the criminology of place. Criminology 27:27–55
Smith TS (1976) Inverse distance variations for the flow of crime in urban areas. Social Forces 54:802–815
Smith WR, Frazee SG, Davison EL (2000) Furthering the integration of routine activity and social disorganization theories: small units of analysis and the study of street robbery as a diffusion process. Criminology 38:489–523
Snook B (2004) Individual differences in distance traveled by serial burglars. J Invest Psychol Offender Profiling 1:53–66
Snook B, Cullen RM, Mokros A, Harbort S (2005) Serial murderers' spatial decisions: factors that influence crime location choice. J Invest Psychol Offender Profiling 2:147–164
St. Jean PKB (2007) Pockets of crime. Broken windows, collective efficacy, and the criminal point of view. University of Chicago Press, Chicago
Taylor RB (1997) Social order and disorder of street blocks and neighborhoods: ecology, microecology, and the systemic model of social disorganization. J Res Crime Delinq 34:113–155
Taylor M, Nee C (1988) The role of cues in simulated residential burglary—a preliminary investigation. Br J Criminol 28:396–401
Thill J-C (1992) Choice set formation for destination choice modelling. Prog Hum Geogr 16:361–382
Tobler WR (1970) A computer movie simulating urban growth in the detroit region. Econ Geogr 46: 234–240
Train K (1998) Recreation demand models with taste variation. Land Econ 74:230–239
Train K (2009) Discrete choice methods with simulation, 2nd edn. Cambridge University Press, New York
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 42:283–322
Weisburd D, Bernasco W, Bruinsma GJN (eds) (2009a) Putting crime in its place: units of analysis in geographic criminology. Springer, New York
Weisburd D, Bruinsma GJN, Bernasco W (2009b) Units of analysis in geographic criminology: historical development, critical issues, and open questions. In: Weisburd D, Bernasco W, Bruinsma GJN (eds) Putting crime in its place: units of analysis in geographic criminology. Springer, New York, pp 3–31
Wicker AW (1987) Behavior settings reconsidered: temporal stages, resources, internal dynamics, context. In: Stokols D, Altman I (eds) Handbook of environmental psychology. Wiley, New York, pp 613–653
Wikström P-OH (2006) Individuals, settings, and acts of crime: situational mechanisms and the explanation of crime. In: Wikström P-OH, Sampson RJ (eds) The explanation of crime: context, mechanisms, and development. Cambridge University Press, Cambridge, pp 61–107
Wilcox P, Madensen TD, Tillyer MS (2007) Guardianship in context: implications for burglary victimization risk and prevention. Criminology 45:771–803
Wiles P, Costello A (2000) The 'road to nowhere': the evidence for traveling criminals (Home Office Research Study No. 207). Home Office, Research, Development and Statistics Directorate, London
Wright R, Logie RH, Decker SH (1995) Criminal expertise and offender decision making: an experimental study of the target selection process in residential burglary. J Res Crime Delinq 32:39–53
Zhang L, Messner SF, Liu J (2007) A multilevel analysis of the risk of household burglary in the city of Tianjin, China. Br J Criminol 47:918–937