

## **Title**

Modelling taste heterogeneity regarding offence location choices

## **Author**

**Name:** Michael J. Frith<sup>1</sup>

**Affiliation:** <sup>1</sup> Department of Security and Crime Science, University College London

**Address:** 35 Tavistock Square, Department of Security and Crime Science, University College London, WC1H 9EZ

**Email:** michael.frith@ucl.ac.uk

## **Key Words**

Crime; offence locations; taste heterogeneity; latent class model, random parameter logit

## **Highlights**

- Serious acquisitive crime offenders vary in their offence location preferences
- Conditional logit estimates of these preferences can be biased due to heterogeneity
- Model fits suggest heterogeneity can be modelled using mixed logit or latent class
- Mixed logit and latent class can however give slightly different results
- Latent class though shows the heterogeneity can be linked to types of offenders

## **Abstract**

One of the central topics in crime research, and one in which discrete choice modelling has been relatively recently introduced, is the study of where offenders choose to commit crime. Since the introduction of this approach in 2003, it has however become relatively popular and used in over 25 published studies covering a range of crime types and study areas. However, in most of these analyses the conditional logit has been used which assumes offenders are homogenous in their offence location preferences. This is despite various research finding offenders vary in their decision-making criteria. As such, while three recent studies (Townsley *et al.*, 2016; Frith *et al.*, 2017;

Long *et al.*, 2018) used the mixed logit and found some evidence of preference heterogeneity between offenders, there are still open questions regarding its nature. To this end, this study uses the latent class (and mixed and conditional logit) to examine the offence location choices of serious acquisitive crime offenders in York (UK). In particular, to understand how the spatial preferences differ between offenders and if there are any observable sources. Like the previous studies, this analysis identifies the presence of preference heterogeneity. This study also finds that the latent class and mixed logit equally fit the data though there are some differences in the results. These findings and other factors therefore raise questions for future crime location choice research regarding the appropriate model for these types of analyses and the true underlying nature of offender preferences.

## 1 Introduction

One of the central topics in criminological research, and one in which discrete choice modelling [DCM] has been introduced relatively recently (Bernasco and Nieuwbeerta, 2003; Bernasco and Nieuwbeerta, 2005) but has since become increasingly popular, is the study of offence location choices. Or, more simply, the study of where criminals choose to offend and why they choose those locations.

Prior to the introduction of DCM, these types of analyses generally employed one of two approaches: the target-based approach (e.g. Sampson, 1985), which analyses the relationship between the characteristics of locations (such as the demographic makeup of the residents) and the number of offences. Alternatively, the offender-based approach (e.g. Wiles and Costello, 2000) analyses the relationship between distance from each offender's home and the number of offences that offender has committed. While both approaches are not unreasonable, various research including qualitative studies supports that offending decisions are generally influenced by *both* types of factors (e.g. Bennett *et al.*, 1984; Rengert and Wasilchick, 1985; Wright and Decker, 1996), and so separate approaches do not account for their joint influence. More specifically, target-based analyses cannot incorporate proximity (for the exceptions, see Bernasco and Luykx, 2003; Bernasco and Block, 2011) and so do not address it directly as a crucial factor. Likewise, offender-based analyses do not incorporate the characteristics of the potential offence locations.

In contrast, DCM in these types of analyses takes the form for each offence of the offender choosing from a set of locations to offend in and can simultaneously include both types of variables. The models can include proximity (and other similar variables) as an individual-alternative-specific variable as it will vary for each individual (offender) based on the locations of

where they live and each alternative. The models can include the characteristics of each location as alternative-specific variables as they will vary for each possible offence location but will remain identical for all offenders. In addition, they can also incorporate individual-specific variables, such as age or ethnicity, to examine their effects on offending decisions.

In 2003 and 2005, Bernasco and Nieuwbeerta introduced this approach to the study of offence location choices by using it to analyse residential burglaries in The Hague (Netherlands) and found that the offenders preferred shorter crime trips and offending in locations with more targets, with more accessible targets and ethnically heterogeneous residents that are expected to be less likely to act as guardians for each other with the latter being especially the case for non-native offenders. Since then the approach has been used in over 25 published studies of offence location choices (Frith, 2019; see also Ruiter, 2017 for a recent review) and has been extended, for example, to find that other sources of residential heterogeneity can impact crime trips (Johnson and Summers, 2015; Frith *et al.*, 2017) and that the residential (e.g. Bernasco, 2010; Bernasco and Kooistra, 2010) and offending (Lammers *et al.*, 2015) history of offenders also impacts their crime location choices. Although many of these studies have investigated acquisitive crimes such as residential burglary (e.g. Bernasco and Nieuwbeerta, 2005; Bernasco, 2006), robbery (e.g. Bernasco and Block 2009; Bernasco *et al.*, 2013) and theft from vehicles (Bernasco, 2010; Johnson and Summers, 2015), others have analysed offences such as violence (Summers, 2012) and terrorism (Marchment and Gill, 2019) and found the offenders generally behaviour similarly. Some studies have also analysed groups of offences altogether where the offenders are assumed and generally found to share similar crime location choice criteria (e.g. Bernasco, 2010, Lammers *et al.*, 2015). These studies have also been conducted, with relatively similar results, in a range of study areas from the Netherlands (e.g. Bernasco and Nieuwbeerta, 2005; Bernasco, 2006) and England (Summers, 2012; Baudains *et al.*, 2013) to USA (e.g. Bernasco and Block, 2009; Bernasco *et al.*, 2013) and Australia (Clare *et al.*, 2009; Townsley *et al.*, 2015).

In all but three of these analyses the conditional logit [CL] (McFadden, 1974) has been used. This model assumes that offender preferences are homogenous, or that they only systematically differ between broad sub-groups, for example between juvenile and adult offenders. Despite this, various criminological research including qualitative (e.g. Bennett *et al.*, 1984; Rengert and Wasilchick, 1985; Wright and Decker, 1996) and other non-DCM quantitative studies (e.g. Townsley and Sidebottom, 2010; Bouhana *et al.*, 2016) find offenders vary in their decision-making criteria. As such, three more recent studies (Townsley *et al.*, 2016; Frith *et al.*, 2017; Long *et al.* 2018) acknowledge that offender preferences are likely to vary including in subtle or unobservable ways. These analyses model this unobserved heterogeneity using the mixed logit [ML] model (McFadden

and Train, 2000; Hensher and Greene, 2003), which assumes the preferences belong to some continuous distribution. Townsley *et al.* (2016) and Frith *et al.* (2017) both investigated individual-level heterogeneity and found significant amounts of variation in the effects of the observed variables across their samples and that their ML models fit the data better than the equivalent CL models. Long *et al.* (2018) instead investigated alternative-level heterogeneity but reported not detecting any significant unobserved heterogeneity specific to the alternatives. Based on these first analyses of heterogeneity in offence location choices there are several open research questions. In particular regarding taste heterogeneity and its scale across (other samples of) offenders, its underlying distributions, and any potentially observable sources.

Based on this, and whilst so-far unused in this literature, there is a second model that is popular in the wider literature that also deals with unobserved heterogeneity, the latent class logit [LCL] (Lazarsfeld and Henry, 1968; McLachlan and Peel, 2000). As an alternative to the continuous distributions in ML, LCL allows for unobserved heterogeneity by estimating joint discrete distributions of preferences (classes) and assigning decision-makers (offenders) to each class on a probability basis. The probabilistic class allocations can also be related to individual-level characteristics. One key advantage of the LCL, compared to ML, is that the distributions of preferences does not need to be specified; though the number of classes do. In fact, in many of comparisons of the two models (Greene and Hensher, 2003; Hess *et al.*, 2011), neither is to be assumed superior - although the extent to which this will be true for (each sample of) offenders will depend on the actual underlying distributions of preferences.

This article furthers this growing area of research through the introduction and testing of the LCL and comparing it to the equivalent CL and ML models. This is through an analysis of the offence location decisions of serious acquisitive crime [SAC] offenders which includes residential burglars, robbers and theft of and from vehicle offenders in York (UK). This study therefore contributes to the criminological literature by:

1. By estimating the presence and scale of preference variation, including for the first time using the LCL, across a different sample of offenders.
2. By testing the extent to which the heterogeneity in this sample is better represented by discrete distributions in the LCL rather than continuous distributions in the ML.
3. Because offender characteristics can be related to class membership in LCL, by investigating observable but not currently hypothesised potential sources of heterogeneity more nuanced than they can be captured using interactions with broad groupings in CL and ML.

The remainder of this article is organised as follows. In the next section, the DCM analytical framework is described in relation to offence location choice research and the models that will be used in this analysis. The second section describes the data and the analytical strategy employed. The third and fourth sections presents and discusses the results respectively.

## 2 Analytical framework

Although more established in other fields, the DCM framework and associated models are relatively straightforward to apply to analyse offence location choices. Here, and following notation from Train (2009), the choice situation can be described as there being a sample of decision makers,  $n = 1, \dots, N$ . In these types of analyses these are offenders deciding where to offend, and thus in this analysis, they are offenders deciding where to commit a SAC which includes residential burglary, robbery, and theft from and of vehicle offences. These analyses, including the present analysis, therefore do not consider potential offenders deciding whether to offend, or not, but just where to offend and so can be considered discrete spatial choice models.

For each offence, each offender is faced with a choice from  $j = 1, \dots, J$  spatial alternatives. Here, offenders are theoretically choosing from every possible target (e.g. every dwelling for a residential burglary). However, it is necessary to set some limit, for example, for computational reasons and the availability of data<sup>1</sup>. Also, because an offender's (geographical) knowledge is bounded (Clarke and Felson, 1993), those distant or otherwise rarely selected targets can be omitted as they will have a negligible chance of being chosen. Here, because the alternatives, and therefore also the sample of offenders and offences (see later), are normally limited to a geo-political region, in this analysis the choice set is restricted to the city of York (UK). Next, because criminological research highlights that offenders appear to follow a spatially-structured decision process (Brown and Altman, 1982; Bennett *et al.*, 1984), meaningful analyses can be conducted, and the choice set appropriately defined using a spatially-aggregated grouping. In general, finer spatial granularity are to be preferred (see also Weisburd *et al.*, 2009). For example, while offenders' mental boundaries between *areas* are effectively inscrutable, and will, to some degree, vary from offender to offender, larger spatial areas tend to be more heterogeneous. As such, if larger areas are used, local variations between sub-areas would be unobserved. Based on this, the alternatives in this analysis are *output*

---

<sup>1</sup> Principally due to the limited sharing of data across geo-political boundaries (such as counties in the UK) between neighbouring police forces and between the police and academia.

*areas*, herein referred to as neighbourhoods, which are the smallest units for which the variables in this study are available.

Each offender selects the neighbourhood from which they expect to derive the most utility,  $U$ . This is such that alternative  $i$  will be chosen if  $U_{ni} > U_{nj} \forall j \neq i$ . Given the utility that would be obtained from each alternative,  $U_{nj}$ , is unknown except to the offender themselves and will probably not be fully described in any model, it is decomposed into  $V_{nj}$  which represents observed factors that influence this utility and  $\varepsilon_{nj}$  which captures unobserved factors (see also later):  $U_{nj} = V_{nj} + \varepsilon_{nj} \forall j$ .

For offence location choice research, four core criminological theories - though they themselves draw upon more general theories, such as those by Bentham (1823) for rational choice, Hawley (1950) for routine activities and Lewin (1951) for crime pattern theory - can inform the observable factors:

- **Rational choice theory** suggests offenders as being broadly rational decision-makers who consider the expected rewards, costs and risks when selecting an offence location (Cornish and Clarke, 1986; Clarke and Felson, 1993).
- **Routine activities theory** states for an offence to occur, the offender must converge in time and space with a suitable target in the absence of capable guardians (Cohen and Felson, 1979). This perspective also argues that the latter is generally provided informally by residents, rather than formally the police, who by their presence or proximity can discourage crime from occurring (Felson, 1998).
- **Social disorganisation theory** describes the ability of residents to deter crime can depend on the amount of social organisation within the community (Shaw and McKay, 1942; Sampson and Groves, 1989). Therefore, in communities which lack communal ties or the ability for them to form, the residents may not feel responsible for their neighbourhood and so will be less willing to act as guardians.
- **Crime pattern theory** explains that an offender's spatial decision-making will be shaped by their everyday lives (Brantingham and Brantingham, 1981, 1993). That is where offenders, like non-offenders, routinely frequent certain activity nodes, such as their homes and workplaces. Together with the paths between these nodes, they develop an awareness space where opportunities should be favoured, for example, because they are more familiar which reduces some of the risks (Beavon *et al.*, 1994).

Based on these, in this study the potential utility from offending in each neighbourhood, regardless of the sub-type of SAC and like how these types of groups of offence types have been analysed together in previous studies (e.g. Bernasco, 2010; Lammers *et al.*, 2015), is hypothesised to be (sufficiently) described by:

- **Distance from the offender's home:** There is a large body of research that show acquisitive offenders, and offenders in general, tend to offend near their homes (Rossmo, 2000; Wiles and Costello, 2000; Townsley and Sidebottom, 2010). Although there may be greater risks by doing this, for example of being identified, doing so should need less effort in terms of reaching or returning from the target. These locations should also be preferred as they will be more familiar and so there is less uncertainty about the risks. Offenders are therefore expected to more likely offend in neighbourhoods close to their homes.
- **Distance from the city centre:** Due to the concentration of facilities and services in or near most city centres, including in the study area, it is expected to be visited more often, including by offenders, than equivalent other areas. The city centre, and nearby areas, are therefore expected to be more likely to appear in an offender's awareness space and the opportunities to be more familiar. Offenders are therefore expected to more likely offend in neighbourhoods closer to the city centre.
- **Ethnic heterogeneity:** Various qualitative research describes acquisitive offenders as preferring to offend in socially disorganised neighbourhoods (Bennett *et al.*, 1984; Wright and Decker, 2002). Therefore, because ethnic heterogeneity will likely impede social ties from forming, for example because of cultural or linguistic differences, it is expected offenders are more likely to offend in more ethnically diverse neighbourhoods.
- **Socioeconomic heterogeneity:** Like ethnic heterogeneity, the forming of social ties is also likely to be inhibited from forming when residents are from very different socioeconomic statuses (Hirschfield and Bowers, 1997). As such, it is expected that offenders are more likely to offend in more socioeconomically diverse neighbourhoods.
- **Residential Churn:** Similar to ethnic and socioeconomic heterogeneity, because neighbourhoods with elevated levels of resident turnover would be expected to inhibit community from forming, it is expected that offenders are more likely to offend in neighbourhoods with higher residential churn.
- **Affluence:** As suggested elsewhere (Rengert and Wasilchick, 1985; Light *et al.*, 1993; Wright and Decker, 2002), acquisitive offenders are primarily driven by financial benefit. They would therefore be expected to be more likely to offend in more affluent

neighbourhoods where more valuable targets are more likely to be found. However, on the other hand, because more valuable targets may be better secured against crime, it is also possible that offenders would be more likely to offend in less affluent neighbourhoods where targets which would be expected to have less security are more likely to be found.

- **Previous offence locations (for repeat offenders):** Because offending in a location will likely either involve gathering information about the area or the targets, these areas are likely to be more familiar, particularly in an offending context, to the offender and so require less effort to reconnaissance and there will be less uncertainty or risk about offending (e.g. Johnson *et al.*, 2009; Bernasco *et al.*, 2015; Lammers *et al.*, 2015; Van Sleeuwen *et al.*, 2018). Offenders would therefore be expected to be more likely to offend in area if they have offended there previously.
- **Number of potential targets:** Lastly, and although also included as a control, even if offenders select targets randomly, then neighbourhoods with larger numbers of potential targets are more likely to be selected for an offence. As such, it is expected that offenders are more likely to offend in neighbourhoods with greater numbers of potential targets.

Assuming  $V_{nj}$  is linear in parameters, the utility function can be expressed as  $V_{nj} = \beta_n x_{nj}$  where  $x_{nj}$  is the vector of the eight observed variables (see above) relating to each alternative and  $\beta_n$  is the associated vector of coefficients to be estimated. Using this, different empirical models can be generated based on different distributional specifications for  $\beta_n$  and  $\varepsilon_{nj}$ .

### 3 Discrete choice models

When it is assumed that  $\beta_n = \beta \forall n$  and  $\varepsilon_{nj}$  are independently and identically from an extreme value distribution, the CL is specified (McFadden, 1974) and the probability of offender  $n$  choosing neighbourhood  $i$  ( $P_{ni}$ ) is given by:

$$P_{ni} = \frac{\exp(\beta x_{ni})}{\sum_j \exp(\beta x_{nj})}$$

Although popular, for example because they can be easily computed due to their closed form, the CL has several limitations.

The first limitation and currently of most interest in crime location research regards  $\beta_n = \beta \forall n$ . Specifically, that the CL assumes the preferences for each attribute are identical across offenders or that they only systematically differ across specified observable sub-groups of offenders. Various

qualitative (e.g. Bennett *et al.*, 1984; Rengert and Wasilchick, 1985; Wright and Decker, 1996) and other (non-DCM) quantitative (e.g. Townsley and Sidebottom, 2010; Bouhana *et al.*, 2016) research suggest, however, offenders substantially vary in their decision-making. Some of this variation may be systematic and relatable to observed variables, for example as assumed with age grouping (e.g. Bernasco and Nieuwbeerta, 2005) where juveniles are expected to have shorter offence trips due to likely having less access to transportation. However, even in this case, this neglects unobserved heterogeneity from factors such as that some younger offenders will have, legal or not, access to vehicles and some older offenders will or will not. Furthermore, it is possible that even two offenders of the same or very similar characteristics can have different preferences.

The CL also implies strict substitution patterns including the independence of irrelevant alternatives [IIA] and that unobserved factors are independent over any repeated choices. Regarding the former, IIA implies that the relative odds of choosing any alternative over another does not depend on any other alternatives. While this seems reasonable, it ignores that some alternatives are more or less similar to each other and so can be more or less substitutable than other alternatives. Lastly, regarding the latter, while state-dependence where an offender's earlier choices influence their future choices can be accommodated, unobserved factors are assumed to be unrelated. This can be an issue as it would be expected there are unobserved factors that influence offenders' decisions and that these factors persist across choice occasions.

To resolve these issues, researchers, including in recent offence location choice research (Townsley *et al.*, 2016; Frith *et al.*, 2017; Long *et al.*, 2018), often specify ML models (McFadden and Train, 2000). Rather than fixed coefficients in CL, ML accommodates unobserved preference heterogeneity by assuming preferences,  $\beta_n$ , follow a continuous probability density function,  $f(\beta_n|\theta)$ , where  $\theta$  describes the distribution of  $\beta_n$ . This is such that  $\beta$  can vary for each offender. ML can also be generalised to explicitly accommodate repeated choices by the same offender where  $\beta$  can still vary over offenders but is consistent across the repeated choice occasions for the same offender. This is such that conditional on knowing  $\beta_n$ , the probability of offender  $n$  choosing alternatives  $i$  in choice occasions  $i_1, \dots, i_t$  is:

$$P_{ni} = \int \prod_t \left[ \frac{\exp(\beta x_{ni_{jt}})}{\sum_j \exp(\beta x_{nj_{jt}})} \right] f(\beta_n|\theta) d\beta_n$$

Lastly, ML also does not exhibit IIA as the denominators in a ratio of two probabilities are inside the integral and therefore do not cancel. The ratio therefore depends on all data, including all other alternatives. The ML however requires the ex-ante specification of the preference distributions,

albeit that normal or log-normal (if the sign of the preference is to be restricted) distributions are typically assumed.

Although it can be considered a special case of the ML and so far unused in crime location choice research, unobserved heterogeneity can also be handled using the LCL (Lazarsfeld and Henry, 1968; McLachlan and Peel, 2000). LCL assumes that  $\beta_n$  take a discrete distribution and that there exist finite unobserved sets of preferences for each attribute and that there are  $c = 1, \dots, C$  distinct classes of preference parameters and that each offender belongs to each class with some probability. This is such that the preferences for each attribute can differ across classes but is identical within each class. Similar to the ML, the probability is given by:

$$P_{ni} = \sum_c \pi_{nc} \prod_t \left[ \frac{\exp(\beta_c x_{nit})}{\sum_j \exp(\beta_c x_{njt})} \right]$$

Where  $\beta_c$  is the parameter for class  $c$  and  $\pi_{nc}$  is the probability that offender  $n$  belongs to class  $c$  and can be given by:

$$\pi_{nc} = \frac{\exp(\delta_c z_n)}{\sum_j \exp(\delta_c z_n)}$$

Where  $z_n$  are observable characteristics (such as age or gender) that can affect the class membership of offender  $n$  and  $\delta_c$  is the associated parameter. While the distribution does not need to be specified in LCL, it does however require the specification of the number of classes.

## 4 Methodology

### 4.1 Study area

The study area in these analyses is the city of York (UK), which covers approximately 270km<sup>2</sup> and in 2011 just under 200,000 residents. The choice set is defined using the 2011 UK Census output areas [OAs]. However, because the crime data was geocoded to the 2001 OA boundaries, and 12 of the (619) 2001 areas were sub-divided for 2011 Census, it is unclear for these sub-divided 2011 OAs where any offenders lived or offended. As such, these boundary changes are addressed by re-merging those affected 2011 OAs to coincide with the 2001 equivalents<sup>2</sup>. This gives a total of

---

<sup>2</sup> The alternative involves using the boundaries and related data from the 2001 UK Census, which given the crime data is for offences committed between April 2008 and March 2012, means using potentially out-dated data.

616 output areas, herein called neighbourhoods, which range in size from 0.01 to 15.4km<sup>2</sup> (mean = 0.44km<sup>2</sup>) and cover from 123 to 866 residents (mean = 309).

## 4.2 Crime data

For these analyses, revealed preference data for 1,105 *recorded*<sup>3</sup> and *solved*<sup>4</sup> SAC in York committed by 687 offenders living in York between April 2008 and March 2012 were collected. This data included the offenders' home and offence locations (to the nearest 2001 OA; see above) and some demographic information about the offender: their date of birth and gender. Because some of these crimes involved multiple offenders, and this can create issues calculating some variables (e.g. the distance from the offender's home). As such, the general approach in other crime location studies (e.g. Bernasco *et al.*, 2013) is followed whereby one of the co-offenders in these crimes is randomly selected as the *single offender* and this offender's data is only included in the analyses<sup>5</sup>. This results in a final dataset of 1,105 offence location choices made by 498 offenders - of which, 325 offenders made one offence location choice, 111 made two or three, and 62 who made up to 92 location choices. This total is equivalent to approximately 10% of the reported SAC in York during the same time-period (ONS, 2018a) which is comparable (e.g. Frith *et al.*, 2017) or greater (e.g. Bernasco and Nieuwbeerta, 2005) than that used in similar studies. Note that the previous offence location variable only includes crimes that are detected within this timeframe and so ignores if they committed an offence prior to April 2008 or where an offence has gone undetected. In this sample of offenders, around 26% of offences were committed by offenders who had committed more than one type of SAC.

## 4.3 Other data

Various data were used to calculate the eight independent variables. For the distance measures, road network data, which included the geometry of all roads, were provided by the Ordnance Survey. For each offender, the distance needed to travel to offend in each alternative (including in the alternative chosen) is calculated by the minimum distance along the road network between the centroids of the neighbourhood where the offender lives and each other neighbourhood. For the distance needed to travel to offend within the neighbourhood where the offender lives, the mean

---

<sup>3</sup> These are incidents reported to or otherwise discovered by the police which amount to a crime by the law and there is no credible evidence to the contrary (that the incident was not a crime) (UK Home Office, 2018).

<sup>4</sup> Although the specific terminology varies, including detected (in the UK), deared (in the US) and solved (more generally), this means in relation to a recorded crime, at least one person has been charged (or otherwise dealt with) due to sufficient evidence and likelihood of conviction (ONS, 2018b).

<sup>5</sup> As with the other analyses that have followed this procedure, this approach is tested by repeating the random selection of offences and analysing the subsequent data. In each of the five instances this was done, similar results were obtained and so is not discussed further.

distance between all roads and all other roads was used. The distance to the city centre is similarly calculated as the distance along the road network from the centroid of the neighbourhood where the offender lives and the geometric centre of the York city centre (which is defined as the area within the city walls).

The social disorganisation and number of potential target variables were calculated using data from the 2011 UK Census. Using the same measures from similar analyses (e.g. Johnson and Summers, 2015; Frith *et al.*, 2017), ethnic and socioeconomic heterogeneity and residential churn are calculated using the index of qualitative variation (Agresti and Agresti, 1978) which is calculated by:

$$IQV = (1 - \sum_{k=1}^n p_{kj}^2) \times 100$$

Where  $n$  are the different groups and  $p_{kj}$  is the proportion of people belonging to group  $k$  that reside in area  $j$ . Following previous analyses and the main groupings in the UK Census, ethnicity is classified into four groups: a) white, b) black, c) Asian and d) other. Socioeconomic status is divided into six groups: a) managerial and professional occupations, b) intermediate occupations and small employers, c) lower, semi-routine or routine occupations, d) full-time students, e) long-term unemployed, and f) other. The groupings for residential churn were: a) lived in the same neighbourhood the previous and b) lived in a different neighbourhood the previous year. These measures can be interpreted as the probability that two persons randomly selected from the same neighbourhood come from different ethnic and socioeconomic groups or did not both live in that neighbourhood the previous year. As such, the values can theoretically range from 0 to 1 where larger values indicate greater heterogeneity in the neighbourhood.

Because the actual number of potential targets vary for the types of SAC, this variable was defined by the number of households (for residential burglary offenders), number of residents (for robbery offenders) and number of vehicles (for vehicle-theft related offenders). Lastly, the level of affluence in a neighbourhood was calculated as the median price paid for a house between 2006 and 2014 in each neighbourhood according to the UK Land Registry. On average, there were 46 houses sold within the time-period per neighbourhood. An analysis of correlations between these variables was also conducted which revealed no signs of multicollinearity. Summary descriptive statistics for the independent variables are presented in Table 1.

**Table 1: Descriptive statistics of the independent variables**

Type of variable	Variable	Mean	SD	Min	Max
<b>Individual alternative specific</b>	Distance (km)	5.52	3.03	0.06	23.18
	Distance to the city centre (km)	4.09	2.51	0.03	13.76
	Previous offence location	0.01	0.08	0.00	1.00
<b>Alternative specific</b>	Ethnic heterogeneity (%)	0.09	0.08	0.00	0.54
	Residential churn (%)	0.28	0.18	0.04	0.98
	Socioeconomic heterogeneity (%)	0.83	0.06	0.13	0.87
	Affluence (£100,000)	1.91	0.59	0.54	4.60
	Number of potential targets – households (100)	1.40	0.38	0.05	4.49
	Number of potential targets – residents (100)	3.21	1.37	1.23	26.69
	Number of potential targets – vehicles (100)	1.36	0.34	0.05	4.28

#### 4.4 Model estimation

The three discrete choice models of offence location choices in these analyses are all computed in Stata 14 (StataCorp, 2015). For the LCL, there are two methods for estimation: maximum likelihood estimation (MLE) and expectation-maximisation (EM). However, because MLE can take substantially longer to compute and can also fail to achieve convergence (Bhat, 1997; Train, 2008), the LCL are estimated through EM using the ‘llogit’ command (Pacífico and Yoo, 2012) and the associated standard errors using the ‘gllamm’ command (Rabe-Hesketh *et al.*, 2002). That said, the EM algorithm can converge at a local, rather than global, maximum. As such, and following Train (2008), each LCL model is estimated 10 times with 10 different random starting values and the model with the largest log-likelihood is inferred as the global maximum. Also, as the number of classes need to be specified, the standard procedure is followed whereby it is estimated with two to eight classes (each of which are repeated 10 times; see above). These models are then assessed using the consistent Akaike information criterion [CAIC] and the Bayesian information criterion [BIC]<sup>6</sup>. Lastly, based on the available data regarding the offenders and previous analyses, including of sub-groups in related CL analyses (e.g. Bernasco and Nieuwebeerta, 2005; Menting *et al.*, 2016), the variables entered into the class membership model are: gender, age

<sup>6</sup> These are used, including over the standard Akaike information criterion, because they more heavily penalise extra parameters (i.e. more classes) and so emphasize parsimony when determining the number of the classes.

group<sup>7</sup> (under 18 years of age or over), and the number of crimes they have committed (in the dataset): one crime, two to three crimes, or four or more crimes.

The CL was estimated using maximum likelihood with the built-in ‘clogit’ command. Whilst otherwise accounted for in the LCL and ML (see earlier), due to the clustering of offences within offenders (i.e. repeated choices), robust standard errors are computed for the CL (White, 1982). For the ML, although it could also be estimated using maximum simulation likelihood, following Frith *et al.* (2017) and Townsley *et al.* (2016), it was estimated using hierarchical Bayes using the ‘bayesmixedlogit’ command (Baker, 2015). More specifically, because the parameter estimates appeared relatively stable at around 10,000 draws from the posterior distribution, these draws are discarded and the following 50,000 draws, with every 10<sup>th</sup> draw retained to prevent correlation between subsequent draws, are used to estimate the final model (see also Train and Sonnier, 2005). Also, all variables are entered non-fixed and modelled with normal distributions which is the default choice in the absence of compelling reasons to indicate otherwise<sup>8</sup>.

## 5 Results

### 5.1 Overall model fits

Prior to comparing the LCL, ML and CL, it is necessary to determine the optimum number of classes for the final LCL model. This is shown in Table 2 for the models with two to eight classes. For these models, their fits are assessed using CAIC and BIC statistics where smaller values indicate better fits relative to the number of parameters and Table 2 indicates the three-class model has the smallest values for both metrics and so is to be preferred. The other models are therefore not discussed further and the results from this model are presented in Table 3 and in Figure 1.

As shown in Table 3, and for easier interpretation, the LCL, ML and CL models are compared using the root likelihood [RLH] statistic which is calculated as the geometric mean of the estimated probabilities of the chosen alternatives. In the case of the LCL, the RLH is calculated based on the class-level preferences based on assigning each offender to the class which they are

---

<sup>7</sup> Because the specified model (see earlier) restricts these variables to remain consistent within offenders, age-group reflects their average age at the time(s) of their offences.

<sup>8</sup> To explain, in DCM some variables are sometimes entered using other distributions, for example, cost is sometimes entered using a log-normal distribution which restricts its sign, here to negative. This is because it is largely unexpected that decision-makers prefer to pay more (see also Train and Sonnier, 2005). In the case of offender spatial preferences related assumptions are not definitive. Taking the obvious example of distance, even though more distant targets need greater effort and so offenders would likely prefer to offend closer to home, offending nearby also likely raises the risk of being identified. Therefore, even though it is likely that most offenders prefer shorter crime trips, it is plausible that some will prefer offending further away.

**Table 2: Model fits for LCL models with two to eight classes**

Number of classes	Log likelihood	Number of parameters	CAIC	BIC
2	-6408.12	24	12989.3	12965.3
<b>3</b>	<b>-6338.18</b>	<b>40</b>	<b>12958.4</b>	<b>12918.4</b>
4	-6282.68	56	12969.2	12934.5
5	-6246.63	72	13012.4	12940.4
6	-6197.48	88	13029.5	12941.5
7	-6184.18	104	13118.3	13014.3
8	-6138.72	120	13142.7	13022.7

NOTE: The model with the smallest CAIC and BIC values are bold and highlighted.

probabilistically estimated to most likely belong. For the ML and CL, the RLH is calculated using the estimated individual-level parameters and the sample-level parameters respectively. Although the value itself can be used, it is more interpretable by comparing it to a reference equivalent model. Here, and using the CL as that model, values below 1 reflect that the model performs worse than the CL and values above 1 indicate how much better the model performs. As shown in Table 3, the RLH values for the LCL and ML are 4.48 and 4.45 respectively. This indicates that both models fit the data better than the CL by around  $4.5 \times$  better. The RLH statistic for the LCL and ML is relatively similar: 4.48 and 4.45 implying that there is no overall significant difference between them in terms of fitting the data.

## 5.2 Preference point estimates

In terms of the preference point estimates and starting with the LCL, Table 3 shows that the first latent class has the largest class share and offenders have a 42% probability of having the preferences indicated by the parameters for this class. Specifically, that their offence location choices are significantly influenced by five variables. These are distance from the offenders' homes, distance to the city centre, socioeconomic heterogeneity and, for offenders with multiple offences and so where its applicable, whether they had previously offended in a location which were all estimated to have a significant negative effect. Ethnic heterogeneity was estimated to have a significant positive impact. Offenders are more likely to have these preferences, compared to the other class-level preferences, if they are less prolific offenders as the coefficients for moderately and highly prolific offending are significant and positive for both other classes. Offenders are also less likely to be male compared to those in class 2 as the coefficient for male is significant for that class.

**Table 3: Estimated results from the LCL, ML and CL analyses of SAC offender location preferences**

	LCL			ML	CL
	Class 1	Class 2	Class 3		
<b>Estimated offence location choice coefficients</b>					
Distance (km)	-1.30** (0.12)	-0.45** (0.06)	0.02 (0.09)	-0.64** (0.04)	-0.41** (0.04)
Distance to the city centre (km)	-0.40* (0.16)	0.33** (0.07)	-0.28** (0.08)	-0.18** (0.05)	0.10** (0.04)
Ethnic heterogeneity (10%)	0.34** (0.12)	-0.08 (0.16)	0.19* (0.09)	-0.05 (0.07)	0.14* (0.07)
Residential churn (10%)	0.05 (0.07)	0.22** (0.07)	0.05 (0.05)	0.02 (0.04)	0.14** (0.03)
Socioeconomic heterogeneity (10%)	-0.27* (0.11)	0.72* (0.36)	0.49 (0.34)	0.64** (0.12)	0.27** (0.10)
Affluence (£100,000)	0.03 (0.15)	0.24 (0.14)	0.09 (0.12)	-0.25* (0.09)	0.14 (0.08)
Number of potential targets (100)	0.09 (0.06)	0.21 (0.17)	0.22 (0.13)	-0.01 (0.07)	0.14** (0.03)
Previous offence location (where applicable)	-1.04* (0.45)	3.90** (0.31)	1.27** (0.26)	1.82** (0.33)	2.08** (0.52)
<b>Estimated standard deviations (of the offence location coefficients)</b>					
Distance (km)	-	-	-	0.55** (0.11)	-
Distance to the city centre (km)	-	-	-	0.65** (0.14)	-
Ethnic heterogeneity (10%)	-	-	-	0.52** (0.13)	-
Residential churn (10%)	-	-	-	0.24** (0.04)	-
Socioeconomic heterogeneity (10%)	-	-	-	0.46** (0.13)	-
Affluence (£100,000)	-	-	-	0.81** (0.23)	-
Number of potential targets (100)	-	-	-	0.31** (0.07)	-
Previous offence location	-	-	-	2.62* (1.20)	-

<b>Estimated class membership coefficients</b>					
Male	0.00	1.43**	-0.84	-	-
		(0.37)	(0.48)		
Juvenile	0.00	0.13	-0.83	-	-
		(0.51)	(0.53)		
Moderately prolific offender (2-3 offences)	0.00	1.08**	1.21*	-	-
		(0.40)	(0.47)		
Highly prolific offender (4+ offences)	0.00	1.16*	2.14**	-	-
		(0.45)	(0.43)		
<b>Class share (%)</b>	42	35	23	-	-
<b>Root likelihood statistic</b> (compared to the equivalent CL)		4.48		4.45	1.00

NOTE: The results presented in the top panel are the estimated coefficients (LCL and CL) or the estimated average coefficients (ML) and the figures in parentheses are their respective standard errors. The results in the second panel are the estimated standard deviations of the average coefficients and their standard errors (for the ML). In the third and fourth panels are the estimated coefficients for class membership (where class 1 is the reference class and so the associated coefficients are all 0) and their standard errors and the class share (for the LCL).

\* indicates the parameter is significantly different to 0 at the 5% level; \*\* indicates the parameter is significantly different to 0 at the 1% level.

For the second class, offenders had a 35% probability of having these class-level preferences and being significantly influenced by five variables. These are distance which was found to have a negative impact while distance to the city centre, residential churn, socioeconomic heterogeneity and, where applicable, whether they had previously offended there were found to increase the likelihood of a location being selected. Offenders were more likely to have these preferences if they are male. They were also more likely to have these preferences than those in the third class if they had committed 2-3 or 4+ offences.

Offenders were estimated to have a 24% probability of having the class-level preferences indicated by the parameters for the third class. These preferences included being significantly influenced by three variables. These include being negatively influenced by the distance to the city centre, and positively influenced by the amount of ethnic heterogeneity as well as, and where applicable, whether they had previously offended there. Offenders were more likely to have these preferences, than the preferences for the other two classes, if they were prolific offenders.

For comparison, the results from the ML showed that offence location choices were, on average, significantly influenced by five of the eight variables. These include being, on average, negatively affected by the distance, distance to the city centre and affluence. Their offence location choices were also, on average, significantly positively influenced by socioeconomic heterogeneity and,

where applicable, whether the offender had previously offended there. For the equivalent CL, the results show that seven of the eight variables were estimated to have a significant impact on the offence location choices. These include where distance has a negative impact while all the other variables - except affluence which is estimated to have no impact – were estimated to positively influence their offence location choices.

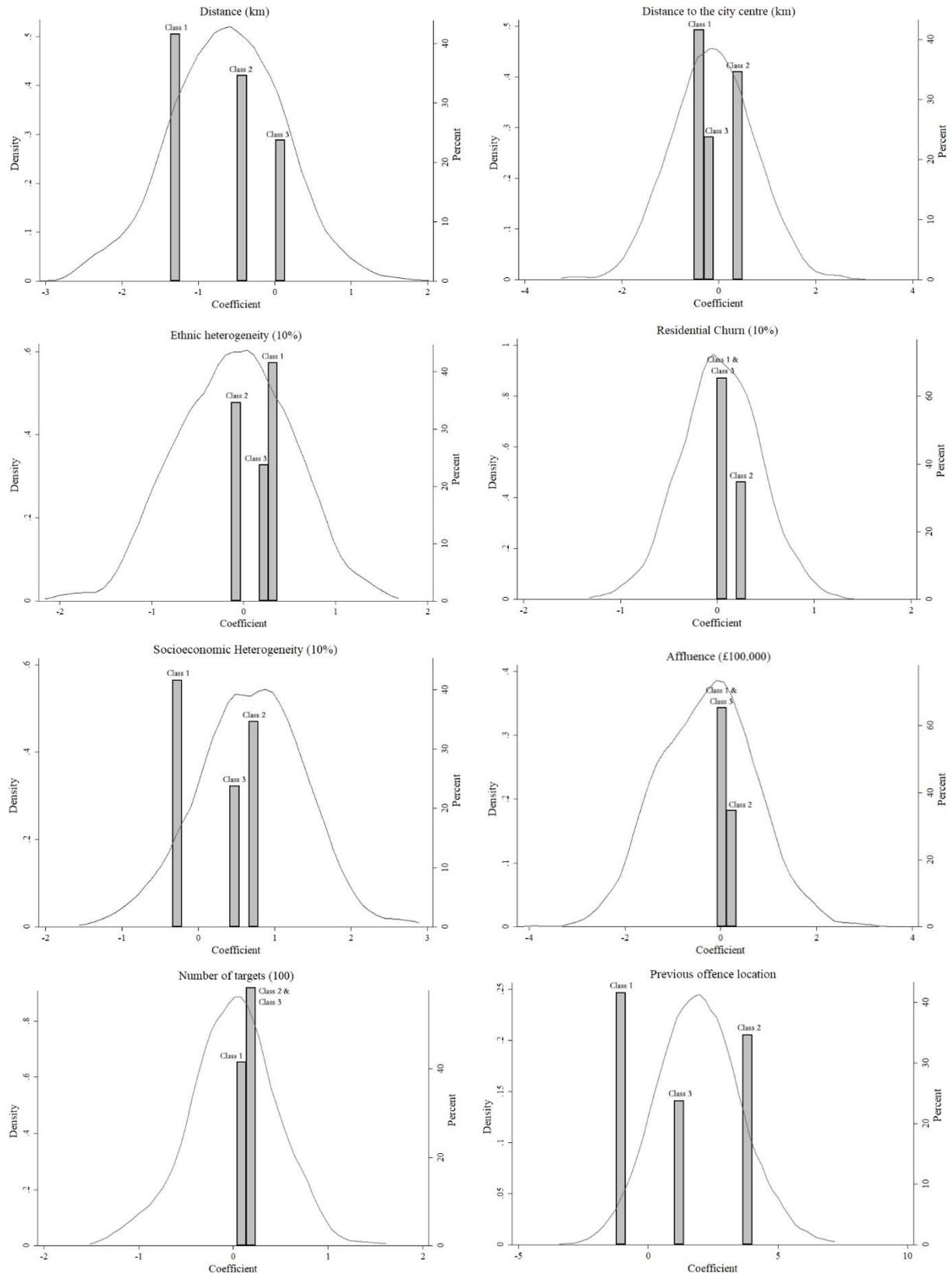
While most of the estimated effects from the LCL, ML and CL are relatively consistent, there are some discrepancies. One key example is that distance to the city centre was estimated to have a statistically significant positive effect in the CL while it has an, on average, significant negative effect in the ML and a significant negative effect for the first and third classes in the LCL. Another example is where in the ML, affluence was estimated to have an, on average, significant negative effect on offence location choices while it was found to have a non-significant positive effect for all classes of offenders in the LCL and all offenders according to the CL. Lastly and while the average effect of ethnic heterogeneity and the number of potential targets was negative, but non-significant, in the ML, their estimated effects are positive in the CL and generally or tentatively suggested to be positive according to the LCL.

### **5.3 Preference variation**

Although arguably clearer in the ML as it estimates a separate parameter for each variable to describe its scale, the LCL and ML both identify preference heterogeneity amongst the sample of offenders. In the ML, the standard deviations associated with the coefficient estimates for each variable were all statistically significant. As such, the effect of each variable was estimated to vary over the sample. The ML estimated distributions for each variable across the sample are shown in Figure 1 which highlights that although the average effect of each variable across offenders may be in one direction or of one magnitude, for some offenders it is estimated to have a much smaller or larger effect and that the effect will also be opposite to the average effect for other offenders.

Figure 1 also shows the distribution of the class-level preferences as estimated by the LCL and similarly shows the variability of preferences. For example, where it is estimated for the first class that offenders are severely deterred by having to travel larger distances while for the second class they are more modestly deterred, and it is estimated to have no effect for the third class. The effect of having previously offended in a location also similarly highly varies across classes where the second class of offenders highly prefer re-offending in the same areas, in class 3 they somewhat prefer it, whilst for the first class they are deterred from returning to the same area to re-offend.

**Figure 1: Kernel densities of the ML-estimated distributions of preferences (line; primary y-axis) and histograms of the equivalent LCL-estimated distributions (bar; secondary y-axis)**



As can also be seen in Figure 1, although generally the estimated distributions for most of the variables from both models are somewhat similar, for other variables there appears to be some disagreement. The three variables where this is arguably most pronounced are ethnic heterogeneity, socioeconomic heterogeneity and having previously offended in a location. Here, for ethnic heterogeneity the class-level preferences appear more in the upper tail of the ML distribution, for socioeconomic heterogeneity they are generally in the lower tail while the class-level preferences for previous offence location are more dispersed than in the estimated ML distribution.

## 6 Discussion

In this study, along with the CL and ML, the LCL was used for the first time to study the offence location choices and taste variation of SAC offenders in York (UK). Although previous studies (Townshley *et al.*, 2016; Frith *et al.*, 2017; see also Long *et al.*, 2018) only used the ML (and CL), like those studies this analysis highlights the presence of heterogeneity between offenders and therefore the benefits of using choice models that account for this such as the ML and LCL. In terms of the model fits, the RLH statistic indicates that allowing for heterogeneity and using the LCL or ML results in the models fitting the data substantially better, by a factor of around 4.5, than the equivalent CL. This improvement in fit is also relatively consistent with that reported in Townshley *et al.* (2016) of 5.4 and in Frith *et al.* (2017) of 6.8 which supports the argument that offenders do tend to vary in their tastes. This variability in tastes is further supported as there are discrepancies in some of the preference estimates between the CL and the LCL and ML models which implies the CL estimates can be biased due to this heterogeneity. Together these findings highlight the benefits of these models and incorporating preference heterogeneity in future analyses of offenders, including of their offence location choices.

By introducing and using the LCL in these analyses, this paper allows for the first time a comparison of the LCL and ML assumed distributions of offender preferences. The relatively similar RLH fits for both models of 4.45 for the ML and 4.48 for the LCL however indicates that neither model can be unambiguously recommended for future crime location choice research. This result though is not entirely unexpected given similar findings in other comparisons from the wider literature (e.g. Greene and Hensher, 2003; Hess *et al.*, 2011). In terms of the model results themselves, the LCL and ML were not in complete agreement for all variables. For example, affluence was found here to have, on average, a significant negative effect on offence location choices in the ML while it had a non-significant but positive effect for all classes in the LCL. Similarly, the LCL suggested relatively different distributions of preferences for the effects of

ethnic and socioeconomic heterogeneity and previous offence locations compared to the ML. This therefore suggests for offence location choice research that the results can depend on the modelling approach and emphasizes the need for further research. That is, not only with the LCL and the current specification of the ML, but also with different distributions such as lognormal in the ML which have yet to be investigated or tested within this field.

For the variables which have been previously analysed using the random-parameter ML (and CL) models, the estimated parameters and shapes of the distributions of effects in these analyses are generally in line with that found in those studies and expected based on theory. For example, and just specifically regarding similar previous ML analyses (Townnsley *et al.*, 2016; Frith *et al.*, 2017), this includes where offenders generally attempt to minimise effort in that distance in has an on average negative effect on offence location choices in both studies and this analysis. Although only included in one of the previous studies, the, on average, significant negative effect of distance to the city centre is like that found in Townnsley *et al.* (2016) and the, on average, significant positive effect of socioeconomic heterogeneity is like that reported in Frith *et al.* (2017).

However, there are disagreements in terms of the average effect, in particular regarding three variables: ethnic heterogeneity, residential churn and affluence. For the former two variables, they are both significant positive predictors in Frith *et al.* (2017) which is as expected based on social disorganisation theory as more heterogeneous neighbourhoods should be less cohesive and therefore the residents are less willing to exert informal social control. They should therefore be more attractive for offenders. Here, however, and also in Townnsley *et al.* (2016) for residential churn, they were both non-significant. Although further replication is needed to examine this; particularly as the effects of each variable was estimated to significantly differ across the samples, one potential explanation regards the size of the spatial units used as alternatives as much smaller units, average size is  $\sim 0.03\text{km}^2$ , are used in Frith *et al.* (2017) compared to here and in Townnsley *et al.* (2016) where the average sizes are  $0.44\text{km}^2$  and  $8.48\text{km}^2$  respectively (Frith, 2019). As such, if the social organisation process generally occurs at a smaller scale (Hipp, 2007), for example between neighbours on the same street, then the average levels of those related variables at a larger scale can give misleading results. In terms of affluence which had a, on average, significant positive effect in Frith *et al.* (2017) but was non-significant here and in Townnsley *et al.* (2016), the expected effect is less clear based on theory (and past research). That is, because while acquisitive offenders are assumed to be rational and often financial motivated and so they should prefer more affluent targets where greater proceeds are expected, those targets may also be better protected, and so offenders could also prefer less affluent areas where the targets are expected to be less well protected. As such, and again considering the estimated significant variation over the sample, the

disagreement in results may result from slight differences in the samples or study areas and this should be explored.

One interesting and novel finding from this analysis though regards the effects of offenders' previous offence location choices. Here, and although the identified effects on relate to offenders with previous offences as it cannot be estimated for offenders with no known offending history, the CL result generally matches that in previous CL analyses (Bernasco *et al.*, 2015; Lammers *et al.*, 2015; van Sleeuwen *et al.*, 2018). Using the LCL (and ML) however allowed the detection that some (of the repeat) offenders prefer to avoid offending in the same neighbourhoods where they have previously done so. On the face of it this is seemingly counterintuitive. For example, based on rational choice theory it is expected that offenders would prefer offending in or near previous locations as they will have already gained knowledge about the area and the targets and so there will be less risk about offending (e.g. Johnson *et al.*, 2009). This result however could follow the same and similar theories as repeatedly returning to an earlier offence location might be increasingly risky due to residents becoming more vigilant or the area receiving greater police attention following their previous offence and so it might be perceived as riskier. Again, this warrants further investigation as, even if only the case for some offenders, which in these analyses were those more likely to be less prolific, this has obvious policy implications in terms of policing. In particular, regarding crime reduction strategies such as 'cocoon watch' where neighbours are alerted to nearby crimes such as burglaries and offered advice related to their increased risk of the offenders re-offending to that area. As such, even if some offenders, which may be predictable based on the type of location where the offence took place, do not return to the same place to re-offend, then these strategies would only be effective on some occasions and the resources could be better utilised in other ways.

From these analyses, the suitability of the LCL as an alternative to the ML in offence location research is promising for several reasons. First, and especially in the context of this type of research, the LCL is arguably simpler to interpret. This is in the sense that any variation in preferences, for which there is now growing evidence, is represented by the fixed class-level parameters in the LCL rather than the average parameter in combination with its standard deviation in the ML. This is particularly relevant in this field as practitioners, such as the police, are not necessarily experts in statistics. Furthermore, the potential end-users of this research are often accustomed to di-chotomous or multi-chotomous classifications, like that provided by the LCL, such as prolific and non-prolific offenders or juvenile and adult offenders (for example, see Ministry of Justice, 2019). As such, any applicable findings are likely to be more easily assimilated into practice.

Secondly, because observable sources of heterogeneity can be exploratorily investigated with arguably greater nuance than with the CL and ML. That is, in terms of not needing to pre-specify and hypothesise differences between offenders and allow the classes of offenders to emerge from the data itself. The class-level variables then allow the investigation of the different types of offenders and their spatial preferences. In the case of these analyses, new information was garnered, especially regarding the role of offending intensity in offence location preferences. For example, where more prolific types of offenders were less concerned with distance which is counter to what is often found in, albeit, non-DCM analyses (e.g. Townsley *et al.*, 2015). In contrast to the LCL, in the ML, potential sources of heterogeneity can only be explored through post-hoc analyses of the offenders' characteristics and their individual-level parameters (Townsley *et al.*, 2016) or through interactions and estimating separate parameters based on the offenders' characteristics. That being said, the LCL (and the ML in the above case) is limited because those characteristics cannot vary across repeated choices for the same decision-maker. As such, offenders that cross some distinction, such as age group cut-offs if their offences cover a sufficient time-period, must be classified into one of the groupings and so their impact on class allocations to test specific hypotheses is not without limitations. Hypothesis testing of expected differences between offenders is also more complicated with the LCL than the ML where, for example, interactions using offender characteristics can be simply used.

Lastly, and although not discussed so-far, is the computational resources needed to estimate ML models, and particularly spatial ML models where the number of alternatives can be as large as 619 (in this analysis) or 5,286 in Frith *et al.* (2017). For comparison, the final LCL model in this analysis took a couple of hours – though the LCL should be repeatedly estimated with different starting values and numbers of classes - while the ML took several days<sup>9</sup>. The computational issue with the ML though could be overcome through sampling from alternatives and taking a random sample of the alternatives for each choice occasion. Here, and while there is no proof that this strategy will generate consistent parameter estimates in non-CL models, it has been found in various studies to only result in small or modest efficiency losses when sampling reasonable numbers of alternatives (e.g. Daly *et al.* 2014; von Haefen and Domanski, 2018).

Analyses of offence location choices, such as in this study, are subject to limitations. The first is that the data analysed are reported, recorded, and detected crimes and this is not the case for all crimes. The sample of offenders and offences included in the analyses may therefore be biased in some way. That said, the one investigation of this using the offence location choice framework

---

<sup>9</sup> All models were computed on a desktop computer with a i7-4770 CPU running at 3.40GHz and with 16GB RAM.

and the weighted exogeneous sample maximum likelihood estimator (Manski and Lerman, 1977) to correct for possible unequal probabilities of being included in the sample found similar results to the standard maximum likelihood estimator (Bernasco *et al.*, 2013). The non-complete knowledge of offences may also be a particular issue here as the analysis assumes the offenders with one offence in the time-period are not repeat offenders. The results and especially the class membership coefficients may therefore be mis-representing the differences between prolific and non-prolific offenders.

In this analysis a range of offence types were grouped together. While the offence choice criteria are expected to similarly effect the different types of offenders and this approach has been used in other analysis (e.g. Bernasco, 2010; Lammers *et al.*, 2015), it ignores that some types of offenders may behave differently with respect to these factors. This should be considered in future analyses especially where there is enough data to separately estimate these types of models (that can incorporate and estimate heterogeneity) on the different types of offenders. Doing so would not only provide more information on how offenders of the same type differ from each other, but also how different types of offenders differ from other types of offenders.

Another key issue regards the lack of substantiation of the offending decision process as modelled in these analyses. For example, where revealed preference analyses, such as in this study, rely on accurately numerating the choices including the alternatives and their (perceived) attributes. As such, complementary analyses such as those of stated preferences, which have yet to be used in the context of offending research, should be considered to support the findings. Stated preference analyses, and future revealed preference analyses, may also consider one of the underlying assumptions of the choice framework currently employed within this research: that the decision-makers are utility maximisers. That is, and while these and other types of offenders are generally considered to be boundedly rational utility maximisers (e.g. Cornish and Clarke, 1986; Clarke and Felson, 1993), there has been no investigation or testing of competing or complementary decision rules in the context of offence location choice decisions. This is even though other rules such as regret minimisation (e.g. Chorus, 2010) are plausible for explaining offending decisions as it will assume offenders avoid offending in locations where they will have a worse outcome (e.g. arrest) and therefore experience regret.

To conclude, in the current study, the LCL, ML and CL were used to investigate the presence of heterogeneity in the offence location choices of SAC offenders. The results from this analysis found that preference heterogeneity exists across the sample of offenders and the LCL and ML therefore fit the data better than the equivalent CL model. There was however little difference in

the overall fits of the LCL and ML models; though there were some differences in the effects that were estimated and their variability amongst the offenders. This research therefore raises questions for future research regarding the true underlying distributions of preferences amongst offenders. In particular, are those preferences best represented as being discretely distributed using the LCL, continuously normally distributed using the ML, or do they follow some other distributions?

## References

- Agresti, A. and Agresti, B. F. (1978) 'Statistical analysis of qualitative variation', *Sociological methodology*, 9, pp. 204–237.
- Baker, M. J. (2015) 'Bayesian mixed logit model.'
- Baudains, P., Braithwaite, A. and Johnson, S.D. (2013) Target Choice During Extreme Events: A Discrete Spatial Choice Model of the 2011 London Riots, *Criminology*, 51(2): pp. 251–85.
- Beavon, D. J. K., Brantingham, P. L. and Brantingham, P. J. (1994) 'The influence of street networks on the patterning of property offenses', *Crime prevention studies*, 2, pp. 115–148.
- Bennett, T., Wright, R. and Wright, R. (1984) *Burglars on burglary: Prevention and the offender*. Gower Aldershot.
- Bentham, J. (1823) *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press.
- Bernasco, W. (2006) Co-offending and the Choice of Target Areas in Burglary, *Journal of Investigative Psychology and Offender Profiling*, 3(3): pp. 139–55.
- Bernasco, W. (2010b) A Sentimental Journey to Crime: Effects of Residential History on Crime Location Choice, *Criminology*, 48: pp. 389–416.
- Bernasco, W., and Kooistra, T. (2010) Effects of Residential History on Commercial Robbers' Crime Location Choices, *European Journal of Criminology*, 7(4): pp. 251–65.
- Bernasco, W. and Block, R. (2011) 'Robberies in Chicago: A block-level analysis of the influence of crime generators, crime attractors, and offender anchor points', *Journal of Research in Crime and Delinquency*, 48(1), pp. 33–57.
- Bernasco, W. and Luykx, F. (2003) 'Effects of attractiveness, opportunity and accessibility to burglars on residential burglary rates of urban neighborhoods', *Criminology*, 41(3), pp. 981–1002.

- Bernasco, W. and Nieuwbeerta, P. (2005) 'How do residential burglars select target areas? A new approach to the analysis of criminal location choice', *British Journal of Criminology*, 45(3), pp. 296–315.
- Bernasco, W. and Nieuwbeerta, P. 2003. 'Hoe kiezen inbrekers een pleegbuurt? Een nieuwe benadering voor de studie van criminele doelwitselectie'. *Tijdschrift voor Criminologie*, 45(3): pp.254–270.
- Bernasco, W., and Block, R. (2009) Where Offenders Choose to Attack: A Discrete Choice Model of Robberies in Chicago, *Criminology*, 47(1): pp. 93–130.
- Bernasco, W., Block, R. and Ruiter, S. (2013) 'Go where the money is: modeling street robbers' location choices', *Journal of Economic Geography*, 13, pp. 119–142.
- Bernasco, W., Johnson, S. D. and Ruiter, S. (2015) 'Learning where to offend: Effects of past on future burglary locations', *Applied Geography*, 60, pp. 120–129.
- Bhat, C. R. (1997) 'An endogenous segmentation mode choice model with an application to intercity travel', *Transportation science*, 31(1), pp. 34–48.
- Bouhana, N., Johnson, S. D. and Porter, M. (2016) 'Consistency and specificity in burglars who commit prolific residential burglary: Testing the core assumptions underpinning behavioural crime linkage', *Legal and Criminological Psychology*, 21(1), pp. 77–94.
- Brantingham, P. J. and Brantingham, P. L. (1981) *Environmental criminology*. Sage Publications Beverly Hills, CA.
- Brantingham, P. J. and Brantingham, P. L. (1993) 'Environment, routine and situation: Toward a pattern theory of crime', *Advances in criminological theory*, 5, pp. 259–294.
- Brown, B. B. and Altman, I. (1982) 'Territoriality and residential crime', in *Environmental criminology*. Sage Beverly Hills, pp. 55–76.
- Bruinsma, G. J. N. Pauwels, L. J. R. Weerman, F. M. and Bernasco, W. (2013) 'Social disorganization, social capital, collective efficacy and the spatial distribution of crime and offenders an empirical test of six neighbourhood models for a Dutch city', *British Journal of Criminology*, 53(5), pp. 942–963.
- Chorus, C.G. (2010) A new model of Random Regret Minimization, *European Journal of Transport and Infrastructure Research*, 10(2): pp. 181-196
- Clare, J., Fernandez, J. and Morgan, F. (2009) Formal Evaluation of the Impact of Barriers and

Connectors on Residential Burglars' Macro-Level Offending Location Choices, *Australian and New Zealand Journal of Criminology*, 42(2): pp. 139-158.

Clarke, R. V. G. and Felson, M. (1993) *Routine activity and rational choice*. Transaction Publishers.

Cohen, L. E. and Felson, M. (1979) 'Social change and crime rate trends: A routine activity approach', *American sociological review*, pp. 588–608.

Cornish, D. B. and Clarke, R. V. G. (1986) 'The Reasoning Criminal: Rational Choice Perspectives on Offending'. New York: Springer-Verlag.

Daly, A. Hess, S. and Dekker, T. (2014) Daly, Andrew, Stephane Hess, and Thijs Dekker. "Practical solutions for sampling alternatives in large-scale models, *Transportation Research Record*, 2429(1), pp. 148-156.

Felson, M. (1998) *Crime and Everyday Life*. 2nd Edition. Thousand Oaks: Pine Forge Press.

Frith, M. J. (2019) A discrete choice and configurational analysis of burglary offence location choices. Ph.D. Thesis. University College London.

Frith, M. J., Johnson, S. D. and Fry, H. M. (2017) 'Role of the street network in burglars' spatial decision-making', *Criminology*, 55(2), pp. 344–376.

Greene, W. H. and Hensher, D. A. (2003) 'A latent class model for discrete choice analysis: contrasts with mixed logit', *Transportation Research Part B: Methodological*, 37(8), pp. 681–698.

Hawley, A. H. (1950) *Human ecology: a theory of community structure*. Ronald Press Co.

Hensher, D. A. and Greene, W. H. (2003) 'The mixed logit model: the state of practice', *Transportation*, 30(2), pp. 133–176.

Hess, S. Ben-Akiva, M. Gopinath, D. and Walker, J. (2011) 'Advantages of latent class over continuous mixture of logit models', *Institute for Transport Studies, University of Leeds. Working paper*.

Hipp, J. R. (2007). Block, Tract, and Levels of Aggregation: Neighborhood Structure and Crime and Disorder as a Case in Point, *American Sociological Review*, 72(5): pp. 659-680.

Hirschfield, A. and Bowers, K. J. (1997) 'The Effect of Social Cohesion on Levels of Recorded Crime in Disadvantaged Areas', *Urban Studies*, 34(8), pp. 1275–1295.

Johnson, S. D. and Bowers, K. J. (2010) 'Permeability and burglary risk: are cul-de-sacs safer?', *Journal of Quantitative Criminology*, 26(1), pp. 89–111.

- Johnson, S. D. and Summers, L. (2015) 'Testing ecological theories of offender spatial decision making using a discrete choice model', *Crime & Delinquency*, 61(3), pp. 454–480.
- Johnson, S. D., Summers, L. and Pease, K. (2009) Offender as forager? A direct test of the boost account of victimization, *Journal of Quantitative Criminology*, 25: pp. 181-200.
- Lammers, M. Menting, B. Ruiter, S. and Bernasco, W. (2015) Biting once, twice: The influence of prior on subsequent crime location, *Criminology*, 53: 309-329.
- Lazarsfeld, P. F. and Henry, N. W. (1968) *Latent structure analysis*. Boston: Houghton Mifflin Company.
- Lewin, K. (1951) *Field Theory in Social Science*. Edited by Cartwright, D. Oxford: Harpers.
- Light, R., Nee, C. and Ingham, H. (1993) *Car theft: The offender's perspective*. HM Stationery Office.
- Long, D., Liu, L., Feng, J., Zhou, S., & Jing, F. (2018) Assessing the Influence of Prior on Subsequent Street Robbery Location Choices: A Case Study in ZG City, China. *Sustainability*, 10(6), 1818.
- Manski, C. F. and Lerman, S. R. (1977) 'The estimation of choice probabilities from choice based samples', *Econometrica: Journal of the Econometric Society*, pp. 1977–1988.
- Marchment, Z. and Gill, P. (2018) Modelling the spatial decision making of terrorists: The discrete choice approach. *Applied Geography*, 104, pp. 21-31.
- McFadden, D. (1974) 'Conditional logit analysis of qualitative choice behavior', in Zarembka, P. (ed.) *Frontiers in Econometrics*. New York: Academic Press.
- McFadden, D. and Train, K. E. (2000) 'Mixed MNL models for discrete response', *Journal of applied Econometrics*, 15(5), pp. 447–470.
- McLachlan, G. J. and Peel, D. (2000) 'Finite mixture models, volume 299 of Probability and Statistics—Applied Probability and Statistics Section'. New York: Wiley.
- Menting, B. Lammers, M. Ruiter, S. and Bernasco, W. (2016) 'Family Matters: Effects of Family Members' Residential Areas on Crime Location Choice', *Criminology*, 54(3), pp. 413–433.
- Ministry of Justice. (2019) *Prolific Offenders: Criminal Pathway: Prison Events & Offender Needs*. London.
- ONS (2018a) 'Recorded crime data at Community Safety Partnership / Local Authority level'.
- ONS (2018b) *User guide to crime statistics for England and Wales*. London.

- Pacifico, D. and Yoo, H. II (2012) 'llogit: A Stata module for estimating latent class conditional logit models via the Expectation-Maximization algorithm'.
- Rabe-Hesketh, S., Skrondal, A. and Pickles, A. (2002) 'Reliable estimation of generalized linear mixed models using adaptive quadrature', *The Stata Journal*, 2(1), pp. 1–21.
- Rengert, G. F. and Wasilchick, J. (1985) *Suburban burglary: A time and a place for everything*. CC Thomas Springfield, Illinois.
- Reynald, D. M. (2014) Informal guardianship. In: Bruinsma G. and Weisburd D. (eds) *Encyclopedia of Criminology and Criminal Justice*. Springer, New York, NY.
- Rossmo, D. K. (2000) *Geographic profiling*. CRC press.
- Ruiter, S. (2017) 'Crime Location Choice: State of the Art and Avenues for Future Research', in Bernasco, W., Gelder, J.-L. van, and Elffers, H. (eds) *The Oxford Handbook of Offender Decision Making*, pp. 398–420.
- Sampson, R. J. and Groves, W. B. (1989) 'Community structure and crime: Testing social-disorganization theory', *American journal of sociology*, pp. 774–802.
- Sampson, R.J., 1985. Neighborhood and crime: The structural determinants of personal victimization. *Journal of research in crime and delinquency*, 22(1), pp.7–40.
- Shaw, C. R. and McKay, H. D. (1942) *Juvenile delinquency and urban areas*. Chicago: University of Chicago Press.
- StataCorp (2015) 'Stata Statistical Software: Release 14.', 2015. College Station, TX: StataCorp LP.
- Summers, L. (2012) The role of place in outdoor serious violence. Ph.D. Thesis. University College London.
- Townsley, M. and Sidebottom, A. (2010) 'All offenders are equal, but some are more equal than others: Variation in journeys to crime between offenders', *Criminology*, 48(3), pp. 897–917.
- Townsley, M., Birks, D., Bernasco, W., Ruiter, S., Johnson, S.D., White, G. and Baum, S.. (2015). Burglar Target Selection. A Cross-national Comparison, *Journal of Research in Crime and Delinquency*, 52(1): 3–31.
- Townsley, M. Birks, D. Ruiter, S. Bernasco, W. and White, G. (2016) 'Target Selection Models with Preference Variation Between Offenders', *Journal of Quantitative Criminology*, pp. 1–22.
- Train, K. E. (2008) 'EM algorithms for nonparametric estimation of mixing distributions', *Journal*

*of Choice Modelling*, 1(1), pp. 40–69.

Train, K. E. (2009) *Discrete choice methods with simulation*. Cambridge university press.

Train, K. E. and Sonnier, G. (2005) ‘Mixed logit with bounded distributions of correlated partworths’, in *Applications of simulation methods in environmental and resource economics*, pp. 117–134.

UK Home Office (2018) *Crime Recording General Rules*. London: Home Office.

12Van Sleeuwen, S. E. M., Ruiter, S., and Menting, B.. 2018. ‘A Time for a Crime: Temporal Aspects of Repeat Offenders’ Crime Location Choices’. *Journal of Research in Crime and Delinquency*, 55(4): pp.538-68.

von Haefen. R.H. and Domanski, A. (2018) Estimation and welfare analysis from mixed logit models with large choice sets. *Journal of Environmental Economics and Management*, 90, pp.101-118.

Weisburd, D., Bernasco, W. and Bruinsma, G. J. N. (2009) *Putting crime in its place: Units of analysis in geographic criminology*, *Putting Crime in its Place: Units of Analysis in Geographic Criminology*.

White, H. (1982) ‘Maximum likelihood estimation of misspecified models’, *Econometrica: Journal of the Econometric Society*, pp. 1–25.

Wiles, P. and Costello, A. (2000) *The ‘road to nowhere’: the evidence for travelling criminals*. Research, Development and Statistics Directorate, Home Office London.

Wright, R. T. and Decker, S. H. (1996) *Burglars on the job: Streetlife and residential break-ins*. UPNE.

Wright, R. T. and Decker, S. H. (2002) *Robbers on robbery: Prevention and the offender*, *National Criminal Justice*.