Determining how journeys-to-crime vary: Measuring inter- and intra-offender crime trip distributions.

William Smith¹

John W Bond²

Michael Townsley³

1 Thames Valley Police

2 Northamptonshire Police Scientific Support

3 School of Criminology and Criminal Justice, Griffith University, Australia

Acknowledgements

Forensic Intelligence Analysts Gemma Lyon and Helen Bates of Northamptonshire Police for providing data, their expertise and a great deal of their time. The willingness and co-operation of other Scientific Support Unit staff was also invaluable and very much appreciated. We would also like to acknowledge the feedback of Shane Johnson at various stages of the research and also thank both Kim Rossmo, George Rengert and the anonymous reviewers who provided insightful comments on an early draft.

Abstract

Journey to crime studies have attempted to illuminate aspects of offender decision making that have implications for theory and practice. This article argues that our current understanding of journey to crime is incomplete. It improves our understanding by resolving a fundamental unit of analysis issue that had thus far not received much attention in the literature. It is demonstrated that the aggregate distribution of crime trips (commonly known as the distance decay) does not take into account the considerable variation that exists between individual offenders' crime trip distributions. Moreover, the common assumption of statistical independence between observations that make up a distribution is something that, until now, has yet to be tested for distributions of crime trips of multiple offenders. In order to explore these issues, three years of burglary data from a UK police force were linked to 32 prolific offenders to generate journey to crime distributions at the aggregate and offender levels. Using multi-level models, it was demonstrated that the bulk (65%) of the variation of journeys to crime exists at the offender level, indicating that individual crime trips are not statistically independent. In addition the distance decay pattern found at the aggregate level was not, in the main, observed at the offender level – a result that runs counter to conventional wisdom, and another example of the ecological fallacy. The implications of these findings are discussed.

Introduction

Until recently criminology 'has been devoted to finding some unicausal source of motivation' (Brantingham and Brantingham, 1993, p 273) for the commission of criminal acts. Crime control policies have traditionally echoed criminological thinking, with criminal justice systems following a predominantly offender-orientated paradigm of apportioning blame and applying penalties to act as a deterrent to whatever motive compels individuals to commit criminal acts.

Eschewing offenders as the unit of analysis, environmental criminology focuses on the 'where and when' of offending (Rossmo, 2000, p112) and is considered a central theoretical foundation underpinning our understanding of the geography of crime. By concentrating on the place of the offence and the factors that directly affect offence instigation, it builds upon the work of human ecology and seeks to explain the spatial clustering of crime events in terms of an individual's interaction with their physical setting.

Most of the contributions to this volume deal with the issue of what the unit of analysis should be in studies of crime at place, and maybe even more importantly, what the appropriate criteria are for choosing a unit of analysis. Brantingham, Dyreson and Brantingham (1976) provide clear examples of the varieties of observed patterns generated by selecting different spatial units of analysis, from national to block level, illustrating the so called geographic 'cone of resolution'.

In general, measurement is not the main problem here as criminal incidents can be measured at very fine geographical resolutions; what is problematic is determining the level of appropriate aggregation. This article will argue that the most appropriate aggregation threshold is the one which *minimises* the heterogeneity *within* groups of observations while attempting to *maximise* the heterogeneity *between* groups of observations. Hierarchical linear models (or multilevel models) provide a natural framework in which to investigate this. These types of models attempt to control for the influence of factors that may be nested at different levels of aggregation (Snijders and Bosker, 1999) by incorporating multiple geographical levels of aggregation in the same model. This means effects operating by different levels can be partitioned with a view to shedding light on unit of analysis problems. Clearly, this allows a check against the well known, but perennial, problem of committing the ecological fallacy, when relationships at a higher level of geographical aggregation are interpreted in terms of relationships at a lower level.

In a study now considered classic, Robinson (1950) calculated for each state in the USA the percentage of residents who were (a) foreign born, and (b) illiterate. At the state level the correlation was -0.53 but when measured at the individual level the correlation is 0.11! The contrasting results arise by trying to equate findings at the state level with comments about individuals, and can be explained by observing that immigrants tend to move to areas where there are high literacy rates. Ultimately, the ecological fallacy is a not necessarily a

spatial unit of analysis issue, but applies to all situations where issues of aggregation need to be resolved before proper conclusions can be arrived at.

While staying within the boundaries of the main themes of this volume, this article attempts to broaden the scope of the discussion of determining appropriate units of analysis by studying a situation in which criminal incidents are grouped not by spatially nested units but by the people who commit them. Specifically, where multiple crime locations are spatially linked to a common point (residence, say) of the offender who committed them. In this case, the 'unit of analysis' challenge is not restricted to finding an appropriate way to assign crimes to spatial entities, but instead the situation forces us to start looking at broader unit of analysis questions. For example, should we be studying offenders and the locations where they reside, should we take crime events and the locations where they take place as our unit of analysis, or should we focus on the links between offenders and their crimes, i.e. should we study crime trips?"

The remainder of this chapter will focus on the journey to crime literature and the well accepted principle of 'distance decay': the frequency of offending decreases with the distance from home. As we shall demonstrate, some issues have become clouded, with researchers interpreting the findings of research conducted at the crime trip unit of analysis as findings at the offender level unit of analysis. Later, we distinguish between these two levels of analysis and assesses whether the 'distance decay' principle not only applies in the aggregate, but also describes the behaviour of individual offenders.

Theory and previous research

Within environmental criminology a number of related theories attempt to explain the prevalence of criminal acts by focusing on factors other than, but not excluding, offender motivation. Brantingham and Brantingham (1993) describe a criminal event as 'an opportune cross-product of law, offender motivation, and target characteristic arrayed on an environmental backcloth at a particular point in space-time' (p259). In essence, multiple ingredients are required for the generation of a crime, motivation representing but one of these. The dominant theories include routine activities (Cohen and Felson, 1979), rational choice (Cornish and Clarke, 1986) and crime pattern theory (Brantingham and Brantingham, 1981).

Crime pattern theory (Brantingham and Brantingham, 1981) postulates that individuals have certain psychologically intimate or familiar locations (such as work, home recreation locations) called *nodes* or anchor points. Certain routes between nodes are usually preferred over others (in terms of speed, cost, aesthetics, familiarity). These *paths* contribute to the generation of an *awareness space* (a mental map). Locations with high crime risk are produced when the cognitive maps (Canter and Hodge, 2000) of motivated offenders overlap with the spatial distribution of available targets (Cohen and Felson, 1979).

Central to crime pattern theory, and by extension spatial profiles of crime risk, therefore is the magnitude and connectedness of offenders' node network. This observation has, in turn, fostered recent academic interest in offenders' *journey-to-crime* (JTC)¹. The purpose of investigating JTC is to analyse the crime-specific travel behaviour of offenders with a view to inferring some form of offender decision making. It also has investigative value in that findings from the JTC literature may hint at directions of inquiry. For example, the JTC literature is pivotal to the research area of geographic profiling, where crime scenes are triangulated to predict an offender's anchor point (Rossmo, 2000).

The overwhelming majority of research into offender JTC has been conducted in North America. There is a general consensus within the academic community that offenders do not tend to travel far to commit crime (Rossmo, 2000; Rengert et al, 1999; Rhodes and Conly, 1981; Paulsen and Robinson, 2004; Chainey and Ratcliffe, 2005). A comprehensive review of the JTC literature can be found in Rossmo (2000, ch 10)² in which he compiles the results of no less than 38 different JTC studies (consisting of 80 offence-specific JTC estimates). He observes that JTC estimates are calculated in one of four ways: (a) mean distance (usually the arithmetic mean but sometimes the geometric mean is used to restrict the influence of outliers); (b) medial circle (the radius of a circle which captures a certain proportion of crime trip distances); (3) mobility triangles and (4) distance decay functions. Rossmo concludes that the latter of these is the most advantageous as it captures a degree of the variation in crime trip distances, but out of the 38 studies considered only eight employed this method. The other three measures suffer due to an inability (to varying degrees) to capture the variability of individual crime trips. That is they are a measure of central tendency but not dispersion. Nonetheless, the key findings of Rossmo's review are that:

- crimes occur in close proximity to the offender's residence;
- there is a distance decay pattern for crime trips;
- juvenile offenders are less mobile than adult offenders; and
- JTC varies by crime type.

Wiles and Costello (2000) conducted one of the few JTC studies outside North America, carrying out extensive research into the travelling habits of offenders throughout South Yorkshire. Their main findings were broadly consistent with the Rossmo review and they further observed that in the main offenders did not set out with the intention to commit crime, but instead responded to the opportunities that presented themselves during their routine activities. They concluded that the interlinked concepts of offender rationality and routine activities varying in accordance with the spatial distribution of opportunities were still the dominant driving force behind the JTC.

¹ Levine (2005a) notes that researcher interest in JTC has considerable history citing studies in the 1930s.

² Levine (2005a) also provides an overview of the JTC literature but usefully places it in the context of the study of travel across a range of fields. Groff and McEwan (2006) also provide a thorough, recent overview of the JTC literature.

Two recent studies have sought to clarify our knowledge of JTC and its relevance to criminology. van Koppen and de Keijser (1997) were sceptical that aggregate distance decay functions represented individual JTC distributions appropriately. They felt it was feasible that JTC patterns at the offender level displaying no distance decay could be combined to show an aggregate distance decay and that JTC researchers could be at the mercy of an ecological fallacy when commenting on patterns observed in the aggregate. They demonstrated this by simulating a sample of offenders, none of whom conformed to a distance decay function. When the crime trips of each offender were combined, the aggregate distribution displayed the conventional distance decay pattern.

A critique by Rengert et al (1999) strongly rebutted van Koppen and de Keijser's argument by focussing on four areas: (a) the interpretation of the ecological fallacy; (b) assumptions of offender movements; (c) the interpretation of geographic profiling; and (d) random target selection assumptions and range of operations. The second and fourth of these are methodological issues relating to how van Koppen and de Keijser implemented their simulations. Rengert et al demonstrated that van Koppen and de Keijser, by not properly accounting for a uniform distribution of opportunities per distance unit, inadvertently simulated offenders that *did* possess distance decay at the individual level.

Rengert et al (1999) also take issue with van Koppen and de Keijser's use of the ecological fallacy to undermine the integrity of JTC studies. van Koppen and de Keijser (1997) state that "the distance decay function is an aggregated function" which "shows that the number of crimes is inversely related to the distance from criminals' residences. This, however, does not necessarily imply that *individual criminals* commit more crimes closer to their home than far away" (p507, emphasis added). Rengert et al reply that because aggregate distance decay functions are generated from individual data then drawing inferences about the individual from the aggregate is not fallacious. They argue that in the context of aggregate JTC distance decay functions that 'the ecological fallacy would not be present since the analyses derive from individual-level data and inferences at this level of analysis are appropriate. This is especially true if distance decay exists at the *individual level*' (1999, p432-3, emphasis added).

It appears that both groups of researchers are talking at cross purposes and the confusion is compounded as the non-specificity of the term 'individual'; it could refer to an individual offender or an individual crime trip. In the interests of clarification we will refer to a single JTC as a *crime trip*. A number of crime trips will be taken by an individual *offender*. A number of offenders will make up the *population* of offenders. The aggregate distance decay function is, therefore, an estimation³ of the population of all crime trips taken by all offenders.

Restating the quotes above: (a) van Koppen and de Keijser argue that an aggregate distance decay function cannot be used to infer patterns at the

³ Assuming that only a sample of all offenders (and their crimes) are represented in the data.

offender level; (b) whereas Rengert et al defend the claim that most crime trips are short because data on individual crime trips has merely been aggregated. In their respective frames of reference, both groups are right! The confusion lies with making comments about offenders. While most crime trips might be short, we cannot be certain that most offenders make short crime trips. It is inappropriate to draw inferences about *offenders* from the aggregation of *crime trips*, as we have no information about how these crime trips are nested at the offender level. A possible exception is if all offender level JTC distributions take on the same (approximate) shape (as intimated in the earlier quote by Rengert et al, 1999) and they all possess similar values of central tendency.

Consideration of the distributions of individual offender crime trips have not featured prominently in the JTC literature. Exceptions include Kocsis et al (2002), Canter and Larkin (1993), along with four replications cited in Canter et al (2000) and Rengert et al (1999) cite a number of studies they claim have demonstrated offender-level distance decay (Alston, 1994; Rengert, 1996; Rossmo, 1993; and Warren et al, 1995).

LeBeau (1992) reports the travel distributions of four serial rapists (shown here in Figure 1)⁴. It is immediately apparent that, while intra-offender JTC patterns conform to conventional distance decay patterns - the aggregate picture is much the same - there is considerable inter-offender JTC variation. The JTC distribution of offender 1 is well removed from the others and is easily explained by his place of employment located about twenty-six kilometres from a city. Offenders 2 and 3 did not travel very far to offend, offender 3 especially so with a very restricted range of JTC. The JTC distribution of offender 4 displays the most variation; although offender 2 has a wider range and greater standard deviation, this is due to the presence of an outlier and the bulk of the distribution is confined to less than five kilometres.

[insert Figure 1 about here]

Snook (2004) provides an exploration of inter-offender JTC differences of serial burglars in St John's, Canada. He found that median travel distances were associated with age, method of transport and value of property stolen, but not experience, series length (events and time), offending with partners and target type. An aggregate distance decay function is provided for the sample (347 crime trips over 51 offenders) as are descriptive statistics (N, median, mean and standard deviation) of crime trip distributions by individual differences (age, transport, etc as above). The degree of relative dispersion is very high. Computing the co-efficient of variation⁵ for the sample across each individual difference, excluding low frequency method of transport categories, reveals ratios ranging from 50 to 120 percent. This suggests that summaries

⁴ LeBeau also usefully provides a host of other data (date, time, approach method, location, relationship to victim) allowing for further inter-offender comparisons.

⁵ The co-efficient of variation is the standard deviation divided by the mean expressed as a percentage. It is often used to assess the degree of variation of a group of observations while controlling for the magnitude of the observations.

of the JTC distribution at the aggregate level mask considerable variation at possibly the inter-offender level, the intra-offender level or some combination of the two.

Research question

To summarise, the consensus in the JTC literature is that ‘most offenders do not travel far to commit crime’ or ‘most crime trips are short(er) than long’. However, the majority of published studies have not accounted for the nesting of crime trips within offenders, making the former statement unjustified based on the evidence and methods used. If observations are nested within a particular factor, it is inappropriate to make comments about the influence of that factor if it has been ignored! Inferring offender travel characteristics from crime trip data risks committing the ecological fallacy. Of course, just because nesting occurs does not mean it is necessarily relevant. Offenders can be grouped according to their hair colour but it would indeed be surprising if this influenced patterns greatly.

The focus of this study is to explore the extent of variability within JTC distributions with a view to better understanding the degree to which uniqueness or generality drive the well established and accepted aggregate patterns of JTC (i.e. aggregate distance decay). In other words, we seek to describe the extent of variation at the inter-offender level and the intra-offender level, with a view to determining whether the influence of grouping is large enough to raise concerns about a possible aggregation bias. If so, then comments about offender travel characteristics are not justified on the basis of aggregated crime trip data and further research is required to recover this information. If not, then comments about offender characteristics can be made on the basis of crime trip data.

This study is primarily concerned with establishing whether a methodological problem, overlooking the nesting of crime trips within offenders, has compromised our understanding of JTC.

Data and Methods

The area chosen for investigation was Northamptonshire Police Force, a semi-rural police force in the East Midlands region of the United Kingdom. As the literature revealed JTC estimates vary by crime type, only one offence was selected for the study, residential burglary. Two primary data sources were used; offence data and offender data. A three year time window for offence data was chosen after a preliminary investigation indicated that any shorter time frame would provide an insufficient sample set of offenders and their associated JTC distributions. There were a total of 14,217 offences in the period 2002 to 2004 inclusive. Of these, 2341 (16.5%) resulted in a detection (i.e. were linked with an offender and resulted in arrest and charging). The full postcode, date reported, crime beat and offender name of the 2341 burglaries were extracted from the recorded crime information system.

The second source of data was offender information. All 1083 offenders featured in the detected burglary list were included. For each individual the following fields were extracted from the Force Intelligence System (FIS): name, date of birth, occupation and full postcodes of all known residences. As some individuals had multiple addresses in the three year time period of this study, the tenure of each different address was determined by entries on the FIS. This information was originally obtained through a combination of housing records, police stop and searches, arrests for related matters and other police intelligence placing offenders as resident at specific addresses. While a certain degree of imprecision remains, we believe the above approach is about as accurate as is possible from police data. The anchor point chosen to examine was the primary residence of the offender as the quality of police data concerning other anchor points was not of sufficient quality to warrant considering.

Location information for crime incidents and offender residence in this study was determined using the full postcode. In the UK a full postcode relates to, on average, 15 'delivery points' – literally letterboxes. In urban areas this typically means distinct dwellings or households (in the study area there was a trivial amount of high density housing). Full postcodes, therefore are a reasonable compromise for data protection concerns. Point information for residences of crime scenes and offender residences were determined by establishing the centroid of the corresponding full postcode areal unit (Ordnance Survey, 2006).

The distribution of offending rates revealed that many offenders were responsible for only one or two burglaries. A natural break in the offending frequency was located between five and ten offences. It was decided only burglars detected for ten or more offences were included in the remaining analysis. This resulted in only 32 offenders who had been detected for 603 burglaries. In addition three postcodes from offender residences and one from burglary locations were unable to be reconciled with the postcode list available through Ordnance Survey. These anomalies resulted in the net loss of 13 crime trip distances (all of the offenders with incorrect postcodes had valid postcodes for other periods of the time frame). The final sample therefore consisted of 32 offenders responsible for 590 burglaries and crime trips.

JTC distances were established by computing the euclidean distance between the point estimate for the offence location and the point estimate of the most likely offender residence. This distance measure was chosen as manhattan distances are usually thought to be more appropriate for grid-based urban layouts, i.e. North American cities, whereas Euclidean distances are used for European areas. Using UK data, Rossmo et al (2004) presented the surprising result that manhattan distances provide more precise estimates than euclidean distances by using a street routing distance as a baseline. The results that follow could easily be replicated using a manhattan metric with trivial differences observed.

It should be noted that a degree of imprecision exists with the manner of the distance computed here. The derived crime trip distance is based on point estimates (centroid locations of postcode area units), not the locations themselves. Two distinct issues arise from this decision; the generation of zero distances (where the offender resides in the same postcode as the victim) and inaccurate distance measurements due to the difference between the true locations and the point estimates. The latter will be dealt with first.

The difference between a single true location and its corresponding point estimate will be called its *offset*. The *offset effect* is the difference between the true distance and the estimated distance (in other words the combined impact of both offsets). The maximum offset for any point estimate will be dictated by the geographic morphology of the postcode polygon. If the size of the postcode polygons is relatively small, the offset for any point will be small and by extension the offset effect should be minimal.

The areas of all postcode boundaries in the Northampton area (Ordnance Survey, 2006) were analysed using a GIS. The distribution of these area amounts was highly skewed toward very small areas. The 75th percentile of the postcode areas was only 0.03 of a square kilometre, or less than three football pitches.

In order to investigate how offsets might influence the estimate of distances for this study, a simulation was performed (for full details see the footnote⁶). The ratio of true to estimated distances was normally distributed around a mean of one and a standard deviation inversely proportional to the distance between the simulated postcodes. We found that when simulated postcodes were about 350 meters apart the distribution of the ratio of true to estimated distance was relatively stable⁷. The observation that the distribution of the ratio of true to estimated distances takes on a symmetric distribution indicates that the estimated distance will both over- and under- estimate the true distance, plausibly cancelling each other out in the aggregate.

A more important point than these methodological issues is the purpose to which these data will be put. The focus of this study is to quantify the influence of individual offenders on the aggregate JTC distribution. That is, we seek to establish whether JTC distributions at the individual level are the same as the JTC distribution at the population level. To this end, we are not concerned with the physical distances observed (at the aggregate, offender or crime levels) as different crime types and different areas will always limit the

⁶ For the simulation two circular 'postcodes' were defined, with radii of one hundred metres, and located a fixed distance apart. These corresponded to the same area as the 75% percentile postcode. The postcode centroids are the centres of each circle, thus the fixed distance is the 'estimated' distance between the two locations. At each iteration, a true location was randomly located at any point within the postcode boundary (by randomly selecting polar co-ordinates). The distance between each true location was computed and compared to the estimated distance. 10,000 iterations were performed. A range of fixed distances were used to test whether the choice of fixed distance was important.

⁷ Well behaved normal distributions typically have a co-efficient of variation of less than twenty percent. When the fixed distance was greater than 350 meters the standard deviation of the ratio of true to estimate distances was less than 0.2.

external validity of findings, but rather whether distances are influenced by different offenders, and whether this degree of influence is large enough to be concerned about. Moreover, due to the study objectives it is the 'relative difference' between two distances which is important rather than the specific units of measurement. For instance, for two distances 500 meters and 1500 meters, the observation that one is much shorter than the other is the more pertinent issue than the values per se, with respect to the purpose of the study. For these reasons we are happy to tolerate a degree of imprecision in the distance calculations as long as we are confident this will not compromise the relative differences in distances.

The second data generation issue concerned generating a higher number of zero distances than expected. This is a problem because the presence of zero distances will tend to broaden the range of that offender's JTC distribution, possibly compromising the analysis. The data contained only 24 observations with the offender and victim residing in the same postcode (about 4% of the total). These 24 observations were distributed among nine offenders, with one offender accounting for nine of the 24 zero distance observations. The presence of the zero distances did not appear to affect the JTC distribution of this offender. The remaining 15 zero distances were distributed relatively evenly among the other eight offenders. It was felt these would not compromise the overall results.

Offender age was determined by the number of days between birth and the start of the time period. The distribution of ages revealed a trimodal distribution, with natural breaks at roughly 17 and 26 years. Offenders were partitioned into three groups - young ($n = 5$), middle ($n=20$) and older ($n = 7$) - according to these natural breaks. These also seemed to mirror significant legal thresholds (obtaining a drivers license) and the termination of criminal careers in early adulthood (Blumstein et al, 1986).

In order to quantify the degree of heterogeneity that resides at each level, multilevel models were employed. Multilevel models are the term used by social scientists to describe modelling approaches which take into account the grouping, or nesting, of data (Pinheiro and Bates, 2000; Venables and Ripley, 2002). The canonical example is that of student test performance of individuals within different school classes, some of which are located in different schools, which in turn might have different catchment areas (neighbourhoods). While factors such as parent education, IQ and gender may influence *individual* test scores, there also potentially exists effects operating at cruder levels of resolution, such as the class level (quality of individual teachers may vary), the school level (some schools may foster scholastic achievement) and the neighbourhood level (some catchment areas vary demographically), all of which would influence *groups* of test scores. The methodological issue is that multiple observations (pupils) exist within units (teachers, schools, etc) and the influence of grouping factors should be partitioned from relationships that exist at the individual level. Without accounting for these grouping effects, the relationships between factors at the individual level are likely to be biased. For criminological applications of multilevel models see Rountree et al. (1994), Sampson et al. (1997), Jang

(1999), Jang and Johnson (2001), Morenoff et al. (2001), Oberwittler (2004), Tseloni (2006) and Gelman et al. (2007).

In the context of longitudinal data, where repeated measures from a group of individuals are studied, multilevel models can assist in separating the correlation between repeated measurements from correlations with potential explanatory factors. The utility of multilevel models for this study is that it will be possible to determine the proportion of total JTC variability that exists at the offender level. This is calculated using the *intra-class correlation co-efficient* (ICC). Imagine a two stage sampling design where measurements are collected at a micro level (crime trips) which can be grouped into a macro level unit (prolific offenders). The ICC is calculated as the variance *between* macro units divided by the total variance of the sample (Snijders and Bosker, 1999); in other words the ICC is the fraction of the total variance which is accounted for by the macro level⁸.

So, the magnitude of the ICC implies the extent of the influence of the macro unit clustering on the outcome variable. Values of zero (or close to) mean that while observations may be clustered, this grouping does not exert an influence at the individual level (alternatively, the use of regression techniques not accounting for grouped data would be justified). It follows that the ICC statistic can be subjected to an hypothesis test to determine whether its true value is equal to zero or not (see Snijders and Bosker, 1999, p 21).

All analysis in this study was conducted in the statistical programming environment R (R Core Development Team, 2004). The *nlme* package was used to implement multilevel models (Pinheiro et al, 2007).

Results

The sample of 590 crime trips yielded the aggregate distance decay distribution, estimated by a kernel density procedure, shown in Figure 2. We use this method of estimating the density of these data rather than a histogram because the latter is susceptible to both choice of interval size and their locations, whereas kernel density estimation is chiefly reliant on the choice of bandwidth. The distribution had a mean of 3.7 kilometres with a standard deviation of 4.4 kilometres. The median distance travelled to burglary was 2.2 kilometres, consistent with the visual impression of skewness. The degree of variation observed is extreme, yielding a co-efficient of variation of 119 percent.

[insert Figure 2 about here]

The aggregate distance decay function in Figure 2 conforms to the functional form commonly referred to in the research literature; that is, short crime trips

⁸ There is another interpretation of the ICC - its numerical value is equal to the correlation between two randomly selected observations in the same randomly selected macro unit - but its meaning will not be appropriate in the context of this study.

are common and long trips are rare. The mode of the distribution is located at approximately 700 metres.

Figure 2 allows a preliminary picture of the general pattern of JTC trips undertaken by prolific burglars, but the magnitude of variation within the crime trip distribution is considerable. The aggregate distance decay function was recomputed, this time conditioned by the offender age group, and is shown in Figure 3.

[insert Figure 3 about here]

There are some fairly striking features in Figure 3. The most obvious is the restricted range on the crime trip distribution of younger prolific burglars compared to the other two groups. Almost all the crime trips committed by offenders 17 years or less are located less than five kilometres from their residence. The middle and older offenders display, on aggregate, a more typical distance decay function where long trips are rare but nevertheless exist.

A second feature of Figure 3 is the bimodal similarity of each age group's crime trip distribution. Each displays a global maximum value at a relatively short distance before declining rapidly, but a local maximum occurs at a distance beyond the central mass of the distribution. For middle offenders this local maximum is around ten kilometres and older offenders have one slightly further than this. When viewed on an absolute scale such as Figure 3 the local maximum appears muted but in relative terms is of considerable magnitude.

Retaining the three age groups, Figure 4 contains box and whisker plots of each prolific offender's crime trip distribution. Box and whisker plots were chosen here for reasons of space; they contain measures of both central tendency and dispersion and succinctly summarise a distribution. For purposes of clarity they have been ordered on the basis of their median value.

[insert Figure 4 about here]

A tremendous amount of information is contained in Figure 4. The first point to note is that within each age group there appears to be offenders whose JTC distribution is sharply restricted to a certain geographic range, whereas other distributions display a great deal of variation (indicated by the height of the boxes and whiskers on each plot). Typically, an offender with a low median crime trip displays smaller variation, although the older offenders appeared to buck this trend somewhat.

The second notable observation is that, consistent with Figure 3, age appears to be correlated with extent of geographic range. Younger offenders have a lower geographic range compared to middle and older offenders. An explanation for the age-conditioned bimodal distribution observed in Figure 3 presents itself. Note that in each age group in Figure 4 there appears to be two clusters of offenders, ones with low medians and those with higher

medians. It seems that the local maxima (located at greater distances than the global maxima) observed in Figure 3 are generated largely by distinct offenders and are *not* the result of contributions from all offenders in each age group.

The last substantial observation is that there is also a variety of combinations of central tendency and dispersion displayed by different offenders. Some offenders show very different ranges of operation, yet have equivalent medians. The degree of variation at the offender level is large. While a large number of offenders do offend at short distances there are also some prolific offenders who predominantly do not.

Figure 4 provides compelling initial evidence that a variation of JTC exists at both the intra-offender and inter-offender levels. The aggregate crime trip distribution does not appear to be consistent with distributions at the offender level, even after controlling for the age groupings. In other words, there appears to be an influence of nesting of crime trip distances with different offenders. This observation justifies efforts to quantify the extent of variation existing within and between offenders. In order to explore the extent of the nesting effect, multilevel models will need to be employed.

Measuring the extent of intra- and inter- offender variation

As implied earlier, the prolific burglars' crime trips are a form of longitudinal data. For the purposes of this study, we are not trying to establish causal relationships (e.g. distance travelled increases with perceived affluence of victim), but merely to ascertain the degree to which the length of individual crime trips are influenced by the offender. In that vein, a random intercept model⁹, where the intercept term is allowed to take on different values for each offender, was chosen to fit the data. As no explanatory variables are included in the model, i.e. an 'empty model', no consideration towards random slopes for each level is required.

An empty model with random intercepts for each offender was fitted to the 32 prolific offenders and 590 crime trips. The estimated ICC was computed at 0.65. This means that nearly two thirds of the total crime trip distance variation for these data exist at the offender level. Alternatively, there was a high degree of correlation of crime trip distance within offenders. A test revealed extremely strong evidence against a null hypothesis of a zero ICC ($F = 32.3$, $df = (31, 558)$, $p < 0.0001$).

Skewness scores

The final analysis conducted was to investigate the JTC distributions of the individual offenders separately. The approach taken here is to compute a skewness score for each offender. Skewness is a measure of asymmetry in data and has positive values if the right tail of a distribution is longer than the left tail and negative if the reverse is true. If shorter trips are more common than longer ones, as suggested by the JTC literature, offenders should display

⁹ Rasbash et al (2004) use the terminology *variance components* models.

positive skewness in the main. Values of zero conform to a symmetric distribution.

It is possible to compute a standard error of a skewness estimate¹⁰, which means the skewness estimate divided by its standard error reduces to a Z score. This means that for individuals with Z scores within two standard deviations of zero, we cannot be confident (at 95%) that the actual skewness is different from zero. For example, the skewness estimate of the distribution shown in Figure 2 is roughly 1.95 and has a standard error of about 0.1. The skewness Z score for the JTC distribution ignoring grouping within offenders is approximately 19.4 ($=1.95/0.1$). Given the estimate is located almost 20 standard deviations from zero, we can be very confident that the distribution in Figure 2 is not symmetric (of course, one only needs look at the figure to tell this).

Based on the consensus in the literature we would expect that most offenders would display positive values of skewness and that many of these would be reliably different from zero (i.e. with values greater than 1.96). Figure 5 contains the skewness Z scores for each offender in the sample.

[insert Figure 5 about here]

The distribution of skewness Z scores for this sample is surprising and unequivocal. Nearly *half* the offenders have *negative* skewness scores, although only two of these have a magnitude reliably different from zero. Of the remainder, those offenders displaying the expected distance decay pattern, only seven possess a degree of asymmetry reliably different from zero (high positive scores)¹¹. It appears that the source of the skewness observed in the aggregated crime trip distribution (Figure 2) comes from either a minority of the individuals in this sample or is a product of ignoring the grouping of these data.

Discussion

In this article, the issue of determining appropriate units of analysis was applied to a situation where multiple criminal incidents are spatially related to the residence of the offender who committed them, a situation that requires careful reflection on the proper unit of analysis. Using multi-level models, we challenged conventional wisdom and discovered another ecological fallacy: the distance decay pattern found at the aggregate level is generally not observed at the offender level.

The results presented in this study suggest that aggregate crime trip distributions contain considerable variation which was not evenly distributed

¹⁰ Equal to $\sqrt{\frac{6}{N}}$ (Tabachnick and Fidell, 1996, p 72).

¹¹ One of these positively skewed JTC distributions belonged to the individual with a sizable number of zero distances mentioned in the Data section. Thus, even though the effect of the presence of zero distances is to bias the observed skewness estimate downwards, this was not sufficient to compromise the statistical reliability for this offender.

among (prolific) offenders. Individual offender crime trip patterns differed remarkably in terms of their location in space (central tendency) and their spread to the extent that aggregate distance decay functions appear to be only appropriate for inferring features of the population of crime trips. Further, the fact that grouping does have an effect means that crime trip data cannot be considered statistically independent. The theoretical and operational implications of this are discussed as follows.

The theoretical importance of JTC research is that it aids in inferring criminal behaviour and explaining crime patterns. Indeed, Rengert et al state no less that '[i]mplications arising from such research touches the majority of criminological theories' (1999, p427). The aggregate distance decay function is commonly thought to imply that the majority of offenders favour targets closer to their home (or another anchor point) compared to potential targets more distant.

The ICC statistic showed that two thirds of the JTC variation resides between offenders, suggesting that the unit of analysis of most relevance is the offender, not individual crime trips. Consequently, a further major implication of this result is that the JTC patterns of individual offenders are far less variable than the distribution of all crime trip distances imply. In other words we can say that, compared to the population of crime trips taken, there was a greater degree of consistency of distances travelled by individual offenders.

Aggregate JTC distributions certainly give the impression that the bulk of crime trips are short: because they are! What the results of the multilevel model suggest is that there is an influence of grouping (offenders) operating which means that the observations (crime trips) cannot be considered statistically independent. Apart from the obvious issue of the presence of an aggregation effect, if statistically dependent observations are combined the resulting distribution will be biased. Thus, the aggregate distance decay function - the accepted conception of JTC patterns - as currently expressed may be invalid.

Finally, the extent of skewness at the population level is not consistent at the offender level. Almost *half* of the sample displayed skewness in the *opposite* direction that the literature would predict. Of those offenders displaying positive skewness (i.e. a distance decay pattern) only a minority could be said to be reliably asymmetric. This is further evidence of the influence of crime trips nested at the offender level

We find support for van Koppen and de Keijser's argument that comments about offenders drawn from an aggregate distance decay function are fallacious by virtue of an ecological bias. Moreover, we found evidence of significant differences between offenders, which weakens (somewhat) Rengert et al's contention that the aggregation of crime trip distances is unlikely to be subject to the ecological fallacy, especially if offenders are similar to each other.

An unexpected finding in this study was the presence of a group of offenders who appeared to make no short crime trips; depending on criteria this group made up between a quarter to a third of the sample. It is conceivable that these offenders represent commuting offending behaviour (Canter and Larkin, 1993). If so, omitting these individuals from the analysis may alter the findings. It could be that a sizable proportion of the inter-offender variation comes from the contrasting JTC patterns of 'commuters' and 'marauders'. Even so, it at least hints at the relative proportion of both types of offending styles.

In practical terms, the results have some operational bearing. The observation that crime trips committed by the same offender are more consistent than crime trips committed by different offenders strengthens the general rationale of geographic profiling. If offenders are relatively invariant in their own JTC, then determining a node for a single offender, usually by a process of triangulation using a distance decay curve, should be possible with greater precision than previously imagined. However, if the function used to weight distances is computed by aggregating crime trip distances made by a group of offenders, our results suggest then the intensity surface generated for the single offender in question may be vulnerable to aggregation bias problems. The extent to which this is the case will depend on how typical the offender being profiled is and the degree of inter-offender variation among the group of offenders. The results of our research indicate that the latter is considerable and the very nature of geographic profiling is such that, apart from complete speculation, the former will never be known prior to apprehension.

In fact the argument could be extended to apply to *any* distance weighting function derived to emulate the general finding that most crime trips are short. If the parameter values are chosen so that the distance weighting function resembles the aggregate distance decay function then the calculation of the intensity surface does commit the ecological fallacy. Of the three major geographic profiling applications (Rigel, CrimeStat¹² and Dragnet), the second explicitly allows the use of an aggregate JTC curve to be used as a weighting function (Rich and Shively, 2004).

There are three main weaknesses to the data used in this study, although they are certainly not unique to studies of this type. The first is the assumption that the most relevant node for any JTC will always be the residence. The assumption here may be arguable, but it nevertheless presents a picture of territorial range. The second major weakness is data quality. These JTC represent only detected crimes, we can say nothing about those crimes not detected, whether committed by offenders in our sample or not. This is a feature common to all JTC studies and no obvious solution presents itself. Nevertheless, we have attempted to improve the quality of data by factoring in changes to offenders' residence. Finally, we have not attempted to control for the supply of opportunities between home residence and crime location. This may very well yield different patterns than those observed using a physical distance measure.

¹² CrimeStat contains a Journey to Crime Estimation routine, but is geographic profiling in all but name.

Further aspects to bear in mind are that these data relate to volume crime, so the sampling of offenders and their activities is more likely to be incomplete compared to the corresponding patterns of serious crime. The perniciousness and rarity of serious crime means that a much higher proportion of crime trips can be determined for serious offenders compared to volume offenders. Also, the offender sample used here are prolific offenders, a minority of the largely low rate opportunistic individuals who might make up the active burglar population. 'Prolifics' were selected to provide a degree of precision in the individual JTC distributions because while only comprising three percent of the known burglar population they account for about a quarter of all detected burglaries.

The design of this study is broadly consistent with other JTC studies that make comments about offenders based on the aggregate crime trip distribution. Thus, despite the shortcomings of these data and the assumptions underpinning the analysis there is no reason to suggest that the results observed here could not also be found in the datasets of other studies had the patterns been sought. In other words, the fact that these 'weaknesses' are present in this study does not take away from the validity of the findings as they also exist in an array of other JTC studies.

Some readers might be surprised at the focus and findings of this study. Surely distance decay is a fairly incontrovertible phenomenon? Our point is that much of what is conventional knowledge in criminology about JTC comes from studies which have overlooked a serious methodological problem. If our argument is persuasive, this need to be addressed. This is *not* the same as saying that everything we know about distance decay is wrong, but that for distance decay to be valid we need to demonstrate it in an appropriate way. It is not enough to be correct, we must be correct for the right reasons.

Given the results of this study, future work on JTC should focus on replicating the analysis presented here on a larger sample as well as estimating the influence of offender-specific factors from crime-specific factors. Snook (2004) attempted to investigate individual differences among a group of prolific burglars and, in a limited way, target attractiveness. A logical extension of the work outlined here would be to include explanatory variables at the crime trip and offender levels as a multi-level model. This would allow the influence of variables which operate at different levels to be isolated and estimated without running the risk of committing the ecological fallacy.

References

- Alston, J.D (1994) *The serial rapist's spatial pattern of target selection*. Unpublished Master's Thesis, Simon Fraser University, Burnaby, Canada
- Brantingham, P. J., Dyreson, D. A., & Brantingham, P. L. (1976). Crime seen through a cone of resolution, *American Behavioral Scientist*, 20(2), 261–73.
- Brantingham, P.J. and Brantingham, P.L. (1981). *Environmental Criminology*. Beverly Hills, CA: Sage Publications.
- Brantingham, P.J. and Brantingham, P.L. (1993). 'Environment, Routine, and Situation: Toward a Pattern Theory of Crime', In R. V. Clarke and M. Felson (Eds.) *Routine Activity and Rational Choice* (pp. 259-294). New Brunswick, N.J.
- Canter, D.V., and Larkin, P. (1993). 'The environmental range of serial rapists', *Journal of Environmental Psychology*, 13, 63-69.
- Canter, D.V., Coffey, T., Huntley, M. and Missen, C. (2000) 'Predicting Serial Killers' home Base Using a Decision Support System', *Journal of Quantitative Criminology*, 16(4), p457-478
- Chainey, S.P. and Ratcliffe, J.H (2005). *GIS and Crime Mapping*. London: Wiley
- Cohen, L. and Felson M. (1979) 'Social change and crime rate trends: A routine activities approach', *American Sociological Review*, 44, pp 588-607
- Cornish, D.B. and Clarke, R.V. (1986), 'Introduction', In D.B. Cornish and R. V. Clarke (Eds.) *The reasoning criminal: Rational choice perspectives on offending*, New York: Springer-Verlag.
- Gelman, A., Fagan, J. Kiss, A (2007) An Analysis of the New York City Police Department's by "Stop-and-Frisk" Policy in the Context of Claims of Racial Bias, *Journal of the American Statistical Association*,
- Jang, S.J. (1999) "Age-Varying Effects of Family, School, and Peers on Delinquency: A Multilevel Modeling Test of Interactional Theory", *Criminology*, 37(3): 643–686.
- Jang, S.J. and Johnson, B.R. (2001) "Neighborhood Disorder, Individual Religiosity, and Adolescent Use of Illicit Drugs: A Test of Multilevel Hypotheses", *Criminology*, 39(1):109–144
- Groff, E.R. and McEwan, T. (2006) Exploring the Spatial Configuration of Places Related to Homicide Events, Alexandria, VA: Institute of Law and Justice

- Kocsis, R.N., Cooksey, R.W., Irwin, H.J. and Allen, G. (2002), 'A Further Assessment of "Circle Theory" for Geographical Psychological Profiling', *Australian and New Zealand Journal of Criminology*, 35(1), pp 43-62
- LeBeau, J. (1992) 'Four case studies illustrating the spatial-temporal analysis of serial rapists', *Police Studies*, 15, pp 124-145
- Levine, N. (2005a). *User Manual for CrimeStat III*, Ned Levine & Associates: Houston, TX National Institute of Justice: Washington, DC
- Levine, N. (2005b) *CrimeStat III: A Spatial Statistics Program for the Analysis of Crime Incident Locations*, (version 3.0). Ned Levine & Associates: Houston, TX National Institute of Justice: Washington, DC
- Morenoff, J., Sampson, R.J. and Raudenbush, S. (2001), 'Neighborhood Inequality, Collective Efficacy, and the Spatial Dynamics of Urban Violence', *Criminology*, 39, 517-560.
- Oberwittler, D. (2004) "A Multilevel Analysis of Neighbourhood Contextual Effects on Serious Juvenile Offending: The Role of Subcultural Values and Social Disorganization", *European Journal of Criminology*, 1(2): 201 - 235.
- Ordnance Survey (2006) © Crown Copyright/database right 2006. An Ordnance Survey/EDINA supplied service
- Pinheiro, J. and Bates, D. (2000), *Mixed-Effects Models in S and S-Plus, Statistics and Computing Series*, London: Springer
- Pinheiro, J., Bates, D., DebRoy, S. and Sarkar, D. (2007). *nlme: Linear and Nonlinear Mixed Effects Models*, R package version 3.1-83.
- R Development Core Team (2004). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Rasbash, J., Steele, F., Browne, W.J. and Prosser, B. (2004). *A User's Guide to MLwiN*, Version 2.0. London: Institute of Education
- Rengert, G.F. (1996) *The Geography of Illegal Drugs*, Boulder, Colo.: Westview Press
- Rengert, G.F., Piquero, A.R. and Jones, P.R. (1999). 'Distance decay re-examined', *Criminology*, 37, 427-445.
- Rhodes, W.M. and Conly, C. (1981). 'Crime and Mobility: An empirical study', In P. J. Brantingham and P. L. Brantingham (Eds.) *Environmental Criminology* (pp.167-188). Beverly Hills, CA: Sage Publications.
- Rich, T. and Shively, M. (2004) *A Methodology for Evaluating Geographic Profiling Software*, Cambridge, MA: Abt Associates

- Robinson, W. S. (1950). Ecological correlations and the behavior of individuals, *American Sociological Review*, 15, 351-357.
- Rountree, P.W., Land, K.C. and Miethe T.D. (1994) "Macro-Micro Integration In The Study Of Victimization: A Hierarchical Logistic Model Analysis Across Seattle Neighborhoods", *Criminology*, 32 (3): 387–414.
- Rossmo, D.K. (2000), *Geographical Profiling*, Boca Raton, FL: CRC Press LLC.
- Rossmo, D.K. (1995). *Geographic profiling: Target patterns of serial murderers*. Unpublished doctoral dissertation, Simon Fraser University, Burnaby, BC.
- Rossmo, D.K. (1993) 'Multivariate Profiles as a Tool in Crime Investigation'. Chicago: Workshop on Crime Analysis Through Crime Mapping
- Rossmo, D.K. (1987). *Fugitive migration patterns*. Unpublished master's thesis, Simon Fraser University, Burnaby, BC.
- Rossmo, D.K., Davies, A. and Patrick, M. (2004) *Exploring the geo-demographic and distance relationships between stranger rapists and their offences*, Special Interest Series Paper No 16, London: Home Office
- Sampson, R.J., Raudenbush, S. and Earls, F. (1997) 'Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy', *Science*, 277, 918-24.
- Snijders, T. and Bosker, R. (1999), *Multilevel Analysis: An introduction to basic and advanced multilevel modelling*, London: Sage
- Snook, B. (2004), 'Individual differences in distance travelled by serial burglars', *Journal of Investigative Psychology and Offender Profiling*, 1, 53-66.
- Tabachnick, B.G. and Fidell, L.S. (1996) *Using Multivariate Statistics* (3rd Ed), New York (NY): HarperCollins College Publishers
- Tseloni, A. (2006) "Multilevel modelling of the number of property crimes: household and area effects", *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169(2):205–233
- van Koppen, P.J. and de Keijser, J.W. (1997). 'Desisting distance decay: On the aggregation of individual crime trips', *Criminology*, 35, 505-515.
- Venables, W.N. and Ripley, B.D. (2002) "Modern Applied Statistics with S-plus", 4th Edition, Springer-Verlag.
- Warren, J., Reboussin, R. and Hazelwood, R. (1995) *The Geographic and Temporal Sequencing of Serial Rape*, Final Report submitted to U.S.

Department of Justice, National Institute of Justice, Office of Justice Programs, Washington, D.C

Wiles, P. and Costello, A. (2000). *The 'road to nowhere': the evidence for travelling criminals*. Home Office Research Study 207. Research, Development and Statistics Directorate. London: Home Office.

Figures

Figure 1 – Distributions of individual offender crime trips (with a lower bound of zero) estimated by kernel density.

Source: LeBeau (1992)

Figure 2 – Aggregate distribution of crime trips (with a lower bound of zero) estimated by kernel density estimation.

Figure 3 – Distributions of offender crime trips (with a lower bound of zero) estimated by kernel density estimation conditioned by age groupings

Figure 4 – Boxplots of individual offender journey to crime distribution. Plots have been ordered by median value within each age group.

Figure 5 – Z score for skewness for each prolific offender. Observations have been sorted according to Z score magnitude.