Understanding the Deterrent Effect of Police Patrol

Oliver Kenneth Hutt

A dissertation submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Security and Crime Science
University College London

October 14, 2019
I, Oliver Kenneth Hutt, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.
Abstract

The fact that crime clusters spatially has been known since at least the early 19th century. However, understanding of the extent and nature of this clustering at different areal units, and the fact that crime also clusters at different temporal scales is relatively new. Where previously the most at-risk areas (or 'hot-spots') of crime were defined over areas the size of city districts and for periods of months if not years, the last decade has seen the focus shift to micro-places - areas of only a few hundred metres across - which are only ‘hot’ for days or even hours.

The notion that visible police presence in crime hot-spots can deter crime is not new and has been the basis of police patrols for two centuries. This deterrent effect has been well evidenced in many previous studies, both by academics and police practitioners. However, evaluations of these more recent micro-level hot-spot patrol strategies face significant analytic challenges and data quality concerns. They also often assume levels of police activity at the micro-area level (an ‘intention-to-treat’ design) rather than measuring it directly. The aim of this thesis is to investigate the accuracy and precision of data that can be used to evaluate micro-level hot-spot patrol strategies and the implications this has for any analysis conducted using such data at these micro-level geographies.

This thesis begins by outlining the relevant literature regarding place-based policing strategies and the current understanding of how crime clusters in both space and time. It continues by highlighting the data challenges associated with evaluating micro-level police interventions through the use of an illustrative analytic strategy before using a self-exciting point process model to evaluate the effects of police foot patrol in micro-level hot-spot under the assumption that the crime and patrol
data being used are accurate. This is followed by two chapters which investigate
the quality of the two datasets. Finally, the point-process evaluation is re-conducted
using simulated data that takes account of the uncertainty of the datasets to demon-
strate how data quality issues effect the result of such an evaluation and ultimately,
the perceived efficacy of these highly-focussed policing strategies.
Impact statement

The research presented in this thesis highlights a growing data quality issue with regard to the specification of police patrol locations and evaluations of the effect of police patrol implementations. The findings which are presented have implications for how evaluations of police patrol are conducted in future and how greater consideration must be given to the accuracy of the data. A key finding of the research is that the results of evaluation studies can differ when the underlying inaccuracy of the data is taken into account.

Suggestions are made for how data error can be better accounted for in future evaluations of police patrol. This is relevant to both academics and police practitioners as these data quality concerns are likely to impact upon their findings and the robustness with which evaluations of police patrol can claim to work - or not.

The analysis conducted in Chapter 6 is particularly relevant to police forces which use GPS data as it provides guidance on how accurately police patrols can be tracked using GPS data collected at different rates. As the cost of collecting data is dependent on the amount and rate of collection this has direct financial implications for police forces.

This research also highlights the importance of conducting analyses at appropriate spatial and temporal scales, whilst discussing and exemplifying the difficulties in doing so.

Finally, the fact that data quality can have a significant impact on the evaluation outcomes provides practitioners with useful information regarding whether two competing systems differ in the results they produce. For instance, when evaluating whether one system performs ‘better’ than another at predicting crime or in helping
officers prevent crime, data quality issues - as highlighted in this thesis - may lead to any evaluated differences between systems being insignificant, which in turn has an affect on procurement of these systems for real-world policing.
Acknowledgements

I would first like to thank Kate Bowers and Shane Johnson for their incredible support during this course of study. Their guidance, insight, and encouragement throughout have been immeasurable and I am greatly indebted to them for all that they have so kindly done for me. I only hope that what follows does not reflect poorly on them.

My thanks also to Tim Pescud, Jim Corbett, James Quirk, and Andre Ramsey from the Metropolitan Police Service who have all expended considerable effort to aid me in gaining access to much of the data which is used in this research. Similarly, I am thankful to Trevor Adams and Christine Leist for not only providing the data but also taking the time to ensure I understood the data - warts and all.

Thank you to West Yorkshire Police for providing me with the opportunity to accompany several of their officers on patrol. They provided much needed practical insight and directly informed the direction of the research.

This work is supported by the Engineering and Physical Sciences Research Council (EPSRC), grant number EP/M506448/1, for which I am incredibly grateful. Regardless of the quality of the work that follows, it would surely be far worse without the financial security this has afforded me and the opportunities it has opened up for me to engage with both academics and police practitioners, to say nothing of mitigating much of the potential stress that I would have endured without this support.
# Contents

1 Introduction .............................................................. 17
   1.1 Structure of the thesis ............................................. 19

2 A review of relevant literature ...................................... 23
   2.1 Crime and place ................................................... 23
       2.1.1 Repeat and near-repeat victimisation .................... 27
   2.2 Police patrol ..................................................... 31
       2.2.1 Introduction .................................................. 31
       2.2.2 Deterrence theory ......................................... 31
       2.2.3 The effects of police patrol ............................... 34
       2.2.4 Measuring dosage ........................................... 37
       2.2.5 The resolution of analysis ................................. 39

3 The effect of police patrol in crime hot-spots .................. 41
   3.1 Introduction ..................................................... 41
   3.2 Background ..................................................... 42
       3.2.1 Police recorded crime data ............................... 42
       3.2.2 Geocoding quality ......................................... 44
       3.2.3 Positional accuracy ....................................... 45
       3.2.4 Completeness ............................................... 49
       3.2.5 Temporal fidelity .......................................... 50
       3.2.6 Police movement data ..................................... 52
       3.2.7 Methods of evaluation .................................... 55
   3.3 Data .............................................................. 57
       3.3.1 Intended Patrol Locations ................................. 57
       3.3.2 Officer Location Data ..................................... 57
       3.3.3 Police Recorded Crime .................................... 58
       3.3.4 Measuring crimes within prospective boxes .......... 59
   3.4 Illustrative analytic strategy .................................. 60
   3.5 Results ........................................................... 62
   3.6 Conclusion ....................................................... 65

4 Modelling the deterrent effect of police patrol ............... 69
   4.1 Introduction ..................................................... 69
   4.2 Data ............................................................. 72
       4.2.1 Recorded crime data ....................................... 73
List of Figures

2.1 Evidence of the spatial clustering of crime .......................... 24
3.1 An illustration of the modifiable areal unit problem. ............... 51
3.2 The urban canyon effect ............................................. 55
3.3 Uncertainty of the timing of specific crime types ...................... 59
3.4 Distribution of dosage duration per (eight hour) shift for prospective boxes that received non-zero dosage ......................... 61
3.5 Cumulative crime count in live and control boxes .................. 63
3.6 Cumulative crime counts with Monte Carlo simulated significance bands ................................................................. 64

4.1 Temporal distribution of crime in Islington (in minutes) ............ 75
4.2 Spatial distribution of crimes in Islington ............................ 76
4.3 Study area with observation window and grid ....................... 80
4.4 Police patrol dosage within study cells ............................. 81
4.5 Total patrol dosage over the whole study period ................... 82
4.6 Endemic-only model of burglary risk ................................ 86
4.7 A simple endemic model incorporating seasonality ................ 88
4.8 Daily seasonality with one (above) and four (below) sine wave terms 89
4.9 An endemic model of burglary incorporating 4 daily and 4 weekly seasonal terms ...................................................... 90
4.10 Temporal interaction functions for endemic-epidemic burglary models. ............................................................. 97
4.11 Gaussian, power-law, and lagged power-law SIAF plots .......... 99
4.12 Gaussian SIAF parameters for different untying distances .......... 100

6.1 Distances between officer and author matched GPS pings .......... 121
6.2 Officer GPS-based paths .............................................. 123
6.3 Interpolated officer path and the true path of a patrol .............. 124
6.4 Officer, observer, and the true path of a patrol .................... 125
6.5 GPS ping distribution during trial period ............................ 126
6.6 Example trajectories for Fréchet distance measurement .......... 129
6.7 Distributions of Fréchet distances for example trajectories ...... 129
6.8 Similarity distributions of police patrol paths using sparse GPS data 131

7.1 Study area with associated observation window and grid .......... 135
7.2 An example of a re-calculated officer path .......................... 136
7.3 The distribution of p-values for simulated model dosage effects . . 140
List of Tables

4.1 Recorded crime attributes ........................................ 73
4.2 Crime geocoding accuracy classifications ....................... 74
4.3 MOPAC7 crimes recorded during the study period .............. 75
4.4 Endemic-only model of burglary risk - Categorical parameters 87
4.5 AICs for endemic-only models incorporating sinusoidal waves 88
4.6 Rate ratios for an endemic model composed of 4 day and 4 week frequencies of seasonality .......................... 91
4.7 Fit of models incorporating patrol in endemic component ....... 92
4.8 Rate ratios for an endemic model incorporating dosage ........ 92
4.9 Parameter estimates for endemic-only models with and without tied data ................................................. 94
4.10 AICs for endemic-only models incorporating sinusoidal waves - for untied data ........................................... 94
4.11 Parameter estimates for endemic-epidemic models .......... 96
4.12 Incorporating dosage into both endemic and epidemic components 100
4.13 Comparison of AIC scores across models .................... 101
5.1 The proportion of crimes geocoded to each accuracy category 107
5.2 MOPAC7 geocoded classifications ................................ 112
5.3 Average MOPAC7 geocoding accuracy (in metres) ............ 113
6.1 Distribution of Fréchet distances by ping frequency ........... 131
7.1 MOPAC7 crime counts - initial evaluation and re-evaluation 135
7.2 Model fits for baseline re-evaluation models .................. 138
7.3 Baseline model estimates ........................................ 138
7.4 Simulated dosage model estimates .............................. 139
Chapter 1

Introduction

This thesis investigates the effect that police foot patrol has on the prevention of crime. These effects have been the subject of many previous studies, both by academics and police practitioners, who have evaluated a range of outcomes including effects on actual crime rates, the public’s perceptions of crime, and perceptions of police performance. Early evaluations of police patrol in the 1970s showed mixed results (see Chaiken, 1976; Kelling, Pate, Dieckman, & Brown, 1974; Press, 1971), and the theory that visible police presence caused a deterrent effect amongst would-be offenders - the founding principle on which police patrols were implemented - was cast into doubt. However, as the body of evidence grew it came to show that focussing police resources in the most at-risk areas or ‘hot-spots’ led to significant reductions in crime (Braga, Turchan, Papachristos, & Hureau, 2019) with minimal displacement to the surrounding areas.

Recent developments in our understanding of where and when crime is most likely to occur have led to police patrol strategies that are more highly focussed in both space and time. That crime clusters in space is not a new revelation; since the early 19th century, there has been consistent evidence that crime does not occur randomly or uniformly in space, but in fact clusters at specific locations (for a recent review see Johnson, 2010). More recent analysis has shown that it also clusters in time (Johnson, Birks, McLaughlin, Bowers, & Pease, 2008). In fact, the ubiquity of this phenomenon at all spatial resolutions has been coining as the ‘law of crime concentration’ which states that:
...for a defined measure of crime at a specific microgeographic unit, the concentration of crime will fall within a narrow bandwidth of percentages for a defined cumulative proportion of crime. (Weisburd, 2015)

Microgeographic units are small geographic areas such as individual buildings, street corners, or small street segments (Eck & Weisburd, 1995) and these small areas of high crime are often referred to as ‘micro-hot-spots’.

These insights have led to the targeting of police patrols at a micro-level; down to hot-spots only a few hundred metres across and ‘hot’ for less than a day (see Mohler et al., 2015, for an example) in contrast to more traditional hot-spots which have been defined based on persistent crime over a prolonged period - usually at least one year (see Ariel, Weinborn, & Sherman, 2016; Fielding & Jones, 2012; Novak, Fox, Carr, & Spade, 2016; Telep, Mitchell, & Weisburd, 2014). Although evaluations of policing strategies focussed on the newer, more precisely defined ‘micro-hot-spots’ have started to appear in the literature (Mohler et al., 2015) one key facet that could significantly influence this research has yet to be addressed: what is the quality of the crime and patrol data behind the planning, execution, and evaluation of these exercises? The answer to this question has repercussions in terms of what areas are defined as hot-spots, when they should be targeted for patrols, and measurement of the effectiveness of those patrols.

Previous research into the accuracy of recorded crime data within the UK has highlighted some significant errors in the recording process. For instance, Harrell (2014) found that robberies in an area of Manchester (UK) were inaccurately coded by on average 190 metres. When the area of analysis is large, this presents less of an issue; however, implementations such as Mohler et al. (2015) use grid-based hot-spots just 150 metres across, at which point such errors could have significant effects on defining an area as a hot-spot and on evaluating the effect of policing such hot-spots.

Another issue that hinders robust evaluation of hot-spot policing is the lack of knowledge of exactly where and when police officers patrol. If patrols are more frequently going to be directed to micro-hot-spots then it follows that the measure-
1.1. Structure of the thesis

This thesis examines the quality of police recorded crime data and the measurement of police patrol to identify the impact of data quality on hot-spot policing strategies. Chapter 2 begins by examining where and when crime occurs. This is followed by a review of the hot-spot policing literature, beginning with an overview of the deterrence objective of police patrols and highlighting some of the challenges encountered to date in trying to measure that effect. Chapter 3 describes an evaluation of a hot-spot policing initiative that was implemented by the Metropolitan Police Service (MPS) in Southwark, an area within London, UK. The chapter highlights the potential data quality issues raised by data quality standards in the recording and measurement of crime data and police officer movements. This is followed in Chapter 4 by an attempt at estimating the effects of a similar police foot patrol strategy implemented in the London borough of Islington, UK. At the time, the MPS were evaluating the benefits of several micro-hot-spot prediction systems across the city and Islington was expending considerable effort to ensure a patrol strategy was properly implemented using the system at their disposal. This is fol-
lowed in Chapter 5 by an investigation into the accuracy of police recorded crime data; specifically the recording of where and when the offences occurred. Whilst police patrol strategies are sometimes measured against other outcomes - such as fear of crime amongst the local population or their confidence in the police - there are reasons for focusing on recorded crime:

- The primary motivation for conducting police patrols has historically been to deter potential offenders. Whilst other beneficial outcomes have been sought or discovered, crime deterrence remains the primary goal for many police forces, including the MPS.

- Evaluations of hot-spot policing initiatives are unfortunately often only considered once the strategy is in place. As such, recorded crime are often the only suitable data for conducting an evaluation.

- As far as the author is aware, all the systems currently in use to predict micro-hot-spot locations (and thus to determine police deployment locations) are predicated first on recorded crime data. Several do incorporate other data sources, such as land use, population demographics, and weather forecasts, but recorded crime data remain the kernel of these models.

Chapter 6 is focussed on the measurement of foot-based police patrols. Whilst it is now generally accepted that hot-spot policing can reduce crime, the question remains as to how much time officers should spend in a hot-spot to maximise the deterrent effect they generate - this is often referred to as the ‘dosage’ of police presence provided in the hot-spot. The chapter uses data collected in West Yorkshire and London to investigate data challenges associated with accurately measuring where officers walked on patrol and how data collection parameters (specifically, the frequency with which their location is recorded through a GPS-enabled device) effect the interpolated path, and thus the estimated dosage delivered in any given area.

The aim of the analyses conducted in Chapters 5 and 6 was ultimately to build a better understanding of how the quality of both crime and patrol data influence the
1.1. Structure of the thesis

estimated deterrent effect of police patrol on crime prevention. Using the findings from these chapters, the evaluation conducted in Chapter 4 is revisited in Chapter 7 and the estimate of the effect of police patrol on preventing crime is recomputed, with the error and uncertainty in the underlying data taken into account. The purpose of this was to provide a more nuanced understanding of how much confidence can be placed in any estimate of the effect of police patrol on deterring criminal behaviour.
Chapter 2

A review of relevant literature

This chapter provides an introduction to the theories and previous research which
are relevant to the subject of this thesis. It begins with an overview of the spatial
and temporal patterns of crime and emphasises the importance of *places* within the
context of studying crime. Theoretical frameworks for why these patterns exist are
then outlined along with a review of research which has tested them. Section 2.1
ends with a discussion of how these theories have been used in a practical sense to
try and predict where and when crime is most likely to occur. Section 2.2 looks
at how police forces have operationalised these theories and frameworks in their
efforts to prevent crime from happening.

2.1 Crime and place

Crime is not a randomly occurring phenomenon. It has been known since at least
the 19th century that crime clusters in space. In France, Balbi and Guerry (1829)
compiled the first comprehensive national data set of ‘moral statistics’ which in-
cluded crime data across all 86 departments of France. Figure 2.1 is a copy of their
original visualisation and this clearly illustrates that, at least at the resolution in-
vestigated, crime clusters spatially. In the figure, the top two maps show the crime
rates per department and darker areas represent higher rates of person (top left) and
property (top right) crime. Whilst not relevant to the current discussion, the map at
the bottom of the figure shows the distribution of male children in education with
darker areas having more male children in formal education.
Figure 2.1: Evidence of the spatial clustering of crime
A comparison of person (top left) and property (top right) crime rates as well as male children in school (bottom) by departments (i.e. regions) in France. (from Balbi & Guerry, 1829. (Source: Bibliothèque nationale de France))

Since then, multiple studies have shown that crime clusters in space (Guerry, 1833; Quetelet, 1842, pp. 87; Johnson, 2010; Shaw & McKay, 1942). So much so that in 2015, Weisburd coined the law of crime concentration to highlight the consistency with which crime has been found to concentrate in space (Weisburd, 2015). One perspective on why this occurs stems from the field of ecology. The Routine Activity approach proposed by Cohen and Felson (1979) emphasises the situational aspects which are necessary for crime to occur, rather than focussing on characteristics of the offender. Routine Activity postulates that crimes occur when three factors converge: a motivated offender encounters a suitable target in a location without capable guardianship. This is sometimes referred to as the Crime Triangle. A suitable target is usually an individual or piece of property. Guardianship can take many forms but is ultimately a mechanism which in some way protects a suitable
2.1. Crime and place

Crime pattern theory builds upon the Routine Activity perspective and describes how individuals develop an ‘awareness space’ - an understanding of the environment around them which they build up as they go about their daily lives. Awareness spaces are cognitive representations of peoples’ activity spaces and are built up of nodes (individual locations such as home and work), paths (the routes taken between nodes), and edges (the boundary between these familiar points and those which are less well known) (P. L. Brantingham & Brantingham, 1993). Whilst everyone will develop awareness spaces, it is the awareness spaces of criminals that are the primary (but not sole) concern of crime pattern theory. The Brantinghams found that, “Criminal targets are usually picked from within this awareness space. Exploration of the unknown is not part of the target search process for most individuals.”

The home is a major activity node for most people. Snook (2004) analysed the distances between burglars’ homes and known targets in the area of the city of St John’s, Newfoundland, Canada. Their sample consisted of 41 series of burglaries committed by 51 burglars through 1998 and 1999 totalling 347 burglaries. They found that the median distance travelled was 1.7 kilometres, with 33% of burglaries being conducted within 1 kilometre of a burglars home and 84% of burglaries being made within 5 kilometres of the burglars home. Furthermore, they found that older burglars travelled significantly farther than younger burglars for their first offence and that on average a burglar over the age of 20 would travel 1.2km farther to commit their first burglary than a burglar 20 years old or younger. This is perhaps due to the larger awareness space of older individuals who have had more opportunity to explore their environment and expand their cognitive map.

The findings of P. L. Brantingham and Brantingham (1993) are also supported by the work of Rengert and Wasilchick (1985) who interviewed burglars about the crimes they had committed, including how they selected their targets. Their research showed that 75% of burglaries occurred within a 45 degree arc between an
offender’s home and their place of work.

As discussed by P. L. Brantingham and Brantingham (1995), the aggregation of peoples’ activity spaces can lead to particular areas having specific profiles of (routine) activity, which can have implications for crime. They identified three types of area whose character can be thought of as being defined by one of the three factors of Routine Activity.

• **Crime generator:** A location with a high number of potential targets, such as shopping centres or transit hubs. These are usually areas where a large number of individuals need to go to or pass through to conduct their day to day activities, not for any particular criminal intent. This creates a high level of opportunity for a potential offender and whilst they might not have travelled to such a location with the intention to commit a crime, they can exploit the large target pool. These are often characterised by a high volume of crime but a low rate of crime due to the high number of potential targets.

• **Crime attractor:** Characterised as a place where offenders go not because it forms part of their non-criminal activity, but with the intention of committing a crime because of well-known opportunities. This might include areas known for prostitution or drug markets where offenders are well motivated due to the expected ‘rewards’ they can gain. Such areas have both a high volume and a high rate of crime due to the ambient population typically being low in such areas.

• **Crime enabler:** A location with poor guardianship such an insecure car park or a city park. These will often have a low volume of crime but a high rate of crime due to the sparse target availability.

A growing body of evidence provides support for such theories (Eck & Weisburd, 2015; Sherman, Schmidt, & Velke, 1992). For a detailed discussion, the reader is referred to Bruinsma and Johnson (2018) which provides detailed reviews of the theoretical and empirical literature. However, the point of central importance here is that crime clusters in space, which has clear implications for the prevention
of crime. More recently, it has also been shown that crime clusters in space and time and this is discussed in the next section.

2.1.1 Repeat and near-repeat victimisation

Recent efforts have also incorporated temporal factors in the study and identification of crime hot-spots. This has been motivated by the empirical finding that locations at which a crime has occurred are likely to be victimised again soon afterwards (Haberman & Ratcliffe, 2012; Polvi, Looman, Humphries, & Pease, 1991). These repeat victimisations of the same target are typically the work of the same offender (Everson & Pease, 2001), and research also demonstrates that as well as returning to the same location, offenders frequently also select targets in the near vicinity of the previous offence - a process known as near-repeat victimisation (Bernasco, 2008; Johnson, Bowers, Birks, & Pease, 2009). Prior victimisation has also been shown to be a good predictor of future victimisation (Farrell, 2013; Pease, 1998) although research demonstrates that this predictive value is short lived. For example, Polvi et al. (1991) and Haberman and Ratcliffe (2012) found that repeats happened very soon after the initiating event; usually no more than a few days afterwards. Similarly, Ashton, Brown, Senior, and Pease (1998) found that offenders would often return up to a few weeks later. Two possible reasons for why repeat and near-repeat victimisations occur were proposed by Tseloni and Pease (2003):

- **The boost explanation**: An offender gains a better awareness of the target location and the potential rewards available once they have committed the initiating offence. With this additional knowledge, the attractiveness of the target (or similar targets nearby) is increased. For example, once a burglar has victimised one property in a street they may assume (often correctly) that other properties on the street will have similar security features, the same internal layout which would aid the speed of their offence, and perhaps items of similar value. In this case, the risk of victimisation will be directly affected (boosted) by a previous offence.

- **The flag explanation**: There are inherent characteristics to a target location
which make it more attractive to offenders than other potential target locations. For example, a clear lack of capable guardianship may make an area more appealing to an offender. In this case, the risk of victimisation is time-stable and unaffected by patterns of crime.

Both explanations have been tested and have found support. Johnson (2008) found that a micro-simulation model that allowed for both boost and flag mechanisms provided the best explanation for near repeat victimisation patterns when simulating patterns of residential burglary data in Merseyside, UK. Neither mechanism (boost or flag) alone generated realistic patterns. In an analysis of crimes detected by the police, Everson (2003) found that burglars would often target nearby homes after successfully committing a burglary. The near-repeat phenomena can also be thought of as a form of contagion that can be modelled using techniques that have their origins in the field of epidemiology (Knox, 1964). Using such techniques, the near-repeat phenomenon has been found in a number of studies conducted in a range of countries and for a range of crime types. For example, Townsley, Homel, and Chaseling (2003) demonstrated that a near-repeat phenomenon was evident amongst residential burglaries in Brisbane, Australia. Ratcliffe and Rengert (2008) found the phenomenon for shooting offences in Philadelphia, USA. Johnson et al. (2008) analyse residential burglary data for two cities in each of five countries and find consistent patterns of near repeats in all of them.

A conceptual framework - which draws on the ideas discussed above - that more explicitly articulates a boost mechanism is the ‘Optimal Forager’ hypothesis (Johnson, 2014; Johnson & Bowers, 2004; Johnson et al., 2009). According to this explanation, burglars are in many respects similar to foraging animals. They seek to maximise benefits whilst simultaneously minimising risk and energy expended. As such, it is suggested that upon finding an area of high opportunity they will continue to exploit it in the near future. This continues until the perceived benefits of remaining or returning to the area are outweighed by the increasing risks. They must also consider the likely amount of effort that needs to be expended to find a more beneficial area. An analysis of police recorded crime data by Johnson, Summers,
and Pease (2006) found that burglaries which occurred close to each other in both space and time were significantly more likely to be the work of the same offender than burglaries which occurred close in space but not time or vice versa. They found the same pattern when analysing thefts from motor vehicles, supporting the idea that offenders return to locations they have recently victimised in order to offend again.

Given the regularity of these patterns of repeat and near-repeat victimisations Bowers, Johnson, and Pease (2004) developed a simple mathematical approach to create crime maps, (which they called ProMap) that incorporated both the spatial and temporal attributes of offences. Evaluations of such ‘prospective’ methods have found that they more accurately predict where crime will happen than traditional, ‘retrospective’ techniques (Johnson et al., 2008; Johnson et al., 2009).

ProMap is not the only predictive model that has been developed to predict where crime will occur. PredPol is a commercial system that has been used in the city of Los Angeles, USA and the county of Kent, UK. It is a prospective crime mapping system, based on the same principles as ProMap, but one that uses a self-exciting point process model to predict crime. This type of model will be discussed in some detail in Chapter 4, but to briefly explain, they model a sequence of events (i.e. points) where the rate of occurrence of events can depend on past events. An event is said to ‘excite’ the probability of another event happening, hence ‘self-exciting’. These models have previously been used to model earthquakes and their after-shocks (Ogata, 1999) as well as the spread of disease (Meyer, Elias, & Höhle, 2012). Predpol represents an implementation of such a model to model crime. Developed by Mohler, Short, Brantingham, Schoenberg, and Tita (2011) an evaluation of the system indicated that it was 1.4-2.2 times more accurate than predictions produced by dedicated crime analysts (Mohler et al., 2015). Both ProMap and PredPol are typically used to generate very near-term predictions; usually for the next few days. Risk Terrain Modelling (RTM) is another prospective crime mapping technique but rather than using the distribution of crime as input it utilises data on aspects of the urban environment that are thought to act as crime generators and attractors (such as bars, ATMs, and shops - recall Section 2.1) to generate individ-
ual, weighted risk layers that are then incorporated into a single risk surface using an additive model (Caplan, Kennedy, & Miller, 2011). Predictions from RTM are usually for much longer time periods. For instance, they may be generated for the following three to six months. One similarity of ProMap, PredPol, and RTM is that the output for each system is a grid where each cell is given a risk score.

In contrast, Rosser, Davies, Bowers, Johnson, and Cheng (2016) have developed a prospective mapping technique based on ProMap that uses the road network rather than areal units. The results of their study show that their road network model is significantly more predictive of residential burglary than the analogous grid-based system, having predicted 20% more crime based on 5% coverage of the potential at-risk area. An important point not yet discussed (see Chapter 3 for a detailed analysis) is that the exact time at which a burglary occurs is often unknown, since the victim will often be away from home (e.g. on holiday or at work) at the time of the offence. For this reason, police recorded crime data include fields for the earliest and latest possible dates and times at which offences may have been committed. Of course, the recorded timing of offences can affect predictions generated and estimates of their accuracy. For example, if a crime occurred on one day (e.g. the day before a prediction to be generated) but was recorded as having happened on another day (e.g. the day after the prediction) then the prediction will underestimate the risk in that area for the day the prediction is produced but potentially over-estimate the risk in that area for the days which follow (as the likelihood of a repeat victimisation is higher closer to the initiative event). In their study, Rosser et al. found that using either the earliest or the latest possible time in their analyses did not significantly change their results. However, this is an important issue, and one that is discussed further in Chapters 3 and 5.

Since the police are unable to be everywhere at all times, rational resource allocation would seem to dictate that the police should deploy resources to those areas where they can have the greatest effect and at the times they are mostly to do so. What effect they might try and generate is covered in Section 2.2 below.
2.2 Police patrol

2.2.1 Introduction

It is not sufficient to ask simply if an initiative (in this case police patrol) ‘works’, but rather it is necessary to ask, ‘What works for whom in what circumstances and in what respects, and how?’ (Pawson & Tilley, 2004). In this research the ‘what’ refers to high visibility police foot patrols which are intended to reduce street crime in an area by deterring potential offenders from committing an offence. The mechanism by which this should be theoretically achievable (the ‘how’) is through the generation of a deterrent effect and is described in the next section. This is then followed by a review of historic evaluations that move towards a better understanding of the circumstances in which reductions in crime have been found to occur. That police patrols can reduce crime is now generally accepted (Braga et al., 2019). However, this consensus has only been reached over time and by elucidating ‘for whom in what circumstances’ and by a better understanding of the challenges of measuring both the effects of police patrols and police patrols themselves.

2.2.2 Deterrence theory

The notion that potential offenders could be deterred from committing a crime by increasing the perceived likelihood of being caught originates in the work of Beccaria (1764/1872) who posited that the certainty, not severity, of punishment held greater sway on the likelihood that someone might try to commit a crime.

*The certainty of a small punishment will make a stronger impression, than the fear of one more severe, if attended with the hopes of escaping; for it is the nature of mankind to be terrified at the approach of the smallest inevitable evil, whilst hope, the best gift of Heaven, hath the power of dispelling the apprehension of a greater; especially if supported by examples of impunity, which weakness or avarice too frequently afford.* (Beccaria, 1764/1872, p.94)

Beccaria was primarily interested in developing a more effective legal system and provided no rationale for why individuals might commit crime beyond basic
self-interest. Bentham (1789) on the other hand provided a more developed theory of general behaviour based in the utility of an individual’s action. This ‘utility theory’ posited that all human decisions are based primarily in weighing potential gains against likely costs and then choosing the action they believe will benefit them the most.

Several authors argued that utility theory, taken as a model of criminal decision making, could inform crime control policies and future research (Becker, 1968; Clarke & Cornish, 1985). This led Cornish and Clarke (1987) to propose the Rational Choice Perspective (RCP) which suggested that offenders make largely rational decisions regarding whether to commit a crime by weighing up the perceived potential benefits against the perceived effort required to commit it and the potential to be caught and punished. The study of journeys to burglary crimes by Snook (2004) which was discussed earlier also looked at the value of property that was stolen. They found that burglars travelled a median distance of 1.0km, 1.6km, and 2.1km for property valued at under $500, $500-$1000, and over $1000 respectfully. However, an important aspect of this perspective is that offenders will not possess perfect information. Moreover, the weight given to each element of the calculation will be biased by previous experience, perceptions, and the fact that decisions are often made quickly. In this regard decisions are said to be bounded; whilst they are based in rational thought they are not perfectly rational. This also means that the decision-making process can be influenced to suggest to a would-be offender that the perceived risks of committing a crime outweigh the benefits (even if they do not).

However briefly offenders might consider them, three deterrence factors have been found to influence their decision making; the certainty, severity, and celerity (swiftness) of punishment (Paternoster, 2010). Of the three, certainty of apprehension - as Beccaria posited - has been found to have the greatest effect (see Durlauf & Nagin, 2011; Nagin, 2013). This is good news for police officers who have a much more limited influence on the severity or celerity of punishment, whilst their visible presence can increase the perceived risk of detection and apprehension, and
thus deter the would-be offender (Riccio, 1974).

Distinctions have been drawn between different types of deterrence. General deterrence is aimed at the wider population, with the aim of dissuading the public from a general desire to commit crime. Specific deterrence refers to the threat of further sanctions for someone who has already been punished for a crime. Marginal deterrence refers to an incremental increase in deterrence provided by an additional or stronger condition. For instance, whether harsher sentencing for a crime elicits a greater deterrent effect (Zimring, 1971). Further to this, Riccio (1974) defines direct deterrence as a mechanism designed to directly block the opportunity for an offender to commit a crime. With these distinctions in mind, police patrols can be argued to operate across several categories of deterrence; they remind the general population that they are expected to obey the rule of law, they provide a visible threat to an offender who might seek to commit a crime imminently, and they may impede, in a very literal sense, a crime from being committed.

Sherman (1990) defined a police crackdown as a sudden increase in police presence which could produce a deterrent effect. He highlighted three important distinctions in the effects a crackdown might have; the initial deterrent effect caused by the sudden surge in police presence, a residual effect after a crackdown where offenders are unsure if the risk of being caught is still higher than previous, and the decay of the initial deterrent effect as offenders become more confident about the risk of offending.

Sherman highlighted how ‘crackdown-back off’ (that is, a crackdown followed by the removal of the increased presence) strategies can increase an offender’s perceived risk of being caught due to the uncertainty of capture this generates. As an example, assume there are two areas, $A_1$ and $A_2$ which are each patrolled constantly and have a 5% probability of capture for any offender in each area. If the patrol resource is combined and alternated between $A_1$ and $A_2$ so that at any given time one of the areas has twice as much ‘patrol coverage’ then at any given time the risk of capture is 10% in one area and 0% in the other. According to Sherman, the offender, not knowing which area has the higher risk, is likely to over-estimate the
overall risk of offending in either area. As Sherman (1990) points out:

*It takes the same number of police to create a continual apprehension risk of 5 percent as it does to vary that risk between 0 and 10 percent. It may be more cost-effective to choose the latter option, since that could keep the average perceived risk of apprehension twice as high as the average resources allocated would justify.*

How long officers should remain in a given area, (to create the initial deterrence) and how long any residual deterrence exists for have proved difficult to quantify. Whilst studies (discussed below) have examined whether patrolling has a deterrent effect, research designed to identify the optimal dosage and frequency of patrol is currently very limited.

### 2.2.3 The effects of police patrol

Visible police patrols, with the objective of preventing crime, have been a cornerstone of police activity since at least the early 19th Century (see: HMSO, 1862). However, despite the strong theoretical basis for police patrols discussed above, empirical evaluations were not conducted until the 1970s and produced mixed results. One of the first evaluations, and certainly one of the most influential (Sherman & Weisburd, 1995) was conducted in Kansas City. The Kansas City Preventative Patrol Experiment (Kelling et al., 1974) was a year-long study from October 1972 to September 1973, covering a 32-square mile area of the city. 15 vehicle patrol beats were classified as either ‘reactive’ (no preventative patrol was conducted; officers only entered the beat in response to a call for service); ‘control’ (the usual level of preventative patrol was maintained at one car per beat); or ‘proactive’ (preventative patrolling was increased by two or three times). The experiment sought to measure the impact of vehicle-based patrolling on both crime incidence and public perceptions of crime (primarily, their fear of being victimised). The results showed almost no significant difference in metrics across the three groups. Where statistically significant differences were observed, they showed no clear pattern; differences benefited the three groups fairly evenly.
The Kansas City Experiment significantly diminished the perceived efficacy of preventative patrol (Sherman & Weisburd, 1995) despite other contemporaneous studies that contradicted it (such as: Chaiken, 1976; Press, 1971) and several issues with its design and implementation. Larson (1975) performed a comprehensive review of the experiment, highlighting that the actual amount of police presence (or ‘dosage’) in the proactive areas may not have been substantially greater than the reactive areas. Further, the experiment was conducted over a large (32-square mile) area rather than being focussed on those areas of high risk. By employing this more general strategy, much of the dosage is likely to have been wasted on areas which had no real risk of crime; in effect diluting the impact of any increased patrolling. Whilst the issue of quantifying dosage remains a challenge (discussed further below) more progress has been made in identifying and targeting areas that are most at risk of crime.

Given that the police have limited resources with which to conduct preventative patrols it seems reasonable that they should target those areas with the greatest risk of crime; it would be neither practical nor desirable to have a police presence on every street. This way they can maximise the efficacy of any deterrent effect they provide by, “changing the cost-benefit analysis that the motivated offender engages in when deciding whether or not to commit the crime in question” (Cornish & Clarke, 2003). The fact that the deterrent effect is implemented at the locations where crime is likely to occur is important, as more distal deterrent effects (such as the severity of punishment if they are caught) carry less weight in the offender’s cost-benefit calculation (Paternoster, 2010). Focussing police presence on smaller areas at greater risk of crime is more likely to produce a deterrent effect where it is actually required and a more concentrated dose becomes possible due to the smaller geographic area, reducing the impact of any dosage dilution - two major issues in Kelling’s study. These small regions of high risk are generally referred to as crime hot-spots and, “can be as small as the area immediately next to an automatic teller machine or as large as a block face, a strip shopping center, or an apartment building.” (Eck & Weisburd, 1995). Hot-spot policing patrol strategies have generally
been found to reduce crime (Ratcliffe, Taniguchi, Groff, & Wood, 2011; Sherman & Weisburd, 1995; Skogan & Frydl, 2004) and reviews of the evidence have been constantly updated (see Braga, 2001, 2005; Braga et al., 2014; Braga et al., 2019). Braga et al.’s 2014 review identified 25 empirical tests (both published and unpublished) which met the inclusion criteria - that is, they focussed on measuring the effects of police interventions at crime hot-spots. 20 out of the 25 tests showed significant crime reductions and, overall, the results suggest that they led to a significant reduction in crime with an effect size of 0.184. By comparison, the latest update (Braga et al., 2019) (which covers studies up to February 2017) identified 78 tests of hot-spot police interventions; a substantial increase in studies conducted since the previous review. Of the 78 tests, 62 reported significant crime reductions and the overall effect size for all of the studies was 0.132 which, whilst small, was still significant.

Police forces have recently begun to use systems that target patrols on spatially small and temporally transient hot-spots. As discussed in Section 2.1.1, police forces in Kent (UK) and Los Angeles (USA) have used a preventative patrol tool called PredPol that designates patrol locations measuring approximately 150 metres across. A randomised control trial found that police patrols in PredPol-defined hot-spots were associated with significant (7.4%) reductions in crime. Similar methods, based on the same near-repeat principles discussed above, have been shown to be effective in Manchester (UK), where an evaluation of a prospective hot-spot policing initiative estimated that 338 domestic burglaries were prevented over a 12 month period (Fielding & Jones, 2012). A hot-spot policing initiative in Peterborough (UK) defined circular hot-spots with a radius of 150m (Ariel et al., 2016). They found that an increase in foot patrol by Police Community Support Officers (PCSOs) of approximately two extra ten minute visits per day in the treatment hot-spots led to 39% less crime and 20% fewer emergency calls-for-service compared to the control hot-spots. At such high spatial and temporal resolutions, the accurate measurement and recording of both crime data and police patrols becomes increasingly important.
2.2.4 Measuring dosage

Whilst it is now accepted that hot-spot policing can reduce crime, the question remains of how much time should officers spend in a hot-spot. Research focussed on the relationship between the quantity of dosage and the benefits produced remains sparse (see: Bowers, Johnson, & Hirschfield, 2004) though there are some notable studies, such as an analysis by Koper (1995) of a preventative patrol experiment conducted in Minneapolis (Sherman & Weisburd, 1995). The work was based on the recordings of trained observers who were positioned in 100 active hot-spots; they recorded the length of time officers spent in the hot-spot and the crime and disorderly behaviour in and around the hot-spot. Koper found that fewer than 10 minutes of police presence had no noticeable improvement in deterrence when compared a quick drive-by. 15 minutes of police presence did have an effect but more than 15 minutes had diminishing returns. However, he cautioned about interpreting the results as definitive as the effect was not statistically significant. None-the-less, this peak in efficiency at 15 minute (known as the ‘Koper Curve’) has become a ‘golden rule’ in preventative patrol (Perry et al., 2013), despite a lack of further study. Separate studies have tested whether 15-minute patrols are effective (Telep et al., 2014) and found this can lead to a significant reduction in crime, but the effect of different amounts of dosage have received little scrutiny.

One reason for this is the significant challenge associated with quantifying exactly how long officers spend in hot-spots. Until recently, this could only realistically be achieved by stationing observers either within a hot-spot or by asking police officers to record exactly when they entered and exited the area (e.g. Ratcliffe et al., 2011; Sherman & Weisburd, 1995), by analysing police logs (e.g. Telep et al., 2014), or assuming that treatment had occurred without actually measuring it; known as an ‘intention to treat’ evaluation model (Andresen & Hodgkinson, 2018; Eck & Maguire, 2005; Novak et al., 2016; Sherman et al., 1995). An example of the latter was conducted by Andresen and Hodgkinson (2018). This study was conducted in a small neighbourhood of North Vancouver, British Columbia and covered the period of January 2007 to December 2014. From 2010 onwards,
the Royal Canadian Mounted Police (RCMP) implemented a foot patrol initiative which ran from May to September in each year - their peak tourist season. The authors conducted a pre-post analysis using monthly crime counts for eight crime types. In their statistical analysis, they included a binary variable to account for the months when the foot patrol operation was in effect. Their findings were mixed, with 10 of the 20 patrol locations seeing significant reductions in at least one crime type. However, 3 locations experienced significant increases and 7 locations experienced no significant change. The increases in crime are potentially due to increased citizen confidence in the police and increased opportunity to report crime due to the increased presence of police officers. However, the fact that the patrols are assumed to take place in the correct locations, rather than being empirically measured, is an inherent weakness of such a study. The intention-to-treat approach is particularly problematic due to the potential for implementation failure (Knutsson & Clarke, 2006); that is, the intended dosage is not realised, or declines during the study period, as was the case in the studies by Sherman, Buerger, and Gartin (1989), Sherman and Weisburd (1995) and Telep et al. (2014).

In Larson (1975), one of the main critiques of the Kansas City Experiment was the issue with measuring the frequency and quantity of patrol dosage across the different areas, with Larson asserting that in reality the intended differences in patrol would not have been realised. These assertions were also challenged by the original authors of the Kansas report (Pate, Kelling, & Brown, 1975). Both sides put forward reasonable views but the fact remains that both rely on estimates on how much dosage was achieved; not real-world measurement. The quantification of dosage has been and remains a considerable challenge. Even when observers or police logs are used, there are often issues of precision; officers can stray outside the patrol zone without realising it (Sorg, Wood, Groff, & Ratcliffe, 2014) or ...through boredom or a perception that they were displacing crime to nearby streets would stray for a time if they were aware of areas of interest just beyond the foot patrol area... (Ratcliffe et al., 2011).

Since Koper’s 1995 study, few have looked at how the amount of police patrol
dosage impacts on crime. One of the difficulties with such research concerns measurement. As patrols are directed to increasingly small micro-areas, it follows that the measurement of patrols needs to be increasingly precise. Although the precise measurement of patrol dosage remains a challenge, there has been a steady growth in the usage of Global Positioning System (GPS) devices by police forces. These provide a new method for tracking where officers move and thus a new way of estimating police dosage. A more detailed discussion of the strengths and weaknesses of this approach is covered in Chapter 3 along with an overview of the few studies which have so far utilised these data.

### 2.2.5 The resolution of analysis

Weisburd and Lum (2005) found that, as of 1999, 59% of US police forces were already using computerised crime mapping software as an integral part of policing and this adoption has continued, with crime mapping now commonly used across US police forces, particularly to identify hot-spots (Koper, 2014). Weisburd & Lum argue that the increased confidence in the effectiveness of hot-spot policing is what led to the widespread adoption of computer-based crime mapping within police forces. Their study, based on surveys conducted amongst 125 police forces in the US, found that 43% of police forces that had adopted computerised crime mapping had initially done so to facilitate hot-spot policing and that in total, 80% of police forces that had adopted computerised crime mapping used it to facilitate hot-spot policing, highlighting the importance of hot-spot identification in regular policing.

Crime mapping has evolved considerably over the last half century, particularly in the move from wide-area patrol strategies exemplified by the Kansas City Preventative Patrol Experiment (Kelling et al., 1974) and the Newark Foot Patrol Experiment (Kelling, Pate, Ferrara, Utne, & Brown, 1981) in the 1970s, to more targeted hot-spot patrol strategies evaluated in, for example, Minneapolis (Sherman & Weisburd, 1995), Philadelphia (Ratcliffe et al., 2011), and Sacramento (Telep et al., 2014) in the US after 1990. Evaluations of micro-level hot-spot policing strategies in Los Angeles (US) and Kent (UK) (Mohler et al., 2015), and Peterborough (UK) (Ariel et al., 2016) have improved the spatial resolution of the analysis even
further. Several studies (e.g. Ariel et al., 2016; Fielding & Jones, 2012; Mohler et al., 2015; Williams & Coupe, 2017) have also explored the relationship between the amount of time that police officers spend patrolling such hot-spots – police dosage – and the volume of crime that occurs within them.

Such evaluations are welcome and necessary if hot-spot patrolling strategies aimed at deterring crime are to be improved. However, potential data error and uncertainty warrant greater consideration for such focussed interventions, as they have the potential to significantly alter the conclusions drawn, and the policy decisions that follow. Specifically, there are three factors of data precision that need to be addressed: spatial attributes of crimes, temporal attributes of crimes, and police patrol tracking. The precision of the crime data will have an impact on which areas will be designated as hot-spots, how well those hot-spots predict actual risk of crime, and evaluations of crime reduction strategies in hot-spots. To robustly evaluate the impact of police presence on crime, precise and accurate patrol data are necessary to understand where and when dosage was actually applied. Hot-spot evaluations do not typically include a detailed assessment of the uncertainty associated with the data on which they are based.

Chapter 3 provides an argument for why data uncertainty is a more pressing issue now than ever before. For the purpose of illustration, data from a study conducted in London (UK) are then presented and the effect of dosage on crime occurrences examined. Through this analysis, the quality of the underlying data is examined and the assumptions made outlined with substantial emphasis placed on their potential impacts on the results. This motivates the chapters which follow. Chapter 4 presents a more sophisticated analytical framework through which the effects of police patrol on crime are evaluated. Chapters 5 and 6 focus on the error and uncertainty within the underlying crime data and patrol data respectively. The evaluation from Chapter 4 is then revisited to illustrate how estimates of the effects of police patrol on crime are influenced by these data quality issues.
Chapter 3

The effect of police patrol in crime hot-spots

3.1 Introduction

The purpose of this chapter is to evaluate a micro-place based hot-spot policing implementation that was conducted in London (UK). Significant emphasis is placed on the potential issues raised by data quality standards in the recording and measurement of crime data and police officer movements. The analysis focusses on an area which used a predictive algorithm to designate micro-place patrol zones for each police shift over a 2-month period. Police officer movements are measured using GPS data from officer-worn radios. Descriptive statistics regarding the crime data commonly used to evaluate this type of implementation are presented and simple analyses presented to examine the effects of officer patrol duration (dosage) on crime in micro-place hot-spots. Some of the findings presented in this chapter have previously been published in Hutt, Bowers, Johnson, and Davies (2018). This chapter begins by reviewing the relevant police recorded crime literature and the challenges associated with measuring police officer movements.
3.2 Background

3.2.1 Police recorded crime data

There is a growing trend towards police patrols being targeted with increased spatial and temporal precision (Andresen & Hodgkinson, 2018; Mohler et al., 2015; Williams & Coupe, 2017). As the precision with which police patrols are targeted increases, spatial and temporal inaccuracies (and omissions) in crime data will have a greater impact on the veracity of analyses conducted. Ratcliffe (2004) clearly articulates that “crime is an inherently spatial phenomenon”. However, a complication of mapping crime locations is that crime events may not occur at a precise addressable location, but (say) on a street or in a park. Inaccuracies in recording where exactly a crime occurred will “...translate into compounding errors as the analytical and dissemination stages of police intelligence work are undertaken”. Police recorded crime data is a fundamental component of determining the location of micro-hot-spots and in evaluating police interventions in them. Hart and Zanbbergen (2012) have expressed concern about geocoding: the process of converting address data (such as building, street name, and postcode) into a set of spatial coordinates. Generally, the process involves taking an location to be geocoded (e.g. an address at which a crime occurred) and assigning to it the coordinates of the ‘best match’ among a dataset of all known addresses. Depending on the details of the location, this could be a specific building, the centre point of a road, or the centre point of an area (e.g. if only a postcode is known). Often, however, these outputs are treated as point locations in subsequent processing, regardless of how they were derived.

These issues add to the challenges of accurately geocoding a crime event, particularly if police officers need to provide a real address in the process of recording a crime. Previous research has already highlighted the problem of the spatial inaccuracy of robbery data which, for an area in Birmingham (UK). In this case, the average robbery was found to be correctly geocoded in only 31% of cases, with the average robbery incorrectly geocoded by 193 metres (Harrell, 2014). Similarly, Johnson et al. (2006) analysed theft of motor vehicle (TOMV) and theft from motor
3.2. Background

Vehicle (TFMV) crimes recorded by Derbyshire (9261 incidents) and Dorset police (5747 incidents). They found that, for the two areas, TOMV records did not have complete address data in 31% and 45% of cases respectively. In the same study, TFMV crimes did not have complete address data in 30% and 59% of cases respectively. Such inaccuracy is likely to have a substantial impact on the identification of micro-hot-spots and evaluations of patrol efficacy in these areas. As these examples illustrate, the level of inaccuracy can vary greatly between areas and crime types.

The temporal uncertainty of some crimes present a similar challenge to analysis, albeit with a very different cause. Whilst spatial inaccuracy may be reduced through better recording and geocoding, these improvements are possible because the actual crime location is known (at least by someone). Temporal inaccuracies are usually due to neither the victim, nor the police, ever being able to identify when exactly the crime occurred and so this requires a different mitigation technique. Methods for accounting for such uncertainty, whilst rare, do exist in the literature. Ratcliffe (2000), for example, proposed a method whereby the timing of an aoristic crime is expressed as a probability density, spread uniformly across the potential range of times it could have occurred. This analysis was implemented in Manchester, UK by Fielding and Jones (2012) in a burglary prevention initiative which saw a 26.6% reduction in burglaries over a year when compared to the previous year. A similar technique was described by Brunsdon (1989) when analysing burglary data. The ‘weight’ of each crime was spread uniformly between the earliest and latest time between which it could have occurred. The average weight at any given time of day was taken and could be interpreted as a ‘risk profile’ of residential burglary.

Ashby and Bowers (2013) considered different estimation techniques to examine how using the earliest, latest, or average times (that the crime could have occurred according to the victim) impact on aggregated crime analyses. Using data from CCTV footage to establish the actual time of (bike) thefts at a train station, they found that using either the aoristic approach described by Ratcliffe, or an equivalent approach, closely estimated the distribution of known times. Boldt and Borg (2016) found similar results for residential burglary in a study in which known times were
determined by either burglar alarms recording the exact time they were triggered, reports from victims who were at home at the time of the offence or third-party witnesses. The following sections outline the process and complexities of spatial matching as well as methods that can be used to minimise geocoding inaccuracy. This is followed by a discussion of temporal fidelity and some potential mitigation strategies.

### 3.2.2 Geocoding quality

The quality of geocoded data is determined by three characteristics: positional accuracy, completeness, and repeatability. Positional accuracy refers to the Euclidean distance between a data point’s actual location and the coordinates to which it is geocoded; completeness refers to the proportion of the data that has been geocoded (that is, a ‘match’ is found between the input data and reference data); and repeatability refers to how sensitive the results are to different matching techniques. It should be noted that very high completeness can obviously be achieved by accepting very poor positional accuracy. For example, if a crime is known to have occurred in ‘London’ it can be matched to the centre of London, UK, though it would likely have terrible positional accuracy - especially if it were a dataset for London, Ontario. Different algorithms will set different thresholds (or confidence levels) for what constitutes an acceptable match and how to resolve conflicts. Consequently, there is some level of subjectivity when determining the quality of geocoding. Previous research has attempted to set a minimal standard for completeness. Ratcliffe (2004) analysed data for five different crime types to establish a potential geocoding minimum completeness threshold that needs to be met to avoid significantly different spatial patterns from emerging. This will be discussed in more detail in Section 3.2.4. The focus of this research is on analysing positional accuracy and completeness, which are particularly influential at the stage of initial geocoding (Zandbergen & Hart, 2009), and how these impact on the generation of patrol locations.
3.2.3 Positional accuracy

There exist several automated methods for geocoding an input address to generate a set of coordinates. Each make certain assumptions about where the point should be located and these are now briefly described:

- **Street geocoding** first identifies the street or street segment on which the input address is located. The exact position of the address along the street is then interpolated based on the known building number range. For example, if a street has ten properties on it, numbered 1 to 10, and a crime occurred at property number three then street geocoding would centre the event 30% of the way along the street. The coordinates are generated to either be directly on the street or, more often, positioned approximately 10 metre along a perpendicular offset from the street to account for the fact that the centre of a building will not coincide with the centre of a road.

- **Land parcel centroids** utilise known plots of land (known as land parcels) to set the coordinates as the centre of the land parcel that matches the input address.

- **Address points** are similar to land parcel centroid geocoding but rather than use the centroid, another point within the parcel is utilised, (usually to coincide with the centre of a building or structure).

Each of the above methods have particular strengths and weaknesses depending on the objective of the geocoding and these are discussed below. Street geocoding is the most common method used in many countries for the simple reason that land parcel or address point data sets often do not exist. However, a considerable weakness of street geocoding is the assumption that address spaces are equally distributed along the road segment, which is not always the case. This can lead to more substantial positional errors, particularly in rural environments where road segments are longer and the distribution of buildings less uniform. A study by Cayo and Talbot (2003) found that positional errors associated with street geocoding were
significantly greater in rural areas (95% geocoded within 2,872m of the actual location) compared to suburban (421m) and urban (152m) areas. Whilst parcel geocoding provided consistently greater positional accuracy than street geocoding, there were still distinct differences across rural, suburban, and urban landscapes (95% within 195m, 39m, and 21m respectively). Zandbergen (2009) produced a review of twelve studies (11 conducted in the US and one from Australia) that specifically investigated street geocoding errors and found that the median positional error for residential addresses ranged from 26 to 168 metres.

When land parcel or address point data sets do exist they often result in lower match rates (Zandbergen, 2009). This is due in part to the fact that whilst changes do occur to individual streets or street segments, the addresses on a street can change more frequently. For example, redevelopment may occur which adds new homes to the street or removes old ones. Alternatively, in the case of commercial properties, the owner of the property may change and thus the name of the address may change as well. Errors and contradictions within both the reference and input address data also lead to positional errors. Zandbergen (2009) highlights how these issues can lead to the incorrect location being matched (if they are matched at all) which often results in errors of several kilometres. There are several such errors that can easily arise:

- **Invalid building number.** If the input address specifies an invalid building number, (above the range of the street for instance) the geocoding process may default to the highest (or lowest, or middle) building number in order to produce a match. Alternatively it may reject a match entirely.

- **Misspelt street name.** Geocoding algorithms will attempt to account for spelling errors and find similar matches but these are not always accurate. Consequently, an address may be geocoded to a location far away from its true position.

- **Incorrect street designation.** Input addresses can sometimes be input with the incorrect suffix, for instance ‘road’ where it should be ‘street’ or ‘place’
3.2. Background

etc.

- **Conflicts within the address.** Perhaps the most difficult issue to account for is when there is conflicting information within the address; for instance a postcode which does not coincide with the town or street name.

These types of errors can lead to significant positional errors when matches are made, but also impact whether a match is found at all. Another critical factor effecting the likelihood of a match is when the input address and reference data were compiled. If the input address specifies a new location (such as a newly built residential development or new shop) then an outdated reference data set will be unable to find a match. Similarly, if the the input data being geocoded is historical then the location may no longer exist in the reference data. Thus the reference dataset can have a significant effect on the quality of the geocoding. For example, Mazeika and Summerton (2017) examined residential burglaries in an area of New Jersey (USA) using two separate reference sets; one was a geocoding process built into Google Earth and the second used the US census ‘TIGER’ street geocoding system. The study found that the average geocoding error associated with using the Google Earth reference set was 27 feet in urban environments and 77 feet in suburban environments. The US Census ‘TIGER’ street geocoding system produced substantially greater errors of 122 and 322 feet, respectively.

As discussed above, the positional accuracy and completeness of a dataset can be highly correlated; accepting very poor matches can lead to high levels of completeness at the expense of positional accuracy. The UK has a national address point database which has been used for the analysis of crime records before (Brimicombe, Brimicombe, & Li, 2007). Of particular concern for the analysis of crime is the fact offences can and do happen on the street and so address points or land parcels may not be as appropriate as street geocoding. Brimicombe et al. (2007) outline a process for geocoding crime events that accounts for on-street locations as well as standard addressable locations. Their process is outlined below and was implemented for a large metropolitan area in the UK.

1. **Pre-processing:** The data are cleaned to deal with common errors and
spelling short-cuts. For example, ‘Rd’ is corrected to ‘Road’ and postcode elements are checked for errors such as the letter ‘O’ being mistakenly used instead of the number zero in a position where only a number element can be used.

2. **Individual address geocoding:** A commercial geocoding software package is used to geo-code those locations which can be matched to known address locations. These locations are given a validation code of ‘L1’ indicating they are geocoded to an individual property address with a high level of presumed accuracy.

3. **Nearby features geocoding:** Non-address locations are geocoded by text mining the free text fields and address data for keywords. Most locations geocoded to this level are road junctions and so keywords such as ‘J/W’, ‘JCT’, and ‘JTN WITH’ are used in conjunction with any road names within the fields. This is then compared with a database of all junctions and their associated coordinates which the authors produced themselves using Ordnance Survey road centre line data to determine the correct coordinates of the relevant road or junction. Locations coded through this process are given the validation code of ‘L2’.

4. **Postcode unit geocoding:** Any remaining records with valid postcode data are geocoded to the postcode level and assigned the validation code of ‘L3’. item **Street pattern geocoding:** If the locational data of a crime event includes a road name and has not been coded to one of the above levels, then an attempt is made to code it to a road section if a road name is included in the crime’s address data. This is aided by the fact that a crime incident record will include a reference to the approximate location of the crime which is recorded by the Call and Dispatch (CAD) centre when a crime is first reported. The location is referenced to a cell on a 250 metre by 250 metre grid called CADref. If a match is made between CADref and the road name recorded for the crime then the crime is geocoded to a random point from the relevant section of road.
The validation code of ‘L4’ is given to these locations and they represent the least accurate level of geocoding.

This process appears to be the system currently being used by the MPS, however the author was unable to verify if any changes have been made to the process since it was published. The following section provides an overview of the current understanding of and challenges associated with the completeness of geocoding.

### 3.2.4 Completeness

Perhaps the first attempt to determine an acceptable geocoding completeness level for crime data was conducted by Ratcliffe (2004). Five datasets were analysed, covering different types of location and crime types in New South Wales (Australia). These included regional, urban, and inner-city locations and datasets including all reported crime, vehicle crime, malicious damage, or burglary. For each dataset a thematic map - generated using census blocks\(^1\) - was produced and then subsequent maps were generated with progressively fewer data points included. Each iteration removed 1% of the data and the map was then compared with the original map using a Mann-Whitney U test to check for significant differences between them at the 0.01 significance level. Once a significant difference was detected the percentage of crimes removed was recorded and the process repeated 250 times to generate a frequency distribution. His analysis produced acceptable minimum geocoding completeness thresholds of 78% to 85% dependent on the crime types and geographic areas.

This represented an important step towards understanding the issue of data completeness in spatial analyses of crime, though several limitations existed within the study. In particular, the removal of crimes from the complete data set was random, whereas the reality of crime data suggests that some crimes will be more likely to be geocoded than others. For instance, burglary is likely to have a higher completeness level than, say, robbery because the former is by definition a crime

---

\(^1\)A census block is the smallest geographical administrative area used by the US Census Bureau and, as the name implies, usually consists of one street block (though in more rural areas can be larger). Unlike UK administrative areas, such as Output Areas, census blocks are not defined based on population and many have a population of zero.
conducted at a property whilst the latter is a crime conducted against an individual. Furthermore, certain locations will be more likely to generate a match than others. For instance, as discussed earlier geocoding matches are less likely to be made when the land use of a particular area has recently changed and if reference data files are not updated. Furthermore, Zandbergen (2008) investigated geocoding quality issues in Florida (US) and found a significantly lower match rate existed within commercial areas when compared to residential areas, and differences were also apparent between single family residences and multi-unit residential addresses.

Further, the analysis conducted by Ratcliffe (2004) aggregated crimes to the census block level to produce thematic maps for comparison; different levels of aggregation are likely to lead to different minimum completeness thresholds. This issue is known as the modifiable areal unit problem (MAUP), and occurs because of the arbitrary selection of boundary locations for most spatial analyses. Although the issue has been recognised since the 1930 it was aptly summarised by Openshaw (1984):

> “the areal units (zonal objects) used in many geographical studies are arbitrary, modifiable, and subject to the whims and fancies of whomever is doing, or did, the aggregating”

The MAUP is clearly illustrated in Figure 3.1 which provides a basic example of the issue. Assume that each square, (a) and (b) represent some area of interest such as a town centre. Say, for example, we had sub-divided the town centre into four regions which might be used to direct police patrols and that we had aggregated our crime counts to these sub-division levels. If our sub-division was as in (a) then we would conclude that patrols should be directed to the north and south regions of the town centre. However, if our sub-division was as in (b) then each region of our town centre appears equally at risk of crime occurring.

### 3.2.5 Temporal fidelity

Issues of spatial inaccuracy within recorded crime data can potentially be improved by better record management, as well as more comprehensive reference sets.
3.2. Background

Figure 3.1: An illustration of the modifiable areal unit problem.
(a) suggests clustering in the north and south of the area; (b) suggests an even distribution of crimes across the whole area.

example, the inclusion of bus stop locations as addressable points in the reference data may make it easier to correctly geocode events. Temporal inaccuracies have a distinct challenge; the issue of uncertainty over when the crime may have occurred as victims may simply not know. For example, in the case of residential burglary a victim is often not at home when the crime occurs or if they are, they may be asleep and only discover the crime has been committed in the morning when they get up.

The temporal variations of crime, across hours, days, weeks, and seasons, has been noted in the literature (Baumer & Wright, 1996; Cohn, 1993) and is recognised as an important factor which is crucial to crime analysis (Ashby & Bowers, 2013; Ratcliffe, 2000). Crimes where the exact time at which they occur is unknown are referred to as aoristic crimes. Methods for accounting for such uncertainty, whilst rare, do exist in the literature. As mentioned in Section 3.2.1, Ratcliffe (2000) provides a method for analysing aoristic crimes by effectively ‘spreading’ the probability density of the timing of the crime across the potential range of times it could have occurred.

Ashby and Bowers (2013) considered different estimation techniques to demonstrate how using, for instance, earliest, latest, or average times impact on aggregated crime analyses. Using data from CCTV footage to establish the actual time of (bike) thefts at a train station, they found, that using the aoristic approach described by Ratcliffe, or randomly selecting a time between the earliest and latest
time on a crime-by-crime basis most closely estimated the distribution of known times. However, it is difficult to know if those crimes for which the exact time was known are representative of all crimes as a selection bias may exist. It may be the case that the crimes for which the exact time of the offence is known could only be pin-pointed at that time because they occurred during a period of the day when officers were more able to dedicate time to detecting them.

A study of residential burglary in Sweden also found that aoristic analysis provided a much better approximation of known crime event times. Known times were determined by either burglar alarms recording the exact time they were triggered, reports from victims who were at home at the time of the offence, or third-party witnesses. Again, random selection of a time between the earliest and latest recorded time produced similar results (Boldt & Borg, 2016). The known times for this study are likely to be more robust due to the method by which they were determined. These examples in the literature demonstrate that, whilst perhaps uncommon, methods for dealing with aoristic crime variables exist and may provide significantly better results than researchers arbitrarily choosing a particular time code to use in their analyses.

3.2.6 Police movement data

The growth in usage of Global Positioning System (GPS) devices by police forces provides a capability for tracking where officers move with much greater accuracy than has previously been possible. Potentially the first study to utilise GPS data to measure dosage was conducted in Peterborough (UK) using radio data for the movement of Police Community Support Officer (PCSO) (Ariel et al., 2016). In that study, GPS ‘pings’ were recorded every minute, and the dosage in a hot-spot was measured as the time between the first GPS ping within the hot-spot, and the first ping without. This study found that police patrols had a significant impact upon crime and disorder, with patrols lasting up to around 15 minutes and on average lasting around 8 minutes. This study was ground-breaking in its use of GPS data to measure dosage but many questions still remain, particularly regarding how frequently an area should be visited and the rate at which initial deterrent effects
3.2. Background

Decay (Sherman, 1990). It is also worth noting that the original intention was for patrols to last 15 minutes each and for three to be conducted in each hot-spot each day for a total of 45 minutes of patrol time per day in each hot-spot. However, in practice this was not achieved and the effect was that more patrols were conducted but for shorter durations, leading to a total of 37 minutes of patrol time per day in each hot-spot - 18% less time than originally planned. This highlights again the issue of employing an intention-to-treat method of evaluation which would have potentially underestimated the impact of patrols in the study. Furthermore, in the analysis conducted, patrol dosage was measured in the aggregate over the entire study period, meaning that variation in patrol dosage per day, or per police shift, was not considered.

Williams and Coupe (2017) also measured police patrol dosage using GPS data. Specifically, the study was concerned with whether more frequent but shorter periods of patrol dosage (nine periods of five minutes each) had a greater or lesser impact on crime than less frequent but longer (three periods of 15 minutes) patrols. Recall that the ‘crackdown, back off’ theory hypothesised by Sherman (1990) stated that the deterrent effect generated by police patrols ‘decays’ once there is no police presence. Williams and Coupe hypothesised that more frequent patrols, “might arguably allow less time for what Sherman calls “deterrence decay” to kick in, so that there would be less crime.” However, their findings suggest that the longer, less frequent patrols were more effective at preventing crime. Unfortunately, Williams and Coupe did not report on the accuracy of their (GPS) data or the sensitivity of their findings to the kinds of problems explored in this chapter, meaning that their findings may be open to errors of inference.

To elaborate, GPS data come with some drawbacks. Chief among these is the fact that signals do not account for every step in an officer’s path. In their study, Ariel et al. (2016) were able to use 1-minute refresh rates. However, discussions with three UK police services suggest that operational ping rates are generally every two to five minutes due largely to data collection costs and radio battery life considerations. Williams and Coupe (2017) did not report the time between GPS
pings in their study. Given the delays between the recording of foot-patrol locations (even if this is only one minute), to establish the paths taken between GPS pings requires interpolation. If employed as a micro-level measure of dosage, this can introduce errors into patrol evaluations (which will increase with the latency between GPS pings).

The studies discussed so far, and the illustration provided later in this chapter, are concerned with foot-based police patrols. Vehicle-based police patrols are less likely to be impacted by the GPS issues discussed for two key reasons. First, Automated Vehicle Locator (AVL) GPS pings usually occur much more frequently; either every 10 to 15 seconds or every few hundred metres of travel (for an example, see Weisburd et al., 2015). Second, vehicles are confined to the road network and as such their potential paths between pings are much more restricted and thus easier to interpolate accurately when compared to foot-based officers who have no such restriction. For this reason, we focus our attention on foot-based patrol dosage and remain mindful that the issues discussed may occur, albeit in a more limited way, for vehicle-based patrols.

GPS data are not a panacea to issues of location measurement. Systematic bias can exist within GPS location measurements and come from various sources; satellite orbital errors and clock bias, receiver clock errors, refraction in the ionosphere and troposphere, and signal multipath (He et al., 2011). Whilst there are methods for reducing or removing most of these errors, signal multipath - where the signal between the satellite and receiver is reflected by large objects, causing non-direct paths to be taken - is distinctly problematic. Within an urban environment this is known as an ‘urban canyon’ issue; whereby tall buildings (or other structures) interfere with the GPS tracking system, causing it to incorrectly interpolate the location as signals are reflected off of buildings, (see Figure 3.2). A study in the town of Gorlitz, Germany found that the average measurement error in areas with broad streets and few tall buildings was 2.5 metres, whilst areas with narrow streets and mostly tall buildings had a average measurement error of 15.4 metres (Modsching, Kramer, & ten Hagen, 2006). Despite these challenges, GPS data provide a valuable method
by which patrol data could be more accurately computed than was possible using previous methods.

![Diagram of the urban canyon effect](image)

**Figure 3.2:** The urban canyon effect
Perceived locations differ from actual locations due to GPS signals reflecting off of tall buildings.

### 3.2.7 Methods of evaluation

As discussed in Section 2.2.3, many studies have evaluated the effect of police patrol on crime. A standard method for such studies is a difference in difference test whereby comparisons are made between a period before, (and sometimes after) an implementation has begun and the period of implementation: the study period. These studies compare changes between the two periods in areas that have experienced the implementation (treatment areas) to areas that were excluded from the implementation (control areas). The usual comparisons made are between aggregated crime rates, fear of crime, perceptions of crime, or other relevant metrics. In the case of hot-spot policing evaluations, the hot-spots have usually been designated based on historic crime trends (for instance, over the previous year) and do not ‘move’ during the study period (for examples see Braga & Bond, 2008; Di Tella & Schargrodsky, 2004; Ratcliffe et al., 2011; Sherman & Weisburd, 1995; B. Taylor, Koper, & Woods, 2011; Telep et al., 2014; Weisburd & Green, 1995). Another
popular method used to evaluate the impact of social policy and law enforcement practices is the interrupted time series (for examples see Koper & Mayo-Wilson, 2006; Novak, Hartman, Holsinger, & Turner, 1999) though Kleck, Britt, and Bوردua (2000) highlight several flaws with this approach:

... users have largely ignored or minimized its [interrupted time-series design] flaws, including:

1. its general inability to rule out alternative explanations,
2. the use of a single or small number of arbitrarily chosen “control” or comparison jurisdictions,
3. arbitrary definition of the endpoints of the time series evaluated,
4. an inability to specify exactly when the intervention’s impact is supposed to be felt, raising problems of the falsifiability of the efficacy hypothesis, and
5. an atheoretical specification of the ... model.

A restriction of both these general evaluation methods is that they lead to the loss of important causal information regarding individual event and the time between events (Freeman, 1989). The remainder of this chapter takes a hot-spot policing operation implemented in the UK to illustrate the above issues, placing particular emphasis on the accuracy and uncertainty of the crime and officer movement data used in an evaluation of the operation.

The present research uses data from a trial conducted by the MPS in the London borough of Southwark (UK), where they used a prospective hot-spot technique to direct foot patrols. For each police shift (of which there were three per day; 7am to 3pm, 3pm to 11pm, and 11pm to 7am) as part of their tasking brief, police officers were given maps on which a series of 250 by 250 metre ‘prospective boxes’ were identified for police patrols.

The aim of the exercise is to illustrate the issues discussed above by working through an example. To do this, the estimated impact on crime of variation in police patrol dosage is examined, placing particular emphasis on the strengths and limitations of the data in terms of data accuracy and usability. Before describing the methods and analyses, the next section outlines the data used in greater detail
3.3 Data

3.3.1 Intended Patrol Locations

Intended patrol locations were identified using a proprietary predictive algorithm developed by a third-party partner working with the MPS. It is based on the principles proposed by Bowers, Johnson, and Pease (2004). The algorithm produces a risk score (essentially time-weighted kernel density estimator) for each 250 by 250 metre cell of a grid that covered the study area. For each shift, the cells most at risk were provided to patrol officers with the intention that they would patrol those areas when possible. These prospective boxes were generated for seven crime types (burglary, theft from a person, theft from a motor vehicle, theft of a motor vehicle, criminal damage, robbery, and violence with injury (VWI)) based on the priorities of The (London) Mayor’s Office for Policing and Crime (MOPAC) - generally referred to by the MPS as ‘the MOPAC 7’. The number of boxes generated for each shift and each crime type varied from three to ten, and a total of 5,697 boxes were identified over the study period. However, officers had some discretion as to which crime types to prioritise and, consequently, exactly which boxes they sought to patrol during any given shift is unknown.

The locations of the boxes changed frequently. Over the two-month period, 388 unique locations were designated as prospective boxes. The least common boxes were identified only once, while the most common were identified 155 times. On average, a location was a prospective box during 21.65 police shifts.

3.3.2 Officer Location Data

Officer movement data were collected in the form of GPS ‘pings’ from body-worn radios. Pings were sent from the radios whenever an officer initiated a call or every five minutes. They have a Circular Error Probability of five metres – that is, for 50% of pings the true location is within 5 metres of the reported location (95% within...
The data include the officer’s call-sign, the time (to the nearest second), and their location (specified at a resolution of one metre). A total of 239,115 officer pings were recorded in Southwark during the study period.

To estimate how much time each officer spent at each location it was necessary to interpolate their location between pings. A variety of approaches could be taken to do this; here, a ‘join-the-dots’ method was used, for which the ‘assumed path’ was taken as the direct line between two sequential pings. Where the assumed path intersected multiple grid cells, the officer was assumed to be walking at constant speed between the two pings and their entry and exit times for each cell were calculated accordingly. The amount of time they spent in each cell was then estimated.

A number of further processing steps were also implemented to address possible sources of error. Where the time between consecutive pings was greater than 15 minutes, this section of the path was discarded from the analysis due to concerns that the officer’s actual path may be substantially different from the assumed path. Similarly, where the speed at which the officer appeared to be moving was greater than 2 metres per second the assumed path was discarded as it was assumed the officer was not on foot.

If more than one officer was in a prospective box at the same time, dosage was calculated as the union of their times in that cell. For example, if two officers were in the same prospective box at the same time for N minutes, the estimated dosage would be N minutes not 2N. This approach was used as previous research suggests that the presence of more officers does not necessarily lead to greater deterrence (Kleck & Barnes, 2010).

### 3.3.3 Police Recorded Crime

Police recorded crime data were provided for the MOPAC7 crime types described in Section 3.3.1 above. Theoretically, each of these crime categories require an offender to pass through a public space in order to reach an offending location and thus, the literature suggests, will be susceptible to a deterrent effect from capable guardians; in this case visible police officers. A total of 2459 incidents occurred during the study period. Data provided detailed the crime type, offence location, and
3.3. Data

the earliest and latest dates and times at which the offence could have occurred. The
temporal uncertainty of three crime types (residential burglary, robbery, and VWI)
is shown in Figure 3.3. These crime types were chosen as they represent known
priorities for the police during the study period and demonstrate the differences
between person-targeted crimes (robbery and VWI) and property-targeted crimes
(burglary). The time at which person-targeted crimes occurred is generally known
quite precisely; on the other hand, the exact timing of a substantial proportion of
burglaries is uncertain.

![Figure 3.3: Uncertainty of the timing of specific crime types](image)

3.3.4 Measuring crimes within prospective boxes

The location of each crime is uniquely defined within the dataset; however, the
precision with which each location is given is not consistent. The most precise
geo-coding level is defined by address point data and 72.1% of MOPAC7 crimes
within our dataset fell within this category. 22.5% of crimes were only coded at the
postcode level, which is less accurate, having a resolution of up to several hundred
metres. The remaining 5.4% of crimes were either coded based on other location
information (such as road junctions or train station), to street level, or their geo-
coding precision was unknown. Again, the level of precision varies by crime type,
with 96.0% of residential burglaries being coded to an exact address, but for VWI
and robbery the figures were 75.8% and 51.9% respectively. While the proportion
of person-targeted crimes that are geo-coded to the postcode level appears high,
there is no reason to believe this is abnormal for these crime types; indeed, this
highlights the challenge of trying to analyse crime data at high spatial resolutions.

These uncertainties associated with the timing and location of offences must
be taken into consideration when interpreting the results of any study that focuses
on such fine temporal and spatial resolutions, including this one. To avoid attrition
in the data, in the analysis that follows all data were analysed, and the dataset there-
fore contains 2459 crimes. Of these, 108 were recorded as occurring within ‘live’
prospective boxes.

### 3.4 Illustrative analytic strategy

The data described above was used to estimate the amount of patrol dosage that
was applied to each ‘box’ and how much was necessary to effectively deter crime.
The unit of analysis was the ‘shift-box’; that is, each instance of a grid cell being
identified as a prospective box in each shift during the study period. Where a loca-
tion was identified for multiple crime types during the same shift, this was treated
as only one shift-box to avoid double counting. The total number of shifts for the
study period was 183. However, prospective box locations for five shifts were not
archived by the police due to a technical error. As such, the data for these shifts are
excluded from all analyses.

Of the 5,697 shift-boxes identified for deployment, 3678 boxes (64.6%) re-
ceived an estimated dosage of zero. This is interesting, but perhaps unsurprising
given the quantity of boxes identified over such a short time frame. It thus appears
that the resourcing required to cover the number of boxes identified at the frequency
they were produced was too great. This also illustrates an important shortcoming of
the ‘intention-to-treat’ approach to evaluation which was discussed in section 2.2.4,
which would wrongly assume that all boxes received intervention. As illustrated in Figure 3.4, of those boxes that did receive dosage, few received more than one hour of estimated dosage in any given shift. Furthermore, the presence of many low (but non-zero) values in Figure 3.4 suggests that many boxes are likely to have received only inadvertent dosage from officers during the course of other duties.

![Distribution of dosage duration per (eight hour) shift for prospective boxes that received non-zero dosage](image)

**Figure 3.4:** Distribution of dosage duration per (eight hour) shift for prospective boxes that received non-zero dosage

Since the recorded crime data are subject to temporal uncertainty (i.e. ‘aoriastic’, as discussed in Section 3.2.5), an analytic choice must be made with respect to how the time of each incident is estimated. The impact of this choice is partly mitigated here since the analyses that follow are conducted at the shift level. For this analysis, the earliest recorded time is used. This is to minimise the possibility that a crime was estimated to have occurred after a patrol when in fact it happened before. This approach guarantees that patrols will not be falsely recorded as having occurred before a crime, but does increase the risk that a crime is erroneously estimated as occurring before a patrol. This trade-off illustrates a further issue with evaluations of this kind and the data on which they are based.

For the purposes of this study, a randomised controlled trial was not possible,
and hence a quasi-experimental approach was adopted. The approach taken here was to compare the count of crime in each box with a suitable control. Defining the set of control boxes was challenging given the fact that the risk of crime is dynamic, varying in both space and time. For this reason, the effect of intervention was estimated by comparing the count of crime in a ‘live box’ with the count of crime in that same location at an earlier time, in this case one week earlier. The option of selecting the same box at an even earlier point in time (e.g. two or three weeks before) was not possible here due to the limited period for which data were available. This design has the advantage of controlling for factors that vary spatially and over the course of the week, but it does not control for other factors. Chief among these is that a control box may itself have been a live box that received some patrolling. While removing such occurrences was considered and tested, this would have led to considerable attrition in the data – which was already limited given the number of crime events examined (N=108). Furthermore, omitting such boxes would, in effect, remove any persistent hot-spots from the analysis and introduce another systematic issue. This highlights a further challenge to evaluations of this kind. Alternative approaches, along with their strengths and weaknesses, are considered in the discussion section.

3.5 Results

Figure 3.5 shows the cumulative count of crime observed in live boxes as a function of the dosage delivered in them. This enables us to see (for example) the total number of crimes that occurred in those live boxes for which up to 10 minutes of patrol dosage (in this case 73 crimes) was delivered, and so on. The curve for the live boxes can then be compared with that for the (matched) control boxes. The difference between the two is also shown.

That the prospective boxes that received no (estimated) dosage during the live period had fewer crimes than during their corresponding control periods seems counter-intuitive. Given that the live boxes are anticipated to be at an elevated risk of crime during the ‘live’ interval, the natural assumption would be that they would
experience more crime in the absence of any intervention. There are, however, reasons why this expectation may be unrealistic.

One explanation is regression to the mean; it is possible that there would be no sustained elevation in risk and the crime rate would naturally subside without any police action. A related explanation is that a selection effect is at play. That is, given that they could not visit every box, officers may have used their local knowledge and avoided patrolling those boxes that (they perceived) were unlikely to actually be at an elevated risk (those for which regression to the mean was likely). It is important to note that this kind of selection effect could equally occur in a randomised control trial and that this would go undetected for an evaluation that used an intention-to-treat design.

If the quantity of patrol dosage had no impact on the ensuing crime rate, the difference between live and control curves would be expected to remain approximately constant. As shown in Figure 3.5, this difference clearly increases between about 10 and 20 minutes of dosage, suggesting that boxes that received these amounts of dosage experienced substantially less crime than their control counterparts. This suggests that patrol dosage has a non-linear impact on crime and that a threshold

**Figure 3.5:** Cumulative crime count in live and control boxes
minimum dosage is required for there to be a deterrent effect, as has been reported elsewhere (Koper, 1995). Again, this is important and would be missed in an evaluation that employed an intention-to-treat design.

To estimate whether the differences observed were statistically significant, a Monte Carlo simulation was used to compare the observed distribution with that expected, assuming the two distributions (those for the live and control periods) were really drawn from the same one (the null hypothesis). To do this, for each live-control pairing, we generated a synthetic distribution by shuffling (or not) the count of crime observed in one period with that observed in the other. Whether the data for each pairing were shuffled (probability 0.5) or not was determined by the output of a uniform random number generator. A total of 5000 synthetic data sets were generated and the difference in the two distributions plotted for each (see Figure 3.6). Figure 3.6 shows the 95%, 99%, and 99.9% confidence intervals - computed assuming the null hypothesis to be true - for the difference between the control and live groups.

![Figure 3.6: Cumulative crime counts with Monte Carlo simulated significance bands](image_url)
3.6 Conclusion

The results shown in Figure 3.6 suggest that the difference between live and control periods was non-significant (at the 99% confidence interval at least) for a patrol dosage of less than 13.5 minutes. For higher doses, the difference is clearly statistically significant, suggesting an impact of patrolling but only when estimated patrol dosage exceeds about 14 minutes. After 14 minutes of dosage, the estimated benefits appear to increase no further.

3.6 Conclusion

Given growing interest in prospective hot-spot techniques (Perry et al., 2013), and the availability of GPS data to evaluate such interventions, the aim of this chapter was to discuss some of the concerns surrounding the quality of GPS and police recorded crime data, to inform the design of future studies. To make the ideas concrete, these issues were illustrated using data from a real-world example. The data available for the illustration was quite limited and so too was the approach to evaluation. As such, some caution is necessary when interpreting the findings. However, it would seem remiss not to at least comment on the findings briefly, regardless of how speculative they might be. Overall, the results presented chime with those of other studies (Ariel et al., 2016; Koper, 1995) and suggest that police foot patrols need to exceed a threshold of about 10 minutes to produce their intended effects and that after approximately 20 minutes they have little further impact.

In terms of implementation practicalities, many of the designated patrol boxes did not receive any dosage at all, something that an ‘intention-to-treat’ study design would fail to uncover. As well as illustrating the need to explicitly measure dosage, this draws attention to the need for all implementations to carefully consider resource limitations. In future trials (or roll outs), police forces may find they can be more effective (and realistic) by matching the number of hot-spots they select for patrol with the resources available to them on the day, rather than trying to patrol a larger number of hot-spots less effectively (or not at all).

The findings also have implications for future research (see Chapter 7). While GPS data offer a more precise picture of where foot-patrol officers are at a particular
time, ping rates are currently relatively infrequent and hence methods of interpolation are necessary to estimate officer paths. Using Euclidean paths will incur some inaccuracy, particularly in cities where potential paths are highly confined by the environment. Improving these estimates warrants greater attention, or at the very least, researchers who use GPS data must be cognisant of the potential errors that are introduced into their analyses. Evaluators need to be transparent about the quality of the data used and methods of interpolation employed. They should also report the sensitivity of their results to different methods of interpolation, or variation in the parameters used to derive the estimates. As a minimum, they should report the ping rate for GPS data. Estimating the exact path taken by officers more accurately could significantly impact on the findings of studies such as this and Chapter 6 looks more closely at how the ping rate influences the assumed path an officer takes.

The use of GPS data also has drawbacks when compared to traditional methods such as the use of police logs or independent observers. For instance, whilst logs and observations can help measure what the officer is doing whilst at a location, GPS data cannot. To reduce uncertainty regarding officer activity, several approaches might be taken in future studies. For instance, officer dispatch logs might be cross-referenced against the GPS data to provide some information regarding officer activities. Furthermore, and looking to the future, the increasing ubiquity of internet enabled devices (including police body worn video and other wearable devices) might provide opportunities to capture activity passively. In defence of the general approach taken here, there is a difference between what officers are doing and how they are perceived by those who might observe them.

The next chapter builds upon some of the findings from this evaluation. The temporal unit of analysis used here was the police shift. This was for practical policing purposes, however, the risk within different boxes is likely to vary at different times of the day depending on the crime type in question (GrubesicMack2008). The next chapter reduces this concern by aggregating dosages to each hour of the day. It also incorporates into the analysis the recent patrol dosage in an area. Whilst the focus here was on the prospective boxes which were targeted for patrols, Chapter 4
analyses the police presence across an entire borough of London. A different analytical approach is also adopted to avoid the challenges outlined here regarding the definition of a suitable control site.

To conclude this chapter, research suggests that hot-spot policing interventions work (Braga et al., 2019). However, evaluations have rarely examined the relationship between patrol dosage and its impact on crime. Understanding this association is important if the police are to make the best use of the resources they have available. The findings presented in this chapter are only speculative but they are at least consistent with the handful of studies that have looked at this issue in the past. They also highlight how intention-to-treat designs are likely to be inadequate for assessing effectiveness as they assume perfect implementation rather than measuring the actual amount of resource allocated.
Chapter 4

Modelling the deterrent effect of police patrol

4.1 Introduction

In 2013, the MPS began a process of evaluating commercial predictive crime mapping systems, having recently developed their own algorithm based on the work of Bowers, Johnson, and Pease (2004), which had been successfully implemented in Manchester, UK (Fielding & Jones, 2012). The MPS internal algorithm had already been operationalised before the evaluation began, and most London boroughs were already using either a first or second iteration of the algorithm to deploy their daily patrols. Although the MPS evaluation did find differences in predictive accuracy between the systems, it rightly cautioned that these were likely influenced by a range of factors that had not be controlled for in the experiment. These included factors such as the total area defined as ‘hot’ by each system, whether or not hot-spots could overlap, and the lack of data regarding police patrol within these hot-spots (Bryant, Azhar, Blackburn, & Falade, 2015). The impact of this last point was aptly stated:

*If dosage rates are sufficiently high then police visibility might have a dissuasive effect on certain crimes and hence reduce the hit rates (based on recorded crimes) and the associated [Predictive Accuracy Index]*

In other words, if police officers were present in these hot-spots and their presence does have the deterrent effect that police patrols are predicated on, then the
predictive crime mapping system may appear to perform worse than if it had not been operationalised successfully and instead officers had been sent to locations where crime was actually less likely to occur.

Although the Bryant et al. report did not investigate the accuracy of the crime data, it did highlight that previous studies have shown that geocoding inaccuracies can be substantial. The results presented in Chapter 3 are based on some of the same data as those used by Bryant et al. and highlighted that recorded crime data is often not exact. They show that the level of spatial and temporal precision varies greatly by crime type; for example whilst 96.0% of residential burglary crimes are accurately spatially recorded, only 75.8% of violent crimes and 51.9% of robberies were geocoded to an exact address point. Conversely, whilst over 85% of violent crimes and robberies are recorded to within a one hour window, this drops to only about 30% of burglaries.

The analysis in this chapter incorporates an estimate for the amount of police patrol time across the entire study area at a resolution based on a 250 metre by 250 metre grid cells. This differs from the analysis conducted in the previous chapter where only the patrol dosage within prospective patrol boxes was utilised. As context, during a police shift only 1-2% of a borough would be designated as prospective hot-spot. Hence the analysis which follows incorporates a much larger, more comprehensive dataset. Patrol dosage in each grid cell is estimated in a similar way to that outlined in Chapter 3; but interpolating officer patrol paths between GPS pings which occur every five minutes. The limitations and challenges of using these data are the focus of Chapter 6.

The MPS in particular (and police services more generally) have a clear interest in understanding how data error and uncertainty inherent in both the crime data and officer patrol data used to evaluate these systems impact the results of these studies. As hot-spot policing initiatives have become more spatially and temporally focussed, it has become more important to understand how the fidelity of the data being used impacts on the determination of where hot spots exist (and thus where officers should be deployed) and the efficacy of hot spot policing strategies. The
4.1. Introduction

The purpose of this chapter is to provide an initial baseline evaluation of a prospective hot-spot policing strategy that was implemented in the borough of Islington, London. The intent here is to construct a suitable methodology for micro-level impact evaluation, but to use the data without giving greater attention to the effect of possible data error than is normal within the literature.

At the beginning of each police shift officers were provided with the locations of several ‘prospective boxes’ - locations which were believed to be at heightened risk of crime during that shift. Officers were tasked with visiting these prospective boxes during their shift in an effort to generate a deterrent effect by providing capable guardianship in the form of high visibility foot patrols. This evaluation is designed to represent a first attempt at estimating the effects of the intervention at preventing crime. In Chapter 7, the effects of the intervention will be re-estimated after identifying the effects of error in police patrol data. The two sets of evaluation estimates will then be compared to assess how biases within the data might affect statistical inference and ultimately, evaluation outcomes.

To make this comparison meaningful it is important that the evaluation methods used are appropriate and well fitted to the spatial and temporal characteristics of the intervention. Evaluations of crime prevention usually model effects for units of time at the resolution of months or years (e.g. Andresen & Hodgkinson, 2018; Ariel et al., 2016), however, the focus of this evaluation is a much finer temporal resolution - hours. This resolution is used to better estimate the effects of police patrol whilst it is happening and immediately afterwards. This is likely to be the point at which the patrol will actually have an effect on crime. The amount of police presence is aggregated to a 250 metre by 250 metre grid based on the Call and Dispatch (CAD) grid used by the police. Given the very small spatial areas being analysed it was not computationally feasible to model each ‘patrol’ in a 250 metre by 250 metre grid cell individually. The rate at which officers move between cells meant that the number of ‘patrols’ were very high and the length of each patrol very short if computed in this way. Hence, the amount of patrol presence in each cell was aggregated to the hour. Further details are provided below.
The modelling approach applied in this chapter takes into account the possible inter-dependent effects of crime events by allowing historic events to affect the likelihood of future ones. A spatio-temporal point process analysis is conducted which allows for these effects to be estimated whilst also incorporating spatio-temporal covariates; in this case, the amount of police patrol dosage. Point process models have previously been used in the study of crime. For instance, Baudains, Belur, Braithwaite, Marchione, and Johnson (2019) constructed a series of models to measure the effect of police presence in countering political dissidents in region of Andhra Pradesh, India for the period of 2000 to 2010. By using data describing the daily number of attacks by insurgents, and the police response to these attacks, a point process model framework allowed them to evaluate the dependence of different types of events on each other as well as evaluating how the likelihood of an event was conditional on historic events and nearby events.

Point process models have also been used to predict the locations which are most at risk of crime. PredPol, which was discussed in Section 2.1.1 uses a point process model to predict where crime is most likely to occur in the near future. This has been used to help guide police patrol deployments and been found to predict where crime will occur better than an experienced crime analyst (Mohler et al., 2015). The rest of this chapter describes a series of models that were generated to estimate the effect of police patrol on crime. Simple models are generated at first, attempting only to estimate the background rate of crime. These are then added to by incorporating police patrol dosage within the models. Finally, the models are expanded to allow for inter-dependence of crime events.

4.2 Data

The evaluation is conducted for the London borough of Islington for the period from the 1st of January 2016 to the 31st of March 2016. This period was selected as there was a concerted effort to implement and follow a predictive crime mapping based patrol strategy during this time. The MPS provided two datasets for this evaluation: police recorded crime events and officer movement data which are now described.
They have broadly the same structure as the data discussed in Chapter 3 only for a different time period and different location. The crime data also included a field specifying the assumed accuracy of the geocoding.

### 4.2.1 Recorded crime data

Police recorded crime data were provided for all seven MOPAC priority crime types. The data fields provided for each offence are detailed in Table 4.1 along with a description for each field. The final field, Geocoding accuracy, is further detailed in Table 4.2 and provides a measure of how the MPS automated geocoding system deals with conflicts with the address data for an offence. For example, a complete and real address would be coded as L1 - the highest level of accuracy; a vehicle collision which happened at the intersection of two streets might be coded as L2; and a crime which has conflicting street and postcode data might be coded as L3. As discussed in Section 3.2.3, this geocoding process appears to be the system designed by Brimicombe et al. (2007), though the author was not able to verify whether or not any changes to the process have been made.

**Table 4.1: Recorded crime attributes**

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRIS number</td>
<td>A unique identifier for each offence recorded in the Crime Reporting Information System (CRIS)</td>
</tr>
<tr>
<td>Offence type</td>
<td>The crime category of the offence</td>
</tr>
<tr>
<td>Offence address</td>
<td>The postal address at which the crime occurred</td>
</tr>
<tr>
<td>X</td>
<td>Geocoded British National Grid Easting coordinate</td>
</tr>
<tr>
<td>Y</td>
<td>Geocoded British National Grid Northing coordinate</td>
</tr>
<tr>
<td>Crime reported date</td>
<td>The date on which the crime was reported</td>
</tr>
<tr>
<td>Crime reported time</td>
<td>The time at which the crime was reported</td>
</tr>
<tr>
<td>Crime start date</td>
<td>The earliest date the crime could have occurred</td>
</tr>
<tr>
<td>Crime start time</td>
<td>The earliest time the crime could have occurred</td>
</tr>
<tr>
<td>Crime end date</td>
<td>The latest date the crime could have occurred</td>
</tr>
<tr>
<td>Crime end time</td>
<td>The latest time the crime could have occurred</td>
</tr>
<tr>
<td>Crime description</td>
<td>The officer’s open-form description of the offence</td>
</tr>
<tr>
<td>CAD recorded location</td>
<td>The offence location details as entered into the system by the dispatch officer when the crime was first reported</td>
</tr>
<tr>
<td>Geocoding accuracy</td>
<td>A classification of how accurate the automated geocoding is</td>
</tr>
</tbody>
</table>
### Table 4.2: Crime geocoding accuracy classifications

<table>
<thead>
<tr>
<th>Code</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Geocoded based on address data and address point data</td>
</tr>
<tr>
<td>L2</td>
<td>Geocoded based on other location information such as road junction data and (train) station name</td>
</tr>
<tr>
<td>L3</td>
<td>Geocoded to the postcode level</td>
</tr>
<tr>
<td>L4</td>
<td>Geocoded to the street level with Grid Reference buffer</td>
</tr>
<tr>
<td>L5</td>
<td>Precision level unknown</td>
</tr>
</tbody>
</table>

A total of 2653 MOPAC 7 crime events were recorded during the study period and Table 4.3 provides a break down by crime type. Theft of and from motor vehicles and theft from a person are all sub-categories of the Home Office major categorisation ‘Theft and Handling’ and are combined from this point onwards. In working out the appropriateness of a point process modelling strategy, it was necessary to first establish whether space-time clustering was actually inherent in the data. To this end, a simple Knox test (Knox, 1964) was conducted to assess whether the MOPAC7 crimes clustered in space and time. A knox test compares the observed pairs of events which are considered ‘close’ in space and time to a simulated distribution that would would be expected if the events were random. The simulated random distribution is generated by shuffling the dates of each event and repeating this process a number of times to generate an expected distribution. This repetition is known as a Monte-Carlo simulation. What counts as being ‘close’ is defined by the user. Based on previous research by Johnson and Bowers (2004) and the fact that the grid cells are 250 metres wide, it was decided that a crime would classed as spatially close if it occurred within 500 metres of another crime and within one week. 200 permutations were conducted and the number of close pairs was found to be significantly greater than would be expected by chance ($p=0.015$), indicating space-time clustering in the data.

The spatial and temporal distributions of these crimes are shown in Figures 4.1 and 4.2. In the analysis that follows all measures of time have been specified in minutes and the code used to perform the analysis also requires all periods to be specified in the same format and so time is measured in minutes from the start of the study at midnight on the 1st of January 2016. The focus of the predictive
patrol strategy in place in Islington at the time was on burglary and so the analysis that follows has been conducted with a focus on this crime type, though results are also presented from an analysis of all MOPAC 7 crimes. During the study period a total of 324 burglaries were recorded at 302 unique locations, giving 22 repeat-victimisations.

<table>
<thead>
<tr>
<th>Crime type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burglary</td>
<td>324</td>
</tr>
<tr>
<td>Criminal damage</td>
<td>443</td>
</tr>
<tr>
<td>Robbery</td>
<td>190</td>
</tr>
<tr>
<td>Theft from a person</td>
<td>632</td>
</tr>
<tr>
<td>Theft from a motor vehicle</td>
<td>379</td>
</tr>
<tr>
<td>Theft of a motor vehicle</td>
<td>161</td>
</tr>
<tr>
<td>Violence against the person</td>
<td>524</td>
</tr>
</tbody>
</table>

**Table 4.3:** MOPAC7 crimes recorded during the study period

**Figure 4.1:** Temporal distribution of crime in Islington (in minutes)
Vertical bars represent occurrences of crimes at the given times.

### 4.2.2 Officer movement data

Police forces within the UK routinely record the locations of their officers by utilising the GPS receivers within their body-worn radio systems. The exact frequency
with which these GPS ‘pings’ are recorded varies by police force, however the MPS have a passive refresh rate of five minutes. That is, when the radio is turned on but not being used by the officer to communicate with someone, the officer’s location (in the form of a longitude and latitude pair), call sign, and the current time are transmitted to a central server every five minutes and recorded and retained. The radio also pings the officers location each time they initiate a call in order to update their location which, given the time between pings, could be several hundred metres away from their last passively recorded location. The MPS provided officer GPS data covering the study period for all officer movements within the boroughs of Brent, Camden, Enfield, Hackney, Haringey, and Islington. Although Islington was the focus of the study, data was provided for the adjoining boroughs to allow for ‘spill-over’ patrols to be accounted for.

As the radios were not constantly measuring an officer’s location, the path they take between successive GPS pings needed to be interpolated. A range of methods could be used for this and here a relatively simple ‘join-the-dots’ approach was
4.2. Data

4.2. Data taken in the same way as the previous chapter. That is, officers are assumed to travel at constant speed in direct lines between consecutive GPS pings. The rationale for adopting this approach is two-fold: first, it was beyond the scope of this research given the time available to learn, implement, and test (with known ground-truths) a sophisticated map-matching algorithm which could realistically model foot-based patrols. Whilst companies such as Google, CityMapper, and Uber will have such algorithms, they are not freely available, nor are they tested against the officer-patrol use case. Instead, they are devised for vehicles and pedestrians, who more often than not are travelling for a specific reason - to get from one location to another - rather than the case of a police officer on patrol who is (or should be) trying to cover a given patrol area. Second, it was reasoned that it would be helpful if the method could be realistically replicated by a police force which is interested in performing their own evaluation. Whilst not a perfect match for the real paths that officers would have taken, absent any other data or a plausible method for ascertaining the real paths officers had historically taken, this method seemed most appropriate and one that a police force might be likely to implement in practice with the resources available to them. The implications of this assumption are explored in detail in Chapter 6.

4.2.3 Measuring police patrol dosage

In order to measure police patrol across the borough a grid of square cells was created to match the Call and Dispatch (CAD) grid cells used by the MPS. The CAD grid is used to deploy officers for a range of duties including foot patrols; locations of prospective hot-spots were provided to officers in the form of highlighted CAD cells. Each cell is a 250 metres by 250 metres square. For the purpose of this evaluation, patrol dosage has been estimated for each hour of the study period by computing the times that officers were in each cell. The general algorithm for computing this was as follows:

1. Select all GPS pings for an officer during the study period by filtering by the officer call signs within the data and order them by date and time.
2. Compare the times for each ordered pair of officer pings. If the time between them is greater than 15 minutes then ‘split’ the officer’s patrol into separate patrols at this point. This section of a patrol path is discounted as it was believed to be too inaccurate to adequately interpolate the officer’s path.

3. For each ordered pair of officer patrol pings determine the average speed at which the officer travelled between them. If the speed is greater than 2.5 metres per second then split the patrol into separate patrols.

4. For each ordered pair of officer pings determine whether both pings occur in the same CAD cell.

   (a) If both pings occur in the same cell, save the ‘sub-patrol’ path as a patrol in that cell from the time of the first ping to the time of the second ping.

   (b) If the pings cross several cells then assume constant speed between the two pings and ‘cut’ the assumed path at any points at which it crosses a cell boundary. Assign the exit and entry times for any cell by interpolating the time of the transition using the time between the first and second ping. Assign each ‘sub-patrol’ path to the relevant cell.

5. For each cell, take the computed sub-patrols and order them by their start time.

6. Iterate through each sub-patrol and compare it to subsequent sub-patrols. If the start time for subsequent patrols is earlier than the end time of the first patrol than combine them into one patrol by taking the new end time as the maximum end time of all the overlapping sub-patrols.

7. For each hour of the study period, estimate patrol dosage for each cell by intersecting the sub-patrols and the hour and summing the length of time for relevant sub-patrols. Where patrols span multiple hours, the start and end of the patrol are bounded by the start and end of the hour.

---

1It is unlikely the officer is conducting a foot patrol if they are proceeding substantially faster than the average pedestrian walking speed of 1-1.4m/s, particularly as police officers carry a considerable weight in the form of their protective vest and tools.
4.2. Data

GPS pings for each officer were joined to form an officer’s trajectory. The time at which they are assumed to cross from one cell into another is estimated by assuming they moved at constant speed and in a direct line between consecutive pings. When two officers were in the same cell at the same time the overlap was not double counted - that is, dosage is measured as the number of minutes in an hour when the box has officer presence, not the total number of officer-minutes spent in the cell. For example, if one officer entered a cell at 10:04 and left at 10:24 and another officer entered the same cell at 10:17 and left at 10:34 the dosage would be measured as 30 minutes not 37 minutes.

Variables were also computed to store the patrol history of each cell. For each hour and each cell the total patrol time in the previous hour, previous 8 hours, and previous 24 hours was computed. This was to allow the analysis to incorporate the history of the cell as the work of Sherman (1990) suggests that a ‘residual’ deterrent effect may exist. This was discussed in Section 2.2. It was hypothesised that police presence in an area might in fact increase the amount of recorded crime during the patrol. There are several reasons for this. Increased police visibility has been shown to increase confidence in the police (Bradford, Jackson, & Stanko, 2009) and thus people living in areas where police visibility is higher may be more likely to report a crime. Furthermore, the availability of a police officer makes the act of reporting a crime (even if the crime seems fairly minor) easier. This hypothesis is shared by Andresen and Hodgkinson (2018) though they were not able to evaluate it. Approximately 12% of crimes in London are recorded ‘in-the-street’ - that is, by a member of the public approaching a police officer in person or by a police officer witnessing an incident and recording it themselves (Heaphy, Davies, Lamb, & Mensah, 2014).

In order to account for patrolling on the boundary of the borough, a buffer of two CAD cells around the perimeter was included when computing dosage so that officers whose patrols crossed the borough boundary would still be properly accounted for. However, only cells that covered some area within the borough of Islington were used in the final analysis. This is shown in Figure 4.3 where the
thicker black outline shows the area of grid cells that were used for the models that follow - this area is called the observation window.

![Figure 4.3: Study area with observation window and grid](image)

Police resources are obviously limited, and in any given hour the amount of patrol dosage that any given cell might receive is likely to be small. Figure 4.4 shows that the vast majority of cells experienced no police presence in any given hour of the study period, whilst at the other extreme, some cells did experience a full sixty minutes of presence during certain hours. Some of these cells (though not all) encapsulate police stations. The locations of these cells are clearly highlighted on Figure 4.5 which shows the logged total dosage per cell over the entire study period and visually highlights the locations of two police stations in the northern and southern areas of the borough as well as one station near the south eastern edge of the borough. The path of higher dosages evident through the centre of the borough follows the course of the main arterial road through the borough.

### 4.3 Modelling framework

This initial evaluation uses a spatio-temporal point process modelling framework to estimate the effect of police patrol on crime. The analysis is conducted for the
Figure 4.4: Police patrol dosage within study cells
London borough of Islington over a three month period where proactive police patrols were a particular focus of the borough’s policing strategy. A series of different models are introduced with varying complexity based on the spatio-temporal two-component epidemic model described by Meyer et al. (2012) which was first developed to model the spread of infectious diseases. The models describe the intensity function which is composed of an endemic component and an epidemic component. In the context of the study of crime the endemic component can be described as the expected background rate at which crime would be expected to occur in the absence of any other influence. The epidemic component comprises a triggering kernel which estimates the increased likelihood of an event occurring because of a triggering event. As discussed in Section 2.1.1 this increased risk might be due to a boost mechanism that has been found to exist for burglary crimes in particular.
The general structure of the model is described below, followed by a simple version of the model which is then subsequently built on.

For a given observation period \((0, T]\) and an observation window \(W\), the cumulative intensity function (CIF) is denoted by \(\lambda(t, s)\) and represents the instantaneous risk of an event occurring at time \(t\) at location \(s\), accounting for all observed events up to that point. The CIF is constructed as the sum of an endemic component - \(v_{[s][t]}\) - which estimates the background risk of an event occurring, and an epidemic component \(e(t, s)\) which represents the influence that an event has on subsequent event likelihood:

\[
\lambda(t, s) = v_{[s][t]} + e(t, s) \quad (t > 0, s \in W)
\]

The endemic component has a multiplicative form \(v_{[s][t]} = \rho(s, t) \exp(\beta z(s, t))\) where \(\rho(s, t)\) is a known spatio-temporal intensity offset. In the models that are generated, this will be used to account for daily and weekly seasonality patterns. \(\beta z(s, t)\) is a linear predictor of endemic covariates and will be used to account for the patrol dosage in cell \(s\) at time \(t\). The index \([s][t]\) references a spatio-temporal grid spanning \(W\) and \((0, T]\) over which the covariates are measured. In this case, our window \(W\) is the grid of cells that cover Islington and \([s][t]\) references specific CAD cells as shown in Figure 4.3.

The epidemic component specifies an 'excitation' effect due to previous events within the set \(I(s, t)\) which consists of all events up to time \(t\) that are considered spatially and temporally proximate enough to influence the likelihood of an event occurring. That is:

\[
I(s, t) = \{j : t_j < t, t - t_j \leq \tau_j, |s - s_j| \leq \delta_j\}
\]

where \(\tau_j\) is the maximum time period over which one event is expected to influence the likelihood of a subsequent event (the temporal bandwidth),\(\delta_j\) is the maximum distance over which one event might influence the likelihood of a subsequent event (the spatial bandwidth. These parameters are specified by the researcher when gen-
erating the model estimates. In the results that follow the spatial bandwidth used is 500 metres and the temporal bandwidth (also known as the period of infectivity) is one week in keeping with the results of the Knox test outlined above and the contagion effect empirically derived from previous studies (Johnson & Bowers, 2004). Other parameter specifications were tested with minimal variation evident.

The epidemic component of the model is then specified as:

\[ e(t, s) = \sum_{j \in I(s, t)} \eta_j f(|| s - s_j ||) g(t - t_j) \]

where \( \eta_j \) is the force of excitation associated with an event and functions \( f \) and \( g \) model the decay of that excitation over space and time respectively. Details of possible decay functions are outlined later. The models that follow are estimated in R using the surveillance package developed by Höhle (2007) and expanded by Meyer, Held, and Höhle (2017). The models are generated by maximising a log-likelihood function to estimate model parameters. At first a series of models are generated that incorporate only endemic components. These are then expanded to incorporate epidemic components. Comparisons between models are conducted using Akaike’s Information Criterion (AIC) to assess the superiority of one model over another at accurately representing the underlying data.

### 4.3.1 Endemic-only models

Several endemic-only models were generated to find a parsimonious way of accounting for background risk at different times of day and on different days of the week. In essence, the desire was to maximise the accuracy of the model whilst minimising the number of coefficients which needed to be estimated. This was necessary as the more complex models required the estimation of a large number of parameters. Unless stated otherwise, to allow easier comparisons and references to the analysis conducted in Chapter 3, the results presented are for models of burglary offences only.

The first model that was calculated incorporated days of the week and hours of the day as categorical (or factor) coefficients, giving 30 parameters to model (an
As the 1st of January 2016 was a Friday, that was taken as the baseline against which other days of the week were compared. Similarly for hours of the day, midnight until 1am is the hour to which all other hours are compared. Unless otherwise stated, the time used to specify when a burglary took place is the earliest time recorded at which the crime could have happened - the crime start time in Table 4.1. The rationale for using this point in time is the same as given in Chapter 3 - to minimise the possibility that a crime was estimated to have occurred after a patrol when in fact it happened before. This approach guarantees that patrols will not be falsely recorded as having occurred before a crime, but does increase the risk that a crime is erroneously estimated as occurring before a patrol.

Figure 4.6 shows the estimated cumulative intensity function generated by the model for just one week. As the only parameters used are hours and days this will repeat for all weeks during the study period. This can be interpreted as the risk of a burglary occurring at a given time point over a one week period. The first week of the study period was used as a ‘lead-in’ period which is used to calibrate the potential source of events, hence the figure begins 1 week (10080 minutes) into our study period.

The results of this model are presented in Table 4.4 as rate ratios; i.e. a ratio of how likely a crime was to occur on a given day or hour compared to the baseline. For example, Tuesday appears to be the day of the week least likely to experience a burglary with the risk being approximately half the risk of a Friday. Compared to Fridays, burglaries are significantly less likely to occur on Mondays, Tuesdays, or Wednesdays.

As can be seen from the table, the confidence intervals for individual days and hours can vary significantly, with the last two hours of a day having a confidence interval for the rate ratios ranging from 0.22 to 1.12. This is due to the small number of events occurring across the study period and within some hours and some days of the week, and thus a robust estimate for the risk is difficult to generate. To reduce the number of parameters which need to be estimated a second model was gener-
ated. This time, the expected daily and weekly seasonality are not parameterised as individual hours and days but through the use of a superposition of sinusoidal waves of different frequencies. This approach has been previously demonstrated by Held and Paul (2012) and Meyer et al. (2017) who produced seasonal terms in this way to model the spread of influenza and invasive meningococcal disease (IMD) in Germany. The simplest wave that can incorporate a day’s ‘seasonality’ is a simple sine wave with a period of one day (1440 minutes) as shown in red in Figure 4.7. Similarly, a week’s seasonality can be incorporated using a wave with a period of 10080 minutes (the blue line). The black line in Figure 4.7 represents the sum of these two waves, to create a simple endemic model for the risk of crime over any given week. The model is generated using a maximum likelihood estimation.

By incorporating sinusoidal waves of different frequencies, a more complex (and hopefully accurate) depiction of risk can be built. For instance, instead of each day having just one ‘peak’ and one ‘trough’ we combine waves of different frequencies. Figure 4.8 shows the simple day model described above and a model
4.3. Modelling framework

Table 4.4: Endemic-only model of burglary risk - Categorical parameters

<table>
<thead>
<tr>
<th>Day/Time</th>
<th>RR</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>0.635</td>
<td>0.42–0.96</td>
<td>0.031</td>
</tr>
<tr>
<td>Tuesday</td>
<td>0.525</td>
<td>0.34–0.82</td>
<td>0.004</td>
</tr>
<tr>
<td>Wednesday</td>
<td>0.637</td>
<td>0.42–0.96</td>
<td>0.032</td>
</tr>
<tr>
<td>Thursday</td>
<td>0.716</td>
<td>0.48–1.07</td>
<td>0.100</td>
</tr>
<tr>
<td>Saturday</td>
<td>0.667</td>
<td>0.44–1.01</td>
<td>0.054</td>
</tr>
<tr>
<td>Sunday</td>
<td>0.827</td>
<td>0.57–1.21</td>
<td>0.330</td>
</tr>
<tr>
<td>01:00 to 02:00</td>
<td>0.556</td>
<td>0.26–1.21</td>
<td>0.140</td>
</tr>
<tr>
<td>02:00 to 03:00</td>
<td>0.444</td>
<td>0.19–1.03</td>
<td>0.057</td>
</tr>
<tr>
<td>03:00 to 04:00</td>
<td>0.056</td>
<td>0.01–0.43</td>
<td>0.006</td>
</tr>
<tr>
<td>04:00 to 05:00</td>
<td>0.111</td>
<td>0.03–0.48</td>
<td>0.003</td>
</tr>
<tr>
<td>05:00 to 06:00</td>
<td>0.277</td>
<td>0.10–0.75</td>
<td>0.012</td>
</tr>
<tr>
<td>06:00 to 07:00</td>
<td>0.333</td>
<td>0.13–0.85</td>
<td>0.021</td>
</tr>
<tr>
<td>07:00 to 08:00</td>
<td>1.108</td>
<td>0.58–2.12</td>
<td>0.760</td>
</tr>
<tr>
<td>08:00 to 09:00</td>
<td>1.775</td>
<td>0.99–3.17</td>
<td>0.053</td>
</tr>
<tr>
<td>09:00 to 10:00</td>
<td>0.835</td>
<td>0.42–1.66</td>
<td>0.610</td>
</tr>
<tr>
<td>10:00 to 11:00</td>
<td>1.001</td>
<td>0.52–1.93</td>
<td>1.000</td>
</tr>
<tr>
<td>11:00 to 12:00</td>
<td>1.389</td>
<td>0.76–2.55</td>
<td>0.290</td>
</tr>
<tr>
<td>12:00 to 13:00</td>
<td>1.054</td>
<td>0.55–2.01</td>
<td>0.870</td>
</tr>
<tr>
<td>13:00 to 14:00</td>
<td>0.779</td>
<td>0.39–1.57</td>
<td>0.480</td>
</tr>
<tr>
<td>14:00 to 15:00</td>
<td>0.999</td>
<td>0.52–1.94</td>
<td>1.000</td>
</tr>
<tr>
<td>15:00 to 16:00</td>
<td>0.833</td>
<td>0.42–1.65</td>
<td>0.600</td>
</tr>
<tr>
<td>16:00 to 17:00</td>
<td>0.776</td>
<td>0.38–1.57</td>
<td>0.480</td>
</tr>
<tr>
<td>17:00 to 18:00</td>
<td>0.832</td>
<td>0.42–1.66</td>
<td>0.600</td>
</tr>
<tr>
<td>18:00 to 19:00</td>
<td>0.886</td>
<td>0.45–1.75</td>
<td>0.730</td>
</tr>
<tr>
<td>19:00 to 20:00</td>
<td>0.666</td>
<td>0.32–1.39</td>
<td>0.280</td>
</tr>
<tr>
<td>20:00 to 21:00</td>
<td>0.500</td>
<td>0.22–1.12</td>
<td>0.090</td>
</tr>
<tr>
<td>21:00 to 22:00</td>
<td>0.278</td>
<td>0.10–0.75</td>
<td>0.012</td>
</tr>
<tr>
<td>22:00 to 23:00</td>
<td>0.500</td>
<td>0.22–1.12</td>
<td>0.091</td>
</tr>
<tr>
<td>23:00 to 00:00</td>
<td>0.500</td>
<td>0.22–1.12</td>
<td>0.092</td>
</tr>
</tbody>
</table>

that incorporates four sine waves (with 1 to 4 periods per day).

Models were generated incorporating a number of different sinusoidal waves to account for both daily and weekly seasonality. The details of every model are not included here, but Table 4.5 specifies the Akaike information criterion (AIC) for the different combinations of weekly and daily sinusoidal waves. A lower AIC is generally desired as it provides a measure of how well a model fits the data it is estimating; in this instance the AIC also accounts for the number of terms included in the model and penalises models with more terms. Table 4.5 shows that includ-
Figure 4.7: A simple endemic model incorporating seasonality
Weekly seasonality (blue line), daily seasonality (red line), and combined (black line)

ing four sinusoidal wave terms to model daily fluctuations and four wave terms to model weekly fluctuations provides a model which most accurately represents the underlying data. This compares to the model presented above where hours and days were categorical and had an AIC of 14846.

Table 4.5: AICs for endemic-only models incorporating sinusoidal waves

<table>
<thead>
<tr>
<th>Wave functions per week</th>
<th>Wave functions per day</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14818.21</td>
<td>14810.90</td>
<td>14795.85</td>
<td>14785.89</td>
<td>14789.05</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>14813.45</td>
<td>14806.13</td>
<td>14791.08</td>
<td>14781.12</td>
<td>14784.29</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>14816.33</td>
<td>14809.01</td>
<td>14793.96</td>
<td>14784.00</td>
<td>14787.16</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>14813.00</td>
<td>14805.69</td>
<td>14790.64</td>
<td>14780.68</td>
<td>14783.84</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>14815.24</td>
<td>14807.93</td>
<td>14792.87</td>
<td>14782.92</td>
<td>14786.08</td>
<td></td>
</tr>
</tbody>
</table>

Thus our 4-day-4-week wave model is a slight improvement of fit and the CIF for this model is shown in Figure 4.9, with daily seasonality (in red), weekly seasonality (in blue), and the combined CIF in black. Unless specified otherwise, all the models that follow incorporate this 4-4 wave model to constitute the endemic seasonality of burglary during our study period. The rate ratios for this model are given in Table
4.3. Modelling framework

Figure 4.8: Daily seasonality with one (above) and four (below) sine wave terms

4.6 for completeness, though they have no real interpretation individually. Going forward, these seasonal terms will be excluded from the results presented. Models which were generated for all MOPAC7 crimes combined yielded a similar outcome; a model composed of three sinusoidal waves to account for daily seasonality and three sinusoidal wave terms to account for weekly seasonality was marginally lower
Figure 4.9: An endemic model of burglary incorporating 4 daily and 4 weekly seasonal terms
Daily seasonality (red line); Weekly seasonality (blue line); combined seasonality (black line)

So far the models generated have not accounted for any potential dependency between crime events; instead they have only tried to estimate the background risk of burglary. Before modelling the self-excitation effect of crime events we incorporate patrol dosage into our underlying risk models. Initially, it was hoped that patrol dosage could also be modelled as a point process, with transitions between grid cells acting as events in the same way as crime-events. However, this proved computationally unfeasible due to the frequency with which officers traversed grid cells - over 700,000 ‘cell transition events’ would have needed to be incorporated, equating to approximately 5 officers per minute moving from one grid cell into another. Patrol dosage was thus aggregated up to hour blocks as described previously. One initial concern was that the risk of crime being recorded might actually increase with police presence due to the increased likelihood of a crime being detected or reported due to the officer presence. As such, dosage was also computed and modelled to account for recent patrol as well. Covariates were included within
Table 4.6: Rate ratios for an endemic model composed of 4 day and 4 week frequencies of seasonality

<table>
<thead>
<tr>
<th>Seasonality</th>
<th>RR</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>sin(2\pi * start/1440)</td>
<td>0.795</td>
<td>0.65–0.97</td>
<td>0.027</td>
</tr>
<tr>
<td>cos(2\pi * start/1440)</td>
<td>0.550</td>
<td>0.46–0.66</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>sin(4\pi * start/1440)</td>
<td>0.696</td>
<td>0.57–0.85</td>
<td>0.0004</td>
</tr>
<tr>
<td>cos(4\pi * start/1440)</td>
<td>1.191</td>
<td>1.01–1.41</td>
<td>0.043</td>
</tr>
<tr>
<td>sin(6\pi * start/1440)</td>
<td>1.061</td>
<td>0.89–1.27</td>
<td>0.51</td>
</tr>
<tr>
<td>cos(6\pi * start/1440)</td>
<td>1.593</td>
<td>1.32–1.93</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>sin(8\pi * start/1440)</td>
<td>1.216</td>
<td>1.03–1.44</td>
<td>0.022</td>
</tr>
<tr>
<td>cos(8\pi * start/1440)</td>
<td>1.296</td>
<td>1.08–1.55</td>
<td>0.0048</td>
</tr>
<tr>
<td>sin(2\pi * start/10080)</td>
<td>0.855</td>
<td>0.72–1.01</td>
<td>0.066</td>
</tr>
<tr>
<td>cos(2\pi * start/10080)</td>
<td>1.100</td>
<td>0.93–1.30</td>
<td>0.26</td>
</tr>
<tr>
<td>sin(4\pi * start/10080)</td>
<td>1.275</td>
<td>1.07–1.53</td>
<td>0.008</td>
</tr>
<tr>
<td>cos(4\pi * start/10080)</td>
<td>0.924</td>
<td>0.79–1.08</td>
<td>0.32</td>
</tr>
<tr>
<td>sin(6\pi * start/10080)</td>
<td>1.063</td>
<td>0.91–1.25</td>
<td>0.45</td>
</tr>
<tr>
<td>cos(6\pi * start/10080)</td>
<td>1.063</td>
<td>0.89–1.27</td>
<td>0.50</td>
</tr>
<tr>
<td>sin(8\pi * start/10080)</td>
<td>1.033</td>
<td>0.87–1.23</td>
<td>0.71</td>
</tr>
<tr>
<td>cos(8\pi * start/10080)</td>
<td>1.250</td>
<td>1.06–1.47</td>
<td>0.0072</td>
</tr>
</tbody>
</table>

the model which summed the dosage in the previous hour, the previous eight hours (to align with police shift times), and the previous 24 hours for each cell. The rationale for including these covariates stems from the work by Sherman (1990) looking at residual deterrence as discussed in Section 2.2. The AICs for these models are presented in Table 4.7 along with the no-dosage model detailed above. The AIC score is penalised for additional parameters however the endemic dosage models for burglary show no real difference to the model without dosage. AIC scores between models generated for different datasets (i.e. burglary and all MOPAC7 crime) should not be compared as the scores are dependent on the size of the data being modelled; hence we would expect the all-crime models to have substantially higher AICs than burglary-only models.

The rate ratios for models incorporating the current dosage and dosage over the past 24 hours are presented in Table 4.8 as this combination provides the better fit for the all-crime data, based on the AIC scores. The model estimates are also included as the rate ratios are very close to 1. This is to be expected even for a significant effect as it is estimating the effect of a single minute of patrol. The
Table 4.7: Fit of models incorporating patrol in endemic component

<table>
<thead>
<tr>
<th>Dosage variables included</th>
<th>AIC - burglary only models</th>
<th>AIC - all crime models</th>
</tr>
</thead>
<tbody>
<tr>
<td>No dosage</td>
<td>14780.68</td>
<td>114068.9</td>
</tr>
<tr>
<td>Current hour</td>
<td>14782.53</td>
<td>113970.3</td>
</tr>
<tr>
<td>Current hour, previous hour</td>
<td>14784.44</td>
<td>113928.8</td>
</tr>
<tr>
<td>Current hour, previous 8 hours</td>
<td>14782.11</td>
<td>113909.1</td>
</tr>
<tr>
<td>Current hour, previous 24 hours</td>
<td>14782.65</td>
<td>113902.8</td>
</tr>
</tbody>
</table>

The burglary model provides no evidence that police patrol dosage has a discernible impact on the underlying risk of crime as neither dosage covariate is statistically significant. However, the all-crime model does suggest that patrol dosage in the current hour significantly increases the likelihood that a crime is reported and patrol dosage over the previous 24 hours significantly suppresses the likelihood of a crime being reported.

Table 4.8: Rate ratios for an endemic model incorporating dosage

<table>
<thead>
<tr>
<th></th>
<th>Model Estimate</th>
<th>RR</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burglary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dosage - current hour</td>
<td>$1.203 \times 10^{-2}$</td>
<td>1.011</td>
<td>1.00–1.03</td>
<td>0.16</td>
</tr>
<tr>
<td>Dosage - previous 24 hours</td>
<td>$-5.977 \times 10^{-4}$</td>
<td>0.999</td>
<td>1.00–1.00</td>
<td>0.17</td>
</tr>
<tr>
<td>All crime</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dosage - current hour</td>
<td>$3.027 \times 10^{-2}$</td>
<td>1.030</td>
<td>1.03–1.03</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Dosage - previous 24 hours</td>
<td>$-9.515 \times 10^{-4}$</td>
<td>0.999</td>
<td>1.00–1.00</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

4.3.2 Epidemic models

The results presented in the previous section have focussed on models that only incorporate endemic variables and do not allow for spatio-temporal interaction between crime events. The endemic-only models are now expanded to incorporate epidemic covariates to allow for such interactions. As outlined in Section 4.3 this allows for an excitation effect, whereby the occurrence of a crime event can increase (or potentially decrease) the expected likelihood of a subsequent event. To recap,
this excitation term can be modelled as:

\[ e(t,s) = \sum_{j \in I(s,t)} \eta_j f(|s - s_j|) g(t - t_j) \]

where \( \eta_j \) is the force of excitation associated with an event and functions \( f \) and \( g \) model the decay of that excitation over space and time respectively. When modelling this excitation the spatial and temporal bandwidth parameters need to be specified. These provide upper-limits for whether one event can be triggered by another. One interpretation is that these bandwidths specify whether one crime might have been committed by the same offender and thus be counted as a repeat or near-repeat victimisation. Models were computed using spatial bandwidths ranging from 50 metres to 5000 metres and temporal bandwidths from five days to four weeks however the specification of these bandwidths had no effect on the results. As such, the results that are presented are based on computations using a spatial bandwidth of 500 metres and a temporal bandwidth of one week. In a purely epidemic model - that is, a model where every event must be caused by a preceding event - these thresholds may well be limitless.

The purpose of our models is to estimate the excitation force by determining suitable parameters for the spatial and temporal interaction functions \( f \) and \( g \). Several standard interaction functions are pre-defined within the surveillance package. Results from different models using these functions are given below. However, events which occur at the same time or place are not allowed within a continuous spatio-temporal point process model and so we must deal with any such ‘tied’ events first. Particularly in the case where a power-law kernel is used for the spatial interaction function, this may lead to model divergence\(^2\) (Meyer & Held, 2014). As such, events which occurred at exactly the same place or time as another event have been shifted. As recommended by Meyer et al. (2017), this shift was by a random amount of up to half the minimum non-zero spatial or temporal separation observed within the data - these were 4.47 metres and 3 minutes respectively. Breaking these

\(^2\)Models using Gaussian, step, power-law, and lagged power-law kernels were all tested and none converged when tied events existed.
ties within the data did not substantially alter the parameter estimates within the endemic models (see Table 4.9). However, AICs did increase as the models do not fit the data as well as previously.

**Table 4.9:** Parameter estimates for endemic-only models with and without tied data

<table>
<thead>
<tr>
<th>Wave Functions</th>
<th>Tied RR</th>
<th>Untied RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>sin(2π * start/1440)</td>
<td>0.795</td>
<td>0.792</td>
</tr>
<tr>
<td>cos(2π * start/1440)</td>
<td>0.550</td>
<td>0.552</td>
</tr>
<tr>
<td>sin(4π * start/1440)</td>
<td>0.696</td>
<td>0.691</td>
</tr>
<tr>
<td>cos(4π * start/1440)</td>
<td>1.191</td>
<td>1.197</td>
</tr>
<tr>
<td>sin(6π * start/1440)</td>
<td>1.061</td>
<td>1.053</td>
</tr>
<tr>
<td>cos(6π * start/1440)</td>
<td>1.593</td>
<td>1.602</td>
</tr>
<tr>
<td>sin(8π * start/1440)</td>
<td>1.216</td>
<td>1.208</td>
</tr>
<tr>
<td>cos(8π * start/1440)</td>
<td>1.296</td>
<td>1.300</td>
</tr>
<tr>
<td>sin(2π * start/10080)</td>
<td>0.855</td>
<td>0.855</td>
</tr>
<tr>
<td>cos(2π * start/10080)</td>
<td>1.100</td>
<td>1.106</td>
</tr>
<tr>
<td>sin(4π * start/10080)</td>
<td>1.275</td>
<td>1.274</td>
</tr>
<tr>
<td>cos(4π * start/10080)</td>
<td>0.924</td>
<td>0.930</td>
</tr>
<tr>
<td>sin(6π * start/10080)</td>
<td>1.063</td>
<td>1.063</td>
</tr>
<tr>
<td>cos(6π * start/10080)</td>
<td>1.063</td>
<td>1.069</td>
</tr>
<tr>
<td>sin(8π * start/10080)</td>
<td>1.033</td>
<td>1.033</td>
</tr>
<tr>
<td>cos(8π * start/10080)</td>
<td>1.260</td>
<td>1.260</td>
</tr>
<tr>
<td>AIC</td>
<td>14781</td>
<td>14824</td>
</tr>
</tbody>
</table>

The overall pattern of which models best fit the data did not change however, and the AICs for the various endemic models (for burglary), using untied data, are presented in Table 4.10. As such, the endemic-component model using four wave functions to model day-seasonality and four wave functions to model week-seasonality continued to be used. For brevity, the rate ratios for the endemic components of the models are not presented.

**Table 4.10:** AICs for endemic-only models incorporating sinusoidal waves - for untied data

<table>
<thead>
<tr>
<th>Wave functions per day</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave functions per week</td>
<td>1</td>
<td>14863.48</td>
<td>14858.08</td>
<td>14841.92</td>
<td>14829.12</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>14859.94</td>
<td>14853.54</td>
<td>14837.38</td>
<td>14824.58</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>14862.77</td>
<td>14856.37</td>
<td>14840.21</td>
<td>14827.41</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>14858.97</td>
<td>14852.58</td>
<td>14836.42</td>
<td>14823.61</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>14861.28</td>
<td>14854.88</td>
<td>14838.72</td>
<td>14825.92</td>
</tr>
</tbody>
</table>
4.3. **Modelling framework**

Endemic-epidemic models were then generated by updating the endemic-only model (excluding patrol dosage) to include epidemic components. The endemic-only model is used as a start point as recommended by Meyer et al. (2017) so that model estimation begins with reasonable start values. The epidemic models all use an exponential decay function for the temporal interaction function, where $\alpha$ is the parameter being estimated:

$$g(t) = e^{-\alpha t}$$

In keeping with the epidemiological origins of the models, the event is sometimes said to be ‘infectious’ during the period after the excitation and before the time at which the risk returns to baseline levels. Models are computed for several spatial interaction functions. These spatial interaction functions (SIAFs) are specified as follows:

**Constant:**

$$f(x) = 1$$

**Gaussian:**

$$f(x) = e^{-x^2/\sigma^2}$$

Where $\sigma$ is the standard deviation parameter being estimated.

**Power-law:**

$$f(x) = (x + \sigma)^{-d}$$

Where $\sigma$ is the standard deviation and $d$ the distance parameter being estimated.

**Lagged power-law:**

$$f(x) = \begin{cases} (x + \sigma)^{-d} & \text{for } x \geq \sigma \\ 1 & \text{otherwise} \end{cases}$$

The constant spatial interaction function (SIAF) produced the highest (i.e. least desired) AIC score (for burglary data - it was not run for the MOPAC7 data) and a smaller intercept parameter than the other models. The other SIAFs all produce similar results as can be seen in Table 4.11. These AICs are all preferable to the
endemic-only model though; recall that the AIC for the best endemic-only model was 14824.

Table 4.11: Parameter estimates for endemic-epidemic models

<table>
<thead>
<tr>
<th>Model</th>
<th>Constant</th>
<th>Gaussian</th>
<th>Power-law</th>
<th>Lagged power-law</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Burglary only models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>14792</td>
<td>14637</td>
<td>14637</td>
<td>14631</td>
</tr>
<tr>
<td>Intercept</td>
<td>$3.0 \times 10^{-10}$</td>
<td>$1.27 \times 10^{-6}$</td>
<td>$3.28 \times 10^{-6}$</td>
<td>$2.60 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$1.61 \times 10^{-3}$</td>
<td>$8.01 \times 10^{-4}$</td>
<td>$1.40 \times 10^{-3}$</td>
<td>$1.34 \times 10^{-3}$</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>-</td>
<td>0.837</td>
<td>-1.03</td>
<td>0.105</td>
</tr>
<tr>
<td>$d$</td>
<td>-</td>
<td>-</td>
<td>0.661</td>
<td>0.682</td>
</tr>
<tr>
<td><strong>MOPAC7 models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>-</td>
<td>111135</td>
<td>110898 †</td>
<td>110884 †</td>
</tr>
<tr>
<td>Intercept</td>
<td>-</td>
<td>$9.54 \times 10^{-6}$</td>
<td>$1.74 \times 10^{-7}$</td>
<td>$1.07 \times 10^{-5}$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-</td>
<td>$1.07 \times 10^{-4}$</td>
<td>$5.87 \times 10^{-5}$</td>
<td>$5.93 \times 10^{-5}$</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>-</td>
<td>-0.96</td>
<td>-5.05</td>
<td>-2.57</td>
</tr>
<tr>
<td>$d$</td>
<td>-</td>
<td>-</td>
<td>0.447</td>
<td>0.458</td>
</tr>
</tbody>
</table>

†- models did not converge.

These last three models, (using the Gaussian, power-law, and lagged-power law kernels) all show similar estimates for the parameter, $\alpha$ of the temporal interaction functions (TIAF) which decay rapidly over the first two days. Figure 4.10 shows the decay in minutes (there are 1440 minutes in a day) for the burglary-only models. Previous research has shown a similar temporal risk pattern for burglary, where the few days immediately after a burglary are much more likely to see repeat or near-repeat victimisations (Johnson & Bowers, 2004). Although the Gaussian model is a worse fit than the power-law models for the full MOPAC7 dataset, the power-law models did not converge.

However, the spatial interaction effect in the models generated here differ from previous research in one major way: whilst previous research has found that burglaries increase the risk of near-repeat victimisation up to a distance of approximately 400 metres (Johnson & Bowers, 2004), Figure 4.11 shows that in the current study the distance decay is nearly instantaneous - with the likelihood of a crime happening returning to baseline levels within only a few metres.

The distance of this spatial decay appears to be entirely an artefact of the necessary untying process. Figure 4.12 shows the impact of different untying distances
4.3. Modelling framework

(a) Temporal decay for endemic-epidemic model using Gaussian SIAF

(b) Temporal decay for endemic-epidemic model using power-law SIAF

(c) Temporal decay for endemic-epidemic model using lagged power-law SIAF

Figure 4.10: Temporal interaction functions for endemic-epidemic burglary models. 95% confidence bands are shown as dashed lines
on the spatial decay parameter, $\sigma$ for a Gaussian SIAF where $d$ is the distance between events.

One possible interpretation of this result is that events in such a densely populated environment only increase the likelihood of a repeat-victimisation, not a near-repeat victimisation. That is, the same offender may victimise the same location (either the same property or another property within the same building) but other proximate buildings are not at a higher risk than what is explained through the underlying risk (the endemic component of the model). As noted above, the constant spatial interaction function produces the worst fit for the data, showing that there is still a spatial component to the risk of contagion - i.e. a heightened risk of repeat victimisation. It may be that the very high population density of London means that theories which have previously been used to describe near-repeat victimisation - such as the Optimal Forager theory discussed in Section 2.1.1 - are less relevant due to the over-abundance of opportunity which effectively removes the need for the offender to forage at all. Accepting that victimisations that occur near previous events appear to be entirely accounted for by the underlying risk, there is still the question of whether patrol dosage has an effect on the risk of victimisation.

As with the endemic-only models presented earlier, the patrol dosage was then incorporated within the updated models. Dosage can be incorporated into the endemic component, the epidemic component, and both. A significant deterrent effect of dosage within the endemic component implies that patrolling reduces the likelihood of an initiating event from occurring whilst a significant deterrent effect of dosage within the epidemic component implies that patrolling deters repeat victimisation. Again, models were computed for each spatial kernel. However, as these have proven to show little difference in the case of burglary, and because only the Gaussian kernel model converged for the MOPAC7 models, only the results utilising a Gaussian kernel are presented here. Table 4.12 shows the estimated effect of patrol dosage when incorporated into both the endemic and epidemic components of the models\textsuperscript{3}.

\textsuperscript{3}Models that incorporated dosage into only the endemic component or only the epidemic component were also generated but did not vary from the models presented.
4.3. Modelling framework

(a) Spatial decay for endemic-epidemic model using Gaussian SIAF

(b) Spatial decay for endemic-epidemic model using power-law SIAF

(c) Spatial decay for endemic-epidemic model using lagged power-law SIAF

**Figure 4.11:** Gaussian, power-law, and lagged power-law SIAF plots. 95% confidence bands are shown as dashed lines. Distances are measured in metres.
Chapter 4. Modelling the deterrent effect of police patrol

Figure 4.12: Gaussian SIAF parameters for different untying distances

Table 4.12: Incorporating dosage into both endemic and epidemic components

<table>
<thead>
<tr>
<th></th>
<th>Model estimate</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burglary-only model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endemic dosage - current hour</td>
<td>$1.20 \times 10^{-2}$</td>
<td>0.190</td>
</tr>
<tr>
<td>Endemic dosage - previous 24 hour</td>
<td>$-1.68 \times 10^{-3}$</td>
<td>0.202</td>
</tr>
<tr>
<td>Epidemic dosage - current hour</td>
<td>$-1.99 \times 10^{-6}$</td>
<td>0.007 **</td>
</tr>
<tr>
<td>Epidemic dosage - previous 24 hour</td>
<td>$-5.08 \times 10^{-7}$</td>
<td>0.010 *</td>
</tr>
<tr>
<td>MOPAC7 model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endemic dosage - current hour</td>
<td>$2.91 \times 10^{-2}$</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>Endemic dosage - previous 24 hour</td>
<td>$-9.55 \times 10^{-4}$</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>Epidemic dosage - current hour</td>
<td>$-4.22 \times 10^{-8}$</td>
<td>0.589</td>
</tr>
<tr>
<td>Epidemic dosage - previous 24 hour</td>
<td>$7.36 \times 10^{-9}$</td>
<td>0.132</td>
</tr>
</tbody>
</table>

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

First, the burglary-only model provides insufficient evidence to conclude that police presence either increases or decreases the risk of an initiating crime event as neither estimate in the endemic component of the burglary model is significant. However, both current and recent patrol dosage is significant in the epidemic component. This suggests that police presence can suppress the likelihood of a repeat victimisation. Although previous research has not disaggregated the effect of police patrol on background rates of crime versus repeat or near-repeat victimisations, this overall reduction from police patrols would fit with previous research, particularly those implementations aimed at reducing repeat and near-repeat victimisations such as Fielding and Jones (2012) and Mohler et al. (2015).

However, the all-MOPAC7 model shows a different pattern. Unlike the burg-
glary analysis, the dosage estimates are now both significant in the endemic component and non-significant in the epidemic component. This is in keeping with the results of the endemic-only models outlined above. When considering the broader context of all MOPAC7 crime types, patrol dosage in a box appeared to significantly increase the likelihood that a crime would be reported within the same hour that the patrol took place. Conversely, recent patrol dosage (over the previous 24 hours) appears to suppress the likelihood of a crime.

Table 4.13: Comparison of AIC scores across models

<table>
<thead>
<tr>
<th>Model</th>
<th>Dosage covariates</th>
<th>AIC - burglary only</th>
<th>AIC - all MOPAC7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endemic only</td>
<td>None</td>
<td>14781</td>
<td>114069</td>
</tr>
<tr>
<td>Endemic only</td>
<td>Current hour, previous 24</td>
<td>14783</td>
<td>113903</td>
</tr>
<tr>
<td>Endemic-epidemic</td>
<td>None</td>
<td>14637</td>
<td>111135</td>
</tr>
<tr>
<td>Endemic-epidemic</td>
<td>Current hour, previous 24</td>
<td>14591</td>
<td>111098</td>
</tr>
</tbody>
</table>

Looking at the various models generated, we can see from Table 4.13 that the endemic-epidemic models which incorporate dosage covariates for both the current study hour and the previous 24 hours have lower AIC scores and are thus the superior fit for the data. Crime events do appear to have both spatial and temporal inter-dependence, and police patrol has a significant effect on the likelihood will be reported, though the direction and magnitude of that effect varies depending on the data being analysed.

4.4 Conclusion

This chapter has investigated the effect of police foot patrol on the risk of burglary offences as well as all MOPAC7 crime types. Whilst several of the findings are in agreement with previous research in this area, there are also disagreements. When only burglary data are included in the models, the findings show a heightened risk of subsequent crime occurring in an area in the days after a burglary, and the rate at which that risk returns to background rates over time is broadly in line with the literature (Johnson & Bowers, 2004). The findings are also in agreement with the literature in finding that offenders are likely to return to the same location to of-
fend again. However, in contrast to previous research, near-repeat victimisations appear to be entirely explained through the underlying risk and there was no spatial contagion effect evident beyond the original victimisation location. Another key difference when only using burglary data was the finding that police patrols do not appear to influence the background risk of victimisation, but do reduce the likelihood that a location will be revictimised. One potential explanation for this finding is that the hot-spot policing strategy that was being followed at the time was being implemented effectively. The strategy used a predictive crime mapping system - similar to ProMap and based on the work of Bowers, Johnson, and Pease (2004) - aimed at disrupting repeat and near-repeat victimisations. Thus, it may be that the lack of significant near-repeat victimisation within the study period is a result of successfully disrupting this behaviour, as intended. Unfortunately, it was not possible to test this as data were not available for a period when such a system was not in operation.

When looking at all MOPAC7 crime types the results were more in-line with expectations and previous research. However, the results from these models still did not produce a near-repeat victimisation effect. Another potential explanation for this finding is that the high levels of opportunity in London may mean that the ‘flag’ explanation of victimisation is much more powerful than the ‘boost’ explanation described by Johnson et al. (2009). Contrary to the analysis of burglary crime, these models did find that patrol within the previous 24 hours had a deterrent effect on crime as anticipated by Sherman (1990). Furthermore, the all-crime models suggest that police patrol caused a significant increase in the likelihood that a crime would be reported during the same hour as the patrol. This is an important new finding. Anecdotally, this possibility has been discussed in the literature (Andresen & Hodgkinson, 2018) but to the author’s knowledge, has not been previously investigated in a quantitative analysis. The analysis in this chapter also improved upon the temporal resolution of the analysis in Chapter 3 by using a temporal resolution of one hour rather than the police shift of eight hours.

There are several potential limitations to this analysis, particularly in regard to
4.4. Conclusion

the data that are used. First, the estimated dosage has been calculated from GPS data collected from officer-worn radios. These only transmit an officer’s location every 5 minutes and so their path through the urban environment has to be interpolated. This, along with inherent inaccuracies with GPS data may add substantial error to these findings that has not been properly accounted for. This will be further examined in Chapters 6 and 7.

Second, police recorded crime are subject to several challenges. These data only represent the crimes that are known to the police and as such provide only a partial picture of the actual risk and victimisation levels. Investigations by independent auditors have also found that crimes are not always recorded by the police when members of the public report them (Her Majesty’s Inspectorate of Constabulary and Fire & Rescue Services, 2018) and that the ‘true’ amount of crime is likely significantly higher than recorded by the police (Maguire & McVie, 2017, p. 169). Furthermore, decisions have been made during the analyses regarding where and when a crime is said to have occurred. For some crimes, such as burglary, the exact time at which the offence took place is usually unknown and a point over a period of time must be specified. Similarly, crimes such as robbery may not always be spatially well defined as the victim may not be able to identify exactly where they were when they were victimised. In this case, the location geocoded location of the crime has been taken as accurate without any further verification that it is. These challenges associated with accurately recording crime details are discussed further in Chapter 5.

Such data limitations are not specific to this study. Almost all previous research on the effect of police patrolling has (to a greater or lesser extent) been subject to the same data quality and boundary definition limitations. The results presented add to the literature in several ways: by adding to the currently sparse literature that analyses the effect of police patrols by using GPS data; by using a two-component self-exciting point process modelling approach to understanding police patrol; and by attempting to disaggregate the effect of police patrol on the background risk of crime and the likelihood of a repeat victimisation event. Furthermore, the eval-
ation conducted in the chapter has been done at a higher spatial and temporal resolution than other analyses on hot-spot police strategies. The higher resolution of the analysis conducted here is more appropriate than those that aggregate over weeks, months, or years and which assume that patrolling has an effect over large areas. The higher resolution of the analysis presented here is more sensitive to detecting any effects - some of which appear to be very short-lived. Such micro-level evaluations also encounter extra challenges. Intra-week and intra-day seasonality - a phenomenon often only considered at an annual resolution - occurs for many crime types such as burglary, assault, and robbery (Andresen & Malleson, 2015) and bicycle theft (Ashby & Bowers, 2013). Accounting for these patterns can be difficult at such resolutions, particularly given how rare crime events are at such spatial and temporal scales. This research has modelled both endemic and epidemic components within the analysis in an attempt to account for some of the unexplained variation.
Chapter 5

Investigating the accuracy of police recorded crime data

5.1 Introduction

This chapter is focussed on the spatial and temporal attributes of police recorded crime data. As discussed in Section 3.2.1, police patrols are being targeted at smaller areas (Andresen & Hodgkinson, 2018) and the time periods for which such locations are designated as ‘hot’ is becoming shorter (Bryant et al., 2015; Mohler et al., 2011; Mohler et al., 2015). The increased resolution of these interventions, in both space and time, raises the questions: how good are the data which are being used to determine these locations? And how do data quality issues impact on the estimated effects of such interventions?

The analyses conducted in Chapters 3 and 4 are reliant on recorded crime data in two key ways. First, as detailed in Section 3.3.1, the predictive crime mapping system that was being used by the MPS to deploy patrols used recorded crime data to determine which areas were most at risk of crime in the near future. Some predictive crime mapping systems do incorporate other data such as land use, weather, and socio-demographics of an area (For examples, see: Caplan et al., 2011; Robertson et al., 2017) but fundamentally they are all based on recorded crime data. The outputs from the system in place at the MPS in 2016 were then used to direct police patrols on a day to day basis. Second, recorded crime data are the dependent
variable in the models created in Chapter 4 when trying to assess the effect of patrol dosage. Hence, recorded crime data were critical to estimating the deterrent effect of police patrols. The challenges associated with accurately recording geospatial characteristics of crime have been discussed in Section 3.2.1 and will not be covered again here. The rest of this chapter outlines an investigation into the accuracy of the police recorded crime data used for the evaluation in Chapter 4. The next section outlines how the geocoding quality of the data were verified. This is followed by the results of the analysis and then implications for evaluations of hot-spot patrol interventions.

5.2 Data and Methods

The MPS maintain a records management system called the Case Overview and Preparation Application (COPA) system. It is intended as a centralised repository for all files and documents related to a crime. The COPA system is a collection of Crime Files, defined as the proper collation and recording of all paperwork relating to a crime or allegation of crime (Head of Records Management Team Met HQ, 2017). A case file can include a record of any emergency service calls, offence type, police dispatch logs related to the crime, witness statements, legal documents related to an offence, cross-referencing links to connect it to other MPS record management systems, etc. In essence, a Crime File would include any information which the MPS might be required to provide to the Crown Prosecution Service if the crime was ever charged. Importantly for this research, it is the digital repository for all witness statements (whether those statements were first written on a computer or handwritten and then scanned into the system and attached to the crime file).

The MPS provided access to the COPA system for all MOPAC 7 crime files related to the period and location of the evaluation in Chapter 4. To recall, the evaluation covered the London borough of Islington for the period of the 1st of January 2016 to the 31st of March 2016 - the study period. The crimes of interest were burglary, theft from a person, theft from a motor vehicle, theft of a motor vehicle, criminal damage, robbery, and violence against the person. A total of 2880 crimes
were recorded which fell within these parameters. In order to assess the accuracy of
the geocoding process, witness statements (where they exist) were analysed for each
crime within the study period by accessing the relevant Crime File on COPA. Where
a witness statement\(^1\) existed, it was read and any locational information regarding
the offence was extracted. The purpose was to compare the location reported by
witnesses (whether they be bystanders, offenders, victims, or attending police offi-
cers) with the geocoded location in the crime record. There were two objectives to
this analysis:

1. To check if multiple accounts of an offence all provided the same locational
   information - i.e. did all witnesses describe the location as the same place
   or was there any contradiction between recollections? If so, how was this
   reflected in the geocoding.

2. Quantify the accuracy of the geocoding process to see if the geocoded coordi-
   nates matched with the known locations of the offences.

First, recall from Section 4.2.1 that within the MPS every geocoded offence
is given an accuracy classification. The definitions of the different geocoding cate-
gories which were provided in Table 4.2 are given again in Table 5.1 and along with
the proportion of crimes that fell within each category for the 2763 crimes being
analysed in this chapter. This shows that approximately three quarters of all crimes
were coded to an exact address.

<table>
<thead>
<tr>
<th>Code</th>
<th>Definition</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Geocoded based on address data and address point data</td>
<td>75.4%</td>
</tr>
<tr>
<td>L2</td>
<td>Geocoded based on other location information such as road</td>
<td>1.4%</td>
</tr>
<tr>
<td></td>
<td>junction data and (train) station name</td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>Geocoded to the postcode level</td>
<td>18.9%</td>
</tr>
<tr>
<td>L4</td>
<td>Geocoded to the street level with Grid Reference buffer</td>
<td>3.3%</td>
</tr>
<tr>
<td>L5+NA</td>
<td>Precision level unknown or Uncoded</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

Of the 2763 crime files, 421 (approximately 16% of all crime files) had any

---
\(^1\)It is important to note that a police officer’s record of the event is also counted as a witness
statement even if the officer is not present during the time of the offence.
witness record associated with them in the COPA system. For those files which did contain witness reports, most reports were scanned in copies of hand-written reports, often directly from officer note books. This highlights the first challenge of utilising these data - searchability. In order to identify locational data (or any information for that matter), it was necessary to read each witness statement in detail; a time consuming process particularly when hand-writing was not always clear. Searching and extracting the required data took approximately three months of full time work despite there being only 441 crime files with usable records. The fact that so many crime files did not include any witness reports - including officer notes - was surprising as it was expected that at the very least an officer’s log of events would have been uploaded to the system. The fact that so few Crime Files had any witness information associated with them was met with some anxiety by the MPS officers overseeing the researcher’s access to the COPA system. Although the COPA system was believed to be the repository for data related to all recorded crimes, it may be that in practice it is used only for those crimes which may lead to prosecution. The proportion would seem to fit as approximately 19% of recorded crimes lead to a prosecution or caution in 1999-2000 (London UK: Home Office, 2001, p.113).

Witness statements related to 421 crimes within the study period included locational information regarding the event. This information was extracted and compared to the geocoded data for each crime event. As these data represent only a subset of the full study period crime data the results that follow are split into two categories: the first part discusses some key qualitative themes that emerged from reading the witness statements and extracting the locational data. The second part of the results provide some descriptive statistics of the geocoding accuracy, though caution must be taken with these findings given the low level of completeness within the dataset.
5.3 Results

5.3.1 Qualitative themes

The number of witness reports varied greatly, from zero (for the vast majority of the crime records) to more than a dozen associated with a single offence. However, several important themes emerged from the information within those records:

- **Addressable locations**: There were 78 offences for which the location of the crime was not the address entered into the crime record for one simple reason - the crime did not occur at an addressable location. As discussed in Section 3.2.3, an addressable location is a land parcel centroid or address point. Assaults which happened at bus stops or on buses (n=13) were often coded to a seemingly random location along the route of the bus, sometimes more than one kilometre from the actual location of the offence. Offences which occurred in public spaces such as pedestrianised areas or parks (n=8) appeared to be coded to a nearby landmark or building which at times could be at the opposite side of the park to which the offence occurred. At first it was assumed that these events might have been coded with an accuracy code other than L1 to specify that the geocoding was not exact, however many of them were in fact coded to the highest level of accuracy. Discussions with MPS officers suggested that the reason for this was that when completing the crime report and absent a true addressable space for the location of the crime event an officer would select any address which they believed to be nearby using a drop-down menu that was available to them. As this address was being selected from a menu of addressable spaces it was coded as completely accurate when geocoded. When the accuracy code was more realistic (usually L3) this was likely due to the recording officer ‘just writing in a road we knew was on the bus route’. Similarly, but less concerning, were events which happened on a street or street corner (n=37) and were recorded to a property close to which they happened - usually the nearest property on the street or a corner building if a crime occurred at an intersection.
• **Consistency of detail:** It was noticeable that for some crimes, where contradictions or slight anomalies might be expected, there was perfect alignment of details. For example, there were several occurrences where a fight had broken out on a residential street between several intoxicated individuals in the early hours of a Saturday or Sunday morning. It might be expected that the exact time and location of such an event would not be precisely known by all those involved, and yet all available witness statements would pin-point the time at which the event occurred to the exact minute and the location to in front of an exact residential address which was in no other way significant to the offenders or victims. It seems more likely that in fact a recording officer, when taking a statement from a relevant person would inform them of the time and location of the event as the officer understood it and that would be recorded in all witness statements. Hence, there would be no way to verify if these details were actually correct and it is impossible to address the first objective specified in the previous section.

• **Understanding crime classifications:** The MOPAC 7 crime types were originally selected for the evaluation in Chapter 4 because it was believed that each of these crime types required a potential offender to pass through a public area and thus they could be susceptible to a deterrent effect created through visible police patrols. However, through reading the details of the witness statements it became apparent that what was generally understood about the nature of the different crime types was inaccurate. Violence against the person is often characterised as a public offence - knife crime that occurs in the streets, altercations which occur in public spaces. Similarly, criminal damage conjures the notion of an offender attacking a victim’s car or a victim’s home. The fact that the MPS was generating predictive crime maps for these crime types - with the intent of using them to guide police patrol strategy - corroborates this belief. In fact, a significant proportion of the crime reports were related to events that occurred in private: many violent offences occurred between family members or individuals in shared accommodation;
criminal damage offences seemed to be mostly related to either disputes between cohabiting individuals where one damaged something belonging to the other, or to individuals who had already been put in police custody and who then caused damage to police property.

• **The geography of offences:** Whilst far less prevalent than the issues highlighted above, there were several crime files which included offences which did not require the offender and the victim to be in the same place at the same time. For example, threats of violence committed via telephone. These crimes pose a unique issue - did the offence occur at the victim’s location, the offender’s location, or is it hubris to assume the offence can or should be geocoded at all? The small number of occurrences in the current data were inconsistent with some geocoded to the victim’s address and other records geocoded to the offender’s location at the time of the offence. Neither is necessarily correct when shorn of context. Particularly in the case of threats of violence against a person this could take the form of multiple calls or text messages from the offender over a period of time where neither the offender or the victim is stationary.

### 5.3.2 Quantitative analysis

Table 5.2 provides the distribution of different crime types based on the automated geocoding categorisations. It also includes the proportions for each crime type based on the subsample with available witness statements. Burglary offences were more likely to be coded to the highest level of accuracy. This was expected given that burglary is inherently a property-based crime and thus is more likely to be geocoded to an addressable location. Similarly, criminal damage is a property-targeted crime and also has a high proportion of accurately geocoded offences. The fact that violent crimes were also estimated to be accurately geocoded is explained by the discovery, discussed above, of how many violent crimes were domestic in nature.

Table 5.3 presents the results of comparing the geocoded location to the loca-
Table 5.2: MOPAC7 geocoded classifications

<table>
<thead>
<tr>
<th>Accuracy code</th>
<th>MOPAC 7 crime type</th>
<th>Res. burglary (n=335)</th>
<th>Criminal damage (n=451)</th>
<th>Robbery (n=195)</th>
<th>TFMV (n=388)</th>
<th>TOMV (n=165)</th>
<th>Theft Person (n=680)</th>
<th>Violence (n=549)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>95.2%</td>
<td>80.9%</td>
<td>66.2%</td>
<td>71.7%</td>
<td>66.7%</td>
<td>63.8%</td>
<td>81.6%</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>0%</td>
<td>0.2%</td>
<td>2.6%</td>
<td>1.3%</td>
<td>1.2%</td>
<td>2.4%</td>
<td>1.8%</td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>4.5%</td>
<td>17.7%</td>
<td>25.1%</td>
<td>24.2%</td>
<td>30.3%</td>
<td>24.4%</td>
<td>12.4%</td>
<td></td>
</tr>
<tr>
<td>L4</td>
<td>0.3%</td>
<td>1.1%</td>
<td>4.1%</td>
<td>2.8%</td>
<td>1.2%</td>
<td>6.8%</td>
<td>3.5%</td>
<td></td>
</tr>
<tr>
<td>L5+NA</td>
<td>0%</td>
<td>0%</td>
<td>2.1%</td>
<td>0%</td>
<td>0.6%</td>
<td>2.7%</td>
<td>0.7%</td>
<td></td>
</tr>
</tbody>
</table>

With witness statements

<table>
<thead>
<tr>
<th>Accuracy code</th>
<th>MOPAC 7 crime type</th>
<th>Res. burglary (n=29)</th>
<th>Criminal damage (n=84)</th>
<th>Robbery (n=17)</th>
<th>TFMV (n=7)</th>
<th>TOMV (n=18)</th>
<th>Theft Person (n=18)</th>
<th>Violence (n=242)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>89.7%</td>
<td>85.7%</td>
<td>76.5%</td>
<td>57.1%</td>
<td>61.1%</td>
<td>72.2%</td>
<td>87.6%</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0.8%</td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>6.9%</td>
<td>14.3%</td>
<td>17.7%</td>
<td>42.9%</td>
<td>38.9%</td>
<td>22.2%</td>
<td>8.7%</td>
<td></td>
</tr>
<tr>
<td>L4</td>
<td>3.5%</td>
<td>0%</td>
<td>5.9%</td>
<td>0%</td>
<td>0%</td>
<td>5.6%</td>
<td>2.1%</td>
<td></td>
</tr>
<tr>
<td>L5+NA</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0.8%</td>
<td></td>
</tr>
</tbody>
</table>

The results should be interpreted with extreme caution given the low base sizes for most crimes. However, the number of violent crimes which were analysed is more robust and there is a clear trend for more accurately estimated geocoding to truly have been more accurate. Mean and median totals are given by crime type to highlight that the distribution of errors is highly skewed; the median error is in fact very low or zero for each crime type.

To summarise, the paucity of witness statements associated with the crime records related to the evaluation conducted in Chapter 4 meant that it was not possible to robustly analyse the geocoding accuracy of the crimes in a quantitative fashion. However, several important insights did emerge which highlight the challenges associated with geocoding crime events and potential data accuracy issues which should be at least acknowledged when conducting evaluations using geocoded crime data.
Table 5.3: Average MOPAC7 geocoding accuracy (in metres)

<table>
<thead>
<tr>
<th>Accuracy code</th>
<th>Res. burglary (n=30)</th>
<th>Criminal damage (n=84)</th>
<th>Robbery (n=19)</th>
<th>TFMV (n=7)</th>
<th>TOMV (n=18)</th>
<th>Theft Person (n=19)</th>
<th>Violence (n=242)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>7</td>
<td>128</td>
<td>12</td>
<td>250</td>
<td>435</td>
<td>55</td>
<td>16</td>
</tr>
<tr>
<td>L2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>50</td>
<td>54</td>
<td>2183</td>
<td>640</td>
<td>440</td>
<td>554</td>
<td>245</td>
</tr>
<tr>
<td>L4</td>
<td>1250</td>
<td>-</td>
<td>1900</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>472</td>
</tr>
<tr>
<td>L5+NA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total mean</strong></td>
<td><strong>118</strong></td>
<td><strong>106</strong></td>
<td><strong>506</strong></td>
<td><strong>417</strong></td>
<td><strong>437</strong></td>
<td><strong>163</strong></td>
<td><strong>96</strong></td>
</tr>
<tr>
<td><strong>Total median</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>0</strong></td>
<td><strong>60</strong></td>
<td><strong>10</strong></td>
<td><strong>28</strong></td>
<td><strong>0</strong></td>
</tr>
</tbody>
</table>

5.4 Conclusion

Valuable insights were gleaned from the opportunity to analyse witness statements where they existed. These findings suggest that researchers must at least acknowledge the potential for police recorded data to be inaccurately geocoded. A particular concern is highlighted by the fact that the geocoding accuracy measure - designed to give some appreciation for the confidence with which the geocoded coordinates represent the true location of the crime - may itself contain substantial errors. Whilst most crimes coded to this classification are clearly coded correctly (as indicated by the median error being zero for most crime types) some substantial outliers exist (as indicted by the mean error being over 100m in almost all cases).

It is also important to recognise that this analysis was conducted for a densely populated area within a major city. As discussed in Section 3.2.3, Cayo and Talbot (2003) found that errors were substantially greater in rural and suburban areas when compared to urban locations.

Historically, geocoding errors within police recorded crime data are unlikely to have affected the results of police patrol evaluations as the areas being studied have been substantially larger than any anticipated geocoding errors. However, with the increasing implementation of micro-hot-spot patrol strategies, geocoding errors may be leading to incorrect locations being designated as hot-spots, which in turn
may lead to sub-optimal patrol designations, and thus a lower estimated efficacy of police patrols to deter crime.

Furthermore, not all crimes that are reported to the police are accurately recorded (Her Majesty’s Inspectorate of Constabulary, 2014a). The most recent inspection by Her Majesty’s Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS) found that 89.5% of reported crimes were properly recorded, with 87.6% of violent crimes being recorded. This means about 28,100 violent crimes a year are not being correctly recorded (Her Majesty’s Inspectorate of Constabulary and Fire & Rescue Services, 2018).

The quantitative inferences which can be drawn from the analysis conducted in this chapter is limited by the data available. However, the effect that geocoding errors have on the designation and evaluation of micro-hot-spot police strategies is likely to vary by crime type, and the lack of completeness due to mis-recording crime (as well as unreported crime) will also have some effect. The evaluation conducted in Chapter 4 was originally planned to be conducted for several MOPAC 7 crime types. The reason for this was that the predictive crime mapping algorithms were being used to predict crime locations for each of these crime types and hence patrols could have been targeted at different locations depending on the crime type the borough was trying to counter on a given day. Based on discussions with several MPS officers, at the very least, these maps were being used to direct patrolling at hot-spots of burglary and violent crimes. Whilst patrols targeted at burglary hot-spots are likely to have been strategically valid, the findings in this chapter suggest that the violent crime hot-spots may have led to ineffective resource allocation. The MPS process of generating predictive crime maps by crime type is, in some instances, misguided.

Directing patrols to hot-spots of violent crime or criminal damage are not going to lead to effective patrol resource allocation if the input data relate substantially to crimes that occur in private spaces and which police patrols are incapable of preventing. A more sophisticated approach needs to be applied which does not assume all crimes of a given type are equally predictive of future crimes. In short, there
must be greater appreciation of the heterogeneous nature of such broad categories as ‘Violence against the Person’ and ‘Criminal Damage’ and micro-hot-spot maps, when used to determine police patrol strategies, need to ensure that they are based on only relevant crime events.
Chapter 6

Measuring the Unremitting Watch

The Inspector on Patrol Duty is to see as much of the Division as possible, and it is desirable that, when practicable, he should see every part of it once at least during his tour of Duty.

HMSO (1862, p.45)

6.1 Introduction

Robert Peel, in setting up the Metropolitan Police Force of London in 1829, envisaged the creation of ‘the unremitting watch’ of police officers across the city. This consisted of police constables whose duty it was to patrol individual beats. Constables were overseen by sergeants\(^1\) who would determine which constables were to patrol each beat during a given shift, and the sergeants were overseen by inspectors. It was expected that whilst on duty police constables would “see every part of his Beat in the time allotted; and this he will be expected to do regularly...” (HMSO, 1862, p.55). Measurement of these patrols was conducted simply by having sergeants check in at specific times and specific places with their constables to ensure they were in the correct beat. Sergeants would then have to write a report at the end of their shift for their inspector. Whilst such a system may have been perfectly adequate for the needs of the 19\(^{th}\) century, it is rarely suitable for modern evaluations of police patrols. Even now, how patrols are measured can vary greatly;

\(^1\)More accurately, they were called ‘serjeants’ which is a now obsolete spelling of the modern word sergeant. For the sake of readability the modern spelling will be used throughout this chapter.
but first an important distinction should be made between planned and actual activity. A failure to deliver what was planned in full or in part is a frequent problem in crime prevention practice (see Knutsson & Clarke, 2006), and means that activity realised in practice may differ substantially from what was intended. This has the potential to undermine studies which adopt an ‘intention-to-treat’ evaluation model (Novak et al., 2016, e.g.), whereby implementation activity is assumed but not measured and highlights the importance of directly measuring policing dosage. Three methods that have previously been utilised to measure dosage were mentioned in Section 2.2.4; using observers to record when officers enter and exit a hot-spot, analysing police logs, and the very recent development of using GPS data from officer-worn radios. GPS data have many advantages over the previous methods such as the passive nature and more frequent rate at which GPS location information can be collected. However, they are not perfect: data cannot be collected every second for practical reasons such as battery life and the infrastructure costs required to process, store, and analyse such volumes of data. Hence, the paths officers take between GPS pings need to be interpolated and the amount of time they spend in a given location must be estimated.

In their study, Ariel et al. (2016) were able to use 1-minute refresh rates. However, as mentioned in Section 2.2.4 this refresh rate is not typical. Discussions with three UK police forces (The MPS, West Yorkshire Police, and Thames Valley Police) suggest that operational ping rates are generally every two to five minutes. This is largely due to data collection costs and radio battery life considerations. Given the delays between the recording of foot-patrol locations (even if this is only one minute), to establish the paths taken between GPS pings requires interpolation. If employed as a micro-level measure of dosage, this can introduce errors into patrol evaluations (which will increase with the latency between GPS pings).

This chapter is motivated by a desire to quantify foot-patrol measurement errors and investigate how significantly they might impact on the measurement of police dosage in micro-places and thus evaluations that try to account for police dosage. With this in mind, it is important to consider the tools and expertise that
practitioners and researchers might have at their disposal. This analysis is in no way trying to improve on the sophisticated and proprietary algorithms used by companies such as Google, Microsoft, Uber, or CityMapper - companies that have all invested heavily in mapping systems which can take raw GPS tracking data and interpolate an individual’s path through the urban network, determine their likely mode of transport, and account for other factors such as traffic conditions, environmental factors, and potentially more accurate path data from other users of their services. These companies also have a broader range of tools with which to measure movement. These include: multilateration (measuring the time that energy waves sent from a mobile phone take to reach different network towers in the area); mobile phone signal strength measured from several network cell towers; crowdsourced WiFi data from nearby receivers; and a much more extensive dataset of pedestrian movement.

This thesis is interested in how a police analyst or academic researcher might measure patrol dosage in a given area using GPS data from officer-worn radios. The number of studies that have already utilised GPS data in the measurement of police dosage is small as the technology is relatively new in law enforcement and thus evaluations have only recently been able to utilise GPS data to measure patrols. Some have counted the number of GPS pings within patrol locations and then multiplied that by the (standard) time between pings. For example, an experiment by Ariel et al. (2016) used GPS data where the location was recorded every two minutes. If, for instance, three pings occurred within a patrol box, they would count this as six minutes of dosage. Alternatively, a ‘join-the-dots’ approach has been used, whereby an officer’s path was assumed to be a straight line between GPS pings and that they moved at a constant speed between pings. When one ping falls outside a patrol location and the next ping falls inside the patrol location, the entry time can be calculated accordingly (Hutt et al., 2018).

For both strategies, the frequency with which pings occur, and where along an officer’s path they occur can have a significant impact on where dosage is assigned. This chapter begins by first describing a pilot experiment conducted with
West Yorkshire Police. The purpose of this experiment was to test whether GPS data could accurately describe where officers had been by comparing it to a known ground truth established by actually walking on patrol with several officers. The second part of the chapter expands on this by collecting more frequent GPS data from approximately 30 foot patrol officers over a two month period.

6.2 Testing the ground truth of GPS patrol data

Early in the course of this research, the author conducted a pilot study in a UK city\(^2\) in conjunction with West Yorkshire Police (WYP) to assess the accuracy of GPS data collected via officer-worn radios. There were two objectives to this experiment: first, did the officer GPS data accurately reflect where the officer actually was; and second, how accurately could officer paths be interpolated from their GPS data. The author accompanied officers on foot patrols in two areas, recording the author’s location every ten seconds using a smartphone-based GPS recording application. This, as well as notes made during the patrols and the author’s recollection of the path, were used to determine the true paths of the officers. GPS data collected from officer radios every two minutes was then provided by WYP for the officers that had been accompanied. The first patrol was completed with two patrol officers between 4:15pm and 6:15pm on the 9\(^{th}\) of December 2015 in a residential neighbourhood. The second patrol was completed with one patrolling officer between 1:30pm and 2:50pm on the 10\(^{th}\) of December 2015 in the city centre. Each officer ping was matched to the author ping that occurred nearest in time so that the spatial distance could be measured. A total of 41 matched pairs were recorded in the residential area and 30 matched pairs recorded in the city centre with median distances between matched pings of 15.3 metres and 18.4 metres respectively. The distribution of these distances are given in Figure 6.1 and has a long tail such that whilst the residential area had three significantly erroneous pairs, most ping pairs were within 20 metres of each other. To account for the fact that the author and officers were moving, only pairs of pings that occurred within 15 seconds of each other are included, with

\(^2\)The city has a population of approximately half a million and covers an area of approximately 25 square miles.
6.2. Testing the ground truth of GPS patrol data

the median time between matched pings being 7 seconds for both groups. Thus, given the average pedestrian walking speed of 1-1.4 metres per second (Levine & Norenzayan, 1999; Snaterse, Ton, Kuo, & Donelan, 2011) and that the GPS receivers were not being held by the same individual it should be expected that there is some distance between the pings.

These results obviously do not assess the measurement error of the GPS receivers used by the officers, but do provide some reassurance that the data is not wildly inaccurate, (at least in the majority of cases) and can broadly represent the paths officers have taken. Maps of the officers’ paths are provided in Figure 6.2 and the residential area map demonstrates that for the majority of the time, the officers’ paths are clear. The only divergence from the actual path is visible at the bottom of the first panel where one GPS ping for the officer represented in red exhibits a significant measurement error and does not follow the same (and true) path as the officer (the blue line). Given the time period between pings and the structure of the network, it is not challenging to ascertain the officer’s paths with a high degree of

Figure 6.1: Distances between officer and author matched GPS pings

These results obviously do not assess the measurement error of the GPS receivers used by the officers, but do provide some reassurance that the data is not wildly inaccurate, (at least in the majority of cases) and can broadly represent the paths officers have taken. Maps of the officers’ paths are provided in Figure 6.2 and the residential area map demonstrates that for the majority of the time, the officers’ paths are clear. The only divergence from the actual path is visible at the bottom of the first panel where one GPS ping for the officer represented in red exhibits a significant measurement error and does not follow the same (and true) path as the officer (the blue line). Given the time period between pings and the structure of the network, it is not challenging to ascertain the officer’s paths with a high degree of
However, in the town centre the true path is less obvious and there are several factors which contribute to this ambiguity: first, the road network is more dense and so it is less clear which path an officer might have taken when there are several in close proximity. Second, GPS measurement errors are compounded by urban canyon effects (see Section 2.2.4) caused by tall buildings which make it more likely that an officer’s GPS location will be situated on a road (or within a building) they did not visit. He et al. (2011) highlight that, “this bias is heavily dependent on the environment surrounding the receiver antenna, but with little relevance in space, so it is difficult to eliminate through differential method[s].” and trying to correct for such errors is complex. There exist several complex map-matching algorithms (algorithms designed to convert GPS data into digitised complete paths which map to a given network) that might, in other circumstances, be able to help mitigate some of these errors. However, as both maps in Figure 6.2 demonstrate, there are times when officers are not traversing known roads or paths. On the residential map, there are two clear sections where the officers ‘jump’ between road sections via paths that do not exist on the map. Foot patrols, unlike vehicle patrols, are not constrained by the ‘official’ paths that exists.

A similar pattern was discovered whilst conducting patrols with officers in London. Figure 6.3 shows the path taken by an officer who was accompanied on a patrol in an inner-London borough so that the true path taken could be compared to the interpolated path from their GPS data. There are several issues to bear in mind. First, there are several occasions when the officers path diverges from roads and pavements as they make their way through housing estates and other non-standard walkways. As such, attempting to ‘correct’ GPS paths by matching them to known road and path structures would in fact add errors and any attempt to mitigate urban canyon effects or innate GPS measurement error are beyond the scope of this research. Second, there occasions where the GPS data is either inaccurate (several pings towards the north of the map are clearly erroneous) or do not get recorded (the short segment of the true path with no apparent GPS data).
6.2. Testing the ground truth of GPS patrol data

Figure 6.2: Officer GPS-based paths

(a) 9th December - Two officers (blue and red paths) patrol a residential area.

(b) 10th December - One officer (black path) patrols a city centre.
Another consideration is how accurately the ping rate allows for meandering paths to be captured. Taking the example of the city centre patrol from Figure 6.2 again, it would appear that when the officer passed along the north eastern edge of the park in the centre of the map they turned right and followed a fairly straight path through the city centre streets. However, this was not the case. Figure 6.4 shows that the officer actually walked three sides of a block; a fact more easily observable using the more frequent GPS data of the observer, though again, an urban canyon effect is clearly visible.

The rest of this chapter describes an experiment conducted in London, UK that was designed to measure how the path interpolated from GPS data differs when using data collected at different rates. Whilst it was not feasible to systematically record the true path that officers took, the purpose of the experiment was to evaluate whether different GPS ping rates would lead to significantly different patrol paths being assumed.
6.3 Data

In 2017, approximately 40 police officers from the MPS agreed to wear a secondary radio whilst out on foot patrol. These radios transmitted the officer’s location every 30 seconds - a significantly faster refresh rate than the 5 minute interval used for standard MPS body-worn radios. The purpose of this more frequent refresh rate was to allow for foot patrol paths to be measured at a higher resolution. Due in part to radio malfunction, a total of 31 officers ultimately participated in the experiment and recorded 214342 ‘location pings’ using the secondary radios. This equates to approximately 1786 hours of recording or 223 eight hour police shifts ranging from the 8\textsuperscript{th} November 2017 to the 16\textsuperscript{th} January 2018. Figure 6.5 shows the distribution.

\footnote{This experiment is focussed entirely on foot patrols, and as such any reference to GPS pings from radios refers to body-worn radios and not vehicle-based radios which do have a significantly faster refresh rate.}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{foot_patrol_paths.png}
\caption{Officer, observer, and the true path of a patrol}
\end{figure}
of the data over the study period and that officers were less likely to carry their secondary radio (or potentially that the radios developed technical issues), particularly after the Christmas break.

**Figure 6.5:** GPS ping distribution during trial period

![Figure 6.5](image)

However, these data were not just recorded when officers were conducting patrols. As such, the data have been cleaned so that only data relating to actual foot patrols were used for the analysis; removing, for instance, time spent in police stations or travelling either in police vehicles or on public transport. Whether an officer’s movements were on foot or by vehicle was determined by calculating the average speed between GPS pings and discounting any travel at more than two metres per second. As discussed above, the standard pedestrian walking speed is 1 to 1.4 metres per second. Given that police officers are carrying a considerable extra weight in the form of their vests and equipment it is assumed they will not be walking any faster than a standard pedestrian when they are on foot patrol. This also had the effect of removing some of the more extreme cases of the urban canyon
6.4 Method

Officer paths were interpolated based on the 30 second refresh rate GPS data, to create a baseline ‘assumed path’ (which, as discussed above, is known to not perfectly match the ‘true path’ but provides an approximation of it). Subsets of the GPS data are then created such that paths are interpolated based on 60, 90, 120 (and so on) second intervals up to five minute intervals. There are many ways of measuring the similarity of two paths. Magdy, Sakr, Mostafa, and El-Bahnasy (2015) provide a useful review, separating the methods into Spatial and Spatio-temporal similarity measures. The spatio-temporal methods can immediately be discounted for the present analysis. The paths we wish to compare (30 second ping rates versus more sparse ping rates) are derived from the same dataset and thus any temporal analysis is not sensible. Of the spatial similarity measures, the Fréchet metric (or distance) is the most popular similarity measure in use according to Gudmundsson, Laube, and Wolle (2011) and its use here is also appropriate.

The Fréchet distance can be described as follows. Assume we have two paths, \( A \) and \( B \). At the start of path \( A \) there is a dog and at the start of path \( B \) the dog’s owner. Both the owner and the dog can walk along their respective paths and they can each vary their speed, though they can not move backwards. The Fréchet distance is the minimum length of leash that would be necessary to connect the dog and its owner for the entire journey from the start to the end of their respective paths. Alternatively, consider the path \( A \) as being made up of infinitely many points. For each point, calculate the shortest distance to path \( B \). The Fréchet distance is the maximum of all these shortest distances. It is important to note that for our data, the distances are not only calculated where the GPS pings occur, but along the entire path. These paths are also often referred to as trajectories.

As one of the trajectories being compared in this analysis is created from a subset of points from the other, there will of course be regular intervals where the
trajectories ‘touch’. However, this is not an issue as the Fréchet distance is the minimum leash length over the entire trajectory. That said, in measuring the similarity of two trajectories, using the Fréchet distance over the full patrol would provide very little useful information as it would only provide us with the maximum distance between the trajectories. The analysis would be highly susceptible to noise within the data such as erroneous pings by urban canyon effects.

Instead of computing the Fréchet distance for the entire path, matched sections of the trajectories are sampled and the Fréchet distance is calculated for each section. This provides a distribution for the similarity between the two trajectories. As a basic example, Figure 6.6 shows a set of points which might represent a (very meandering) patrol officer as they head (approximately) east from the point (0,3). Three trajectories have been created, $T_1$: based on taking every ‘ping’ (the ‘assumed path’ shown as a solid black line), $T_2$: every other ping (dashed blue line), and $T_3$: every third ping (dashed red line). Note that $T_2$ has an alternative specification - rather than being composed of every odd-numbered ping it could be every even numbered ping. Similarly, $T_3$ has two alternative specifications; the path constructed by taking pings 2,5,8, and 11 and the path constructed by taking 3,6,9, and 12.

The Fréchet distance between $T_1$ and $T_2$, when the full trajectories are taken into account, is exactly four. Whilst the Fréchet distance between $T_1$ and $T_3$ is 2.2. Rather than use just this single number, the Fréchet distance is computed for matched subsections of the trajectories. For instance, comparing $T_1$ and $T_2$ again, the Fréchet distance is computed for the Easting intervals of [0,2],[1,3],[2,4],[3,5] etc. This provides a distribution of how similar the trajectories are. Given the very simple (and low sample count) example provided, the distributions for this example are given as boxplots in Figure 6.7.

In order for this analysis to be conducted accurately, the data must include pings every 30 seconds. As such a second phase of data cleaning was required. Data for each officer were split into ‘sub-patrols’ by selecting pings that occurred within 40 seconds of each other and where the maximum distance between two pings was
Figure 6.6: Example trajectories for Fréchet distance measurement

Figure 6.7: Distributions of Fréchet distances for example trajectories
150 metres. A threshold of 40 seconds was allowed as the radios did not always ping exactly every 30 seconds, perhaps due to delays in the data being received. Any sub-patrol with only one or two GPS pings was removed and after inspection of the data, any sub-patrol with more than 60 GPS pings (i.e. more than 15 minutes of persistent 30-second pings) were also removed. This produced 22951 ‘sub-patrols’. Removing 1 and 2-ping patrols reduced this to 11483 patrols and removing the 15 minute patrols reduced this to 10925. These long un-interrupted patrols were found to be due to a radio being switched on and left in one place for an extended period of time or because an officer was stationary. For instance, if an officer was at a hospital or school. The purpose of the data collection was to measure actual patrol movements and so removing these ‘stationary periods’ is not a concern. Similarly, if during a sub-patrol an officer did not move on average 0.3 metres per second the patrol was discarded due to the officer being mainly stationary. The final number of sub-patrols used in the analysis was 6064. The median number of GPS pings in a sub-patrol was 8.98 and the inter-quartile range was 4 to 11.

6.5 Results

The Fréchet distances were calculated between officer patrol paths using 30 second GPS ping rates (the officer’s assumed path) and more sparse data sampled from this same dataset. The distributions of these distances are shown in Figure 6.8 The similarity of the patrol paths reduces as the assumed path is interpolated based on sparser data. Although there is a long tail to the distribution, a sixty second ping rate provides a very similar path to that produced by a 30 second ping rate. Table 6.1 provides some basic descriptive statistics of the distributions. The median Fréchet distance when comparing a path using 30 second ping rates to a path using 5 minute ping rates was 60.1 metres and again, there is a long tail to the distribution.

6.6 Conclusion

An assumed officer patrol path was interpolated using GPS data collected with a 30 second refresh rate. The similarity of this path was then compared to paths based on increasingly more sparse data to evaluate how the frequency of data collection
Figure 6.8: Similarity distributions of police patrol paths using sparse GPS data

### Table 6.1: Distribution of Fréchet distances by ping frequency

<table>
<thead>
<tr>
<th>Ping frequency (seconds)</th>
<th>Frechet distance (metres)</th>
<th>1st quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>0</td>
<td>0</td>
<td>6.81</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>0</td>
<td>9.32</td>
<td>21.67</td>
<td>34.25</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>10.47</td>
<td>27.85</td>
<td>34.94</td>
<td>51.97</td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>17.09</td>
<td>35.80</td>
<td>41.81</td>
<td>60.54</td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>23.21</td>
<td>42.80</td>
<td>47.85</td>
<td>67.72</td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>27.31</td>
<td>47.85</td>
<td>52.64</td>
<td>73.75</td>
<td></td>
</tr>
<tr>
<td>240</td>
<td>31.24</td>
<td>52.68</td>
<td>57.24</td>
<td>79.16</td>
<td></td>
</tr>
<tr>
<td>270</td>
<td>34.74</td>
<td>56.65</td>
<td>61.21</td>
<td>83.49</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>38.00</td>
<td>60.07</td>
<td>64.92</td>
<td>87.78</td>
<td></td>
</tr>
</tbody>
</table>
effects the path an officer is assumed to have taken. Police forces are known to use GPS refresh rates of between two and five minutes. Based on the results of this study, they would produce assumed paths which, on average, deviate from these more regularly measured paths by a median distance of 28m and 60 metres respectively. Whilst this distance may seem relatively small it is only an average path deviation and when the sum total of patrol officer paths is considered it may lead to substantially different estimates of patrol dosage. This is particularly the case when analysis is conducted at small spatial resolutions as if often the case with hot-spot policing strategies.

Several limitations exist within the analysis however. The true paths walked by the officers is unknown. The GPS data still only approximates the officers’ locations; their path between pings has been interpolated and measurement error still exists with in the collection of GPS data. Urban canyon effects can significantly distort the estimated location of an officer. An attempt was made to mitigate some of these issues by removing ‘noisy’ data - where the distance between pings was unrealistically far or the speed travelled between pings too great for the movement to have been by foot.

Still, to the author’s knowledge, this is the first study to examine the similarity of police patrol paths using GPS data. The use of GPS data to estimate police patrol dosage is a relatively new development. So far, very few studies (Ariel et al., 2016; Hutt et al., 2018; Williams & Coupe, 2017) have used GPS data to try and measure the deterrent effect of police patrols. None, other than the paper by this author, have acknowledged the potential error associated with these data. However, as patrols have become more targeted the impact of the accuracy of the measurement of police patrols requires greater consideration.

The findings of this Chapter directly lead into the next, where the evaluation conducted in Chapter 4 is re-examined in light of the expected error in the measurement of patrol dosage.
Chapter 7

The impact of data error on evaluating police patrol

7.1 Introduction

Chapters 5 and 6 investigated the issues of data quality in police recorded crime data and the GPS data used in the measurement of police foot patrol respectively. Whilst these issues are important to understand in their own right, the ultimate objective of these analyses was to apply the findings to the evaluation conducted in Chapter 4 in order to better understand how they might affect the estimates generated. Specifically, the intention was to test whether data error might be sufficient to change the results of an evaluation of police patrolling. If it does, this has significant implications for the reliability of the evaluative evidence-base relating to such a crime prevention intervention. The analysis of police recorded crime data conducted in Chapter 5 did not produce robust quantitative results as had originally been hoped. The data available represented only a small proportion of all crimes and thus it cannot be assumed to be representative of the recorded crime data used in Chapter 4. Hence, this chapter does not attempt to account for the geocoding error of recorded crime within the analysis that follows. Instead, the findings from Chapter 4 are re-examined with a focus only on the measurement of police patrol and how that affects the outcome of evaluations of patrol. Whilst only accounting for the potential influence of one type of data error, if outcome variation is observed
this would be sufficient for demonstrating the impact data inaccuracy can have on the evidence base. In order to represent possible data error, in the analysis that follows, rather than interpolate the officer patrol paths as being a straight line between GPS pings, the Fréchet distance distributions calculated in Chapter 6 are used to simulate other paths that officers may have taken.

### 7.2 Data and methods

The analysis conducted in Chapter 4 used police recorded crime data for the London Borough of Islington and officer location data in the form of GPS pings, collected by the MPS from officer-worn radios. That analysis was conducted using the full borough as the study area and using three months of patrol data. In the analysis that follows however, due to the computational intensity of calculating the office patrol dosage per cell it was necessary to specify a smaller area and time period.\(^1\) The three most northern wards of Islington (called Hillrise, Junction, and Tollington) were selected as the study area and as before, a 250 metre by 250 metre grid was create to match the CAD grid used to assign patrols. Figure 7.1 shows the study area and associated observation window used for the analysis. As before, a buffer of cells is included around the study area when calculating the estimated dosage in each cell but only dosages within the observation window are used when computing the models.

The study period begins on 1\(^{st}\) January as before, however the length of the study was shortened to one month; ending on the 31\(^{st}\) January 2016\(^2\). A total of 138 MOPAC7 crimes occurred during the study period and Table 7.1 provides a breakdown by crime type and includes the breakdown for the initial evaluation conducted in Chapter 4 for comparison. Due to the smaller study area and time period, models were only generated for the combined MOPAC7 crime data.

---

\(^1\)The analysis was at first attempted using the full study area and time period as in Chapter 4, however the computation requirements for this were greater than what was available to the researcher. After several attempts to conduct the analysis ‘piece-meal’ the decision was taken to reduce the study area and time period.

\(^2\)This allowed the re-estimated patrol dosage and subsequent hourly patrol dosage per CAD cell to be computed in approximately one hour. When the computation was originally attempted using the entire study area and study period used in Chapter 4 the computer would crash after 12 hours,
7.2. Data and methods

**Figure 7.1:** Study area with associated observation window and grid

**Table 7.1:** MOPAC7 crime counts - initial evaluation and re-evaluation

<table>
<thead>
<tr>
<th>Crime type</th>
<th>Initial study</th>
<th>Re-evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burglary</td>
<td>324</td>
<td>31</td>
</tr>
<tr>
<td>Criminal damage</td>
<td>443</td>
<td>19</td>
</tr>
<tr>
<td>Robbery</td>
<td>190</td>
<td>9</td>
</tr>
<tr>
<td>Theft from a person</td>
<td>632</td>
<td>23</td>
</tr>
<tr>
<td>Theft from a motor vehicle</td>
<td>379</td>
<td>32</td>
</tr>
<tr>
<td>Theft of a motor vehicle</td>
<td>161</td>
<td>7</td>
</tr>
<tr>
<td>Violence against the person</td>
<td>524</td>
<td>17</td>
</tr>
</tbody>
</table>

Other than the reduced size, the crime data remains the same. However, the method by which patrol dosage was estimated has been changed in order to estimate other potential patrol paths.

### 7.2.1 Re-estimating patrol dosage

The results presented in Chapter 6 suggest that, were the GPS ping rate to be 30 seconds rather than five minutes, the assumed paths that officers take could be substantially different. Based on the computed Fréchet distances, the median ‘perturbation’ from the straight line between 5-minute interval pings was 60 metres. The distribution of Fréchet distances which compared a 30-second ping rate to a 5-minute ping rate provides a measure of similarity between the two estimates of

_having not yet completed the computation_
the officers’ paths. This ‘similarity distribution’ was used to simulate potential patrol paths for the officers in the Islington evaluation. For every pair of consecutive GPS pings \((p_i, p_{i+1})\) that form an officer’s path, an intervening mid-point, \(p_k\), was created. This point was then moved from the direct line connecting the two pings. It was moved perpendicular to the line and the distance it was moved was selected randomly from the Fréchet distance distribution calculated in Chapter 6. Whether the point was moved ‘right’ or ‘left’ of the path was determined by random selection. The assumed path of the officer was then altered to pass through this point and the time at which they were at this point was defined as the mid-point in time between the two actual pings. In other words, it was assumed that the officer moved in a straight line from \(p_i\) to \(p_k\) and then in a straight line from there to \(p_{i+1}\). A basic example of a ‘perturbed path’ is shown in Figure 7.2. As the perturbed paths were derived from the analysis conducted in the previous chapter, it is entirely possible that an officer may have walked these paths rather than the previously assumed straight line between GPS pings.

![Figure 7.2: An example of a re-calculated officer path](image)

As in Chapter 4, a series of self-exciting point process models were then generated. Four ‘baseline’ models were produced using the standard five-minute ping
intervals (i.e. using a direct route and not assuming a perturbed path) and the subset of data (these models are thus the equivalent of the models presented in Chapter 4). For these models, the dosage in a CAD cell for any given hour was estimated in the same way as in Chapter 4 (i.e. calculated for one hour units for 250 metre grid cells). As in Chapter 4, the evaluation models incorporated endemic components - to model the underlying (or background) risk of crime happening at a given time - and the self-exciting epidemic components which model how the occurrence of a crime modifies that expected risk. Each model was generated with and without patrol dosage covariates and thus the four models were:

- **Endemic without dosage:** An endemic-only model which accounted only for seasonality.

- **Endemic with dosage:** An endemic-only model which incorporated the patrol dosage within the hour being estimated and the patrol dosage over the previous 24 hours.

- **Endemic and epidemic without dosage:** An endemic-epidemic model which takes into account any spatio-temporal dependence between crime events.

- **Endemic and epidemic with dosage:** An endemic-epidemic model which accounts for both event dependence and patrol dosage during the current hour and over the previous 24 hours.

Both models which specify an epidemic component included a Gaussian kernel to account for spatial interaction. As in Chapter 4, it was found that the distance over which crime was ‘contagious’ was an artefact of having to ‘untie’ the data.

Patrol dosage in each grid cell was then recalculated by *simulating* potential paths which officers may have taken between pings by using the path perturbation method outlined above. Fifty simulations were conducted and the estimated dosages used to generate estimates for endemic-epidemic models with dosage.
Chapter 7. The impact of data error on evaluating police patrol

7.3 Results

The AIC scores for the four baseline models are given in Table 7.2. Of the four, the final model incorporating both endemic and epidemic components and including dosage covariates in the endemic component, produced the best fit with the data. The model estimates for the dosage covariates are provided in Table 7.3. Both the endemic-only and endemic-epidemic models show a significant positive effect of police patrol in a given hour - that is, police patrol increases the likelihood that a crime will be reported during that hour and is in line with the results presented in Chapter 4 when considering all MOPAC7 crime types. Recent police patrol (over the previous 24 hours) shows a significant result in the same direction as Chapter 4 (demonstrating deterrence) in the endemic-epidemic model but is not significant (at a 95% confidence level) in the endemic-only model.

Table 7.2: Model fits for baseline re-evaluation models

<table>
<thead>
<tr>
<th>Model specification</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endemic without dosage</td>
<td>6066</td>
</tr>
<tr>
<td>Endemic with dosage</td>
<td>6061</td>
</tr>
<tr>
<td>Endemic and epidemic without dosage</td>
<td>5945</td>
</tr>
<tr>
<td>Endemic and epidemic with dosage</td>
<td>5938</td>
</tr>
</tbody>
</table>

Table 7.3: Baseline model estimates

<table>
<thead>
<tr>
<th>Model estimate</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dosage - current hour</td>
<td>0.0349</td>
</tr>
<tr>
<td>Dosage - previous 24 hour</td>
<td>-0.0010</td>
</tr>
<tr>
<td>Endemic-epidemic model</td>
<td></td>
</tr>
<tr>
<td>Dosage - current hour</td>
<td>0.0405</td>
</tr>
<tr>
<td>Dosage - previous 24 hour</td>
<td>-0.0013</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

Hence, the results when taking this subset of the initial study presented in Chapter 4 appear to be broadly in line with the larger study period and area. The fact that the endemic-model does not show a significant effect of recent dosage might be explained by the smaller study size which would require a larger effect in order to detect a significant result.
The final full point process model, incorporating endemic and epidemic components as well as dosage covariates, was then re-run using patrol data which had been perturbed as outlined above. Fifty simulations were run. Table 7.4 and Figure 7.3 provide a summary of the distribution of the \( p \)-values for the dosage effect estimates. Of the 50 iterations, in contrast to the results shown in Table 7.3, none estimated that current-hour patrolling had a significant impact (at the 95% confidence level) on the likelihood of crime (either positively or negatively). Only one of the current-hour dosage estimates had a \( p \)-value less than 0.1 (though this was for a positive effect from current-hour dosage; and thus was in agreement with the findings from the baseline model). There was more evidence - or rather, more consistency in the results across simulations - to suggest that police patrol over the previous 24 hours deterred crime. 32 of the 50 simulations produced a significant negative result for that parameter estimate, which would mean that more police presence over the previous 24 hours reduces the likelihood that a crime will occur. Five of the 50 simulations produced an estimate that was significant at the 99% confidence level. However, the fact that 18 of the 50 simulations did not produce a significant result suggests that the estimated deterrent effect of police patrol is highly dependent on how the officer’s path is interpolated. This is an important point as it suggests that the results presented from evaluations of police patrol in micro-places may vary based solely on the (largely ignored) underlying error associated with measuring patrols.

<table>
<thead>
<tr>
<th></th>
<th>Dosage - current hour</th>
<th>Dosage - previous 24 hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum estimate</td>
<td>0.0179</td>
<td>-0.0003</td>
</tr>
<tr>
<td>Median estimate</td>
<td>0.011</td>
<td>-0.001</td>
</tr>
<tr>
<td>Mean estimate</td>
<td>0.010</td>
<td>-0.001</td>
</tr>
<tr>
<td>Minimum estimate</td>
<td>-0.0043</td>
<td>-0.0015</td>
</tr>
<tr>
<td>Estimates with ( p &lt; 0.1 )</td>
<td>1</td>
<td>44</td>
</tr>
<tr>
<td>Estimates with ( p &lt; 0.05 )</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>Estimates with ( p &lt; 0.01 )</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>
Chapter 7. The impact of data error on evaluating police patrol

7.4 Conclusion

This chapter evaluated how the method of estimating police patrol dosage impacts the estimated effect of that patrol. The assumed paths that officers had taken were simulated using an empirically derived measure of the expected error that results from assuming that officers walk in straight lines between locations measured only every five minutes. Previous research that has utilised GPS data in order to measure police patrol (Ariel et al., 2016; Williams & Coupe, 2017) has not investigated the assumption that these GPS data are accurate. The results presented above highlight some of the challenges of evaluating micro-place based policing strategies and in particular, accurately measuring police patrols. Before the availability of GPS data to measure police officer movements, evaluations of police patrol strategies relied on either independent observers who recorded the amount of time officers spent in specific locations (Sherman & Weisburd, 1995), officer-initiated call logs (Sorg et al., 2014), or assumptions that the intended intervention was correctly implemented (Andresen & Hodgkinson, 2018; Eck & Maguire, 2005; Novak et al., 2016). All these methods have measurement difficulties which are not easily overcome. Whilst

Figure 7.3: The distribution of p-values for simulated model dosage effects
GPS data carry many advantages over these historic methods, the weaknesses of GPS data have received no real scrutiny within the policing literature and the results presented in this chapter suggest that greater consideration should be given to trying to understand these weaknesses better and, where possible, account for them in future evaluations that use police patrol.

The analysis conducted here is not without its own limitations. First, the GPS data used are already subject to measurement error before paths between GPS pings are interpolated. The GPS pings themselves have been assumed to be accurate measurements of the officers’ locations at those times when this is unlikely to be entirely true. Particularly in a built-up city such as London, the urban canyon effect is likely to distort the true locations. Second, assumptions have had to be made as to the nature of the officers’ movements. Whilst some parameters have been set regarding what constitutes a patrol and what data should be discarded, it is not possible to verify what officers were actually doing during the study period.

Furthermore, officer paths have still had to be interpolated. For reasons discussed in Chapter 6, it was decided that trying to match officer patrols to known roads and pathways was unlikely to yield more accurate estimations of their true paths. As the data were collected before this research began it was not possible to measure the officer patrols in another way.

The method used to simulate potential officer patrol paths whilst based on an empirically derived distribution of expected error was ultimately quite simple. For reasons of computational feasibility, only one additional point was added between GPS ping-pairs. A minor evolution of this system would be to include more points along the path. Furthermore, the known urban environment could be incorporated into the path interpolation. For instance, the current system does not prevent an officer’s assumed path from passing through buildings or other barriers. Designing and implementing a more sophisticated path estimation technique was beyond the scope of this thesis, but this avenue deserves considerable attention in the future in order to allow more accurate measurement of patrol.

The study area was necessarily small and the time period short. For this reason
alone it would be sensible to repeat the analysis using other locations and time periods. Replication of this study would also provide a stronger evidence base from which to judge the adequacy of GPS-based evaluations of police patrol strategies.

The study conducted by Ariel et al. (2016) used two-minute ping intervals. Future studies should certainly investigate how actual GPS refresh rates impact not only the assumed patrol paths but also any evaluations of patrol effects.
Chapter 8

Discussion

8.1 Summary

The aim of this thesis has been to investigate how police patrols effect the amount of crime that happens in micro-places. Chapter 2 provided an review of why police patrols were originally implemented and highlighted why the places and times at which crime occur are as important to understand (if not more so) as the people who commit them. The rest of that chapter then discussed the history of research on police patrol strategies, which have shifted from large area interventions (such as the Kansas City experiment evaluated by Kelling et al. (1974)) to patrols targeted at areas no more than a few hundred metres across (such as the recent interventions in Los Angeles, USA and Kent, UK which were evaluated by Mohler et al. (2015)).

As discussed, the fact that police patrols can reduce crime is now well evidenced (Braga et al., 2019). However, as the spatial and temporal resolution of these patrol strategies has become more focussed, so too (by necessity) should evaluations of the implementations. This thesis has sought to highlight the importance of data quality in these evaluations.

Chapter 3 presented a limited illustration of the challenges faced by evaluations of micro-place based policing initiatives. In particular, it focussed on the challenges associated with using police recorded crime data and with using GPS data to measure police officer movements. The findings of the illustration that was presented broadly matched the findings of other studies and suggested that in order for police
patrols to have a deterrent effect, a minimum amount of time - about 10 minutes - needs to be spent in a hot-spot. Furthermore, the deterrent effect had little further impact when the patrol lasted more than about 20 minutes. This was in keeping with previous research on the subject. Koper (1995) found that 14-15 minutes of police patrol in a given two-hour window was the optimal amount of time to spend in a hot-spot. Williams and Coupe (2017) found that patrols of about 15 minutes were more effective than patrols lasting only 5 minutes. Telep et al. (2014) found that 15 minute patrols reduced the amount of crime in a hot-spot.

Other practical issues were also highlighted in that chapter; the number of designated hot-spots which were provided to police officers during the study period were too great for them to be able to visit each of them as intended. Other evaluations such as that by Ariel et al. (2016) found that patrols rarely lasted as long as planned. Implementation failure is a significant issue for crime reduction strategies (Knutsson & Clarke, 2006) and the findings from Chapter 3 highlight why evaluations that use an intention-to-treat method (such as those by Andresen & Hodgkinson, 2018; Eck & Maguire, 2005; Novak et al., 2016) need to ensure they are not over-estimating the amount of patrol that was delivered, and thus potentially under-estimating the effect of police patrol. Furthermore, the evaluation conducted in Chapter 3 used GPS data to measure police officer movements and to estimate patrol dosage. Few studies to date have utilised GPS data, which have clear advantages over other methods which have been used to estimate patrol dosages. Perhaps the most obvious advantage is the fact that using GPS data does not require the patrols officers to actively engage with the data collection task. The fact that methods that have previously been used - such as making officers log their entry and exit times when entering and leaving hot-spots - are potentially affected by the Hawthorne Effect (Landsberger, 1958). That is, officers may modify their behaviour because they are more aware that they are being monitored. In such instances, an evaluation of their behaviour may suggest that a certain effect is created but that effect may be reduced, or at least altered, once officers are no longer being actively monitored. However, GPS data are not perfect and their limitations are discussed below.
8.1. Summary

The analysis described in Chapter 4 expanded on the preceding chapter in a number of ways. First, the analysis incorporated dosage estimates across an entire London borough, rather than focussing only on the approximately 1% of a borough that might be designated as hot-spots. Second, the temporal resolution of the analysis was improved; dosage was calculated in hour units rather than the eight-hour period of a police shift. Third and perhaps most significantly, a different method of evaluation was employed which attempted to estimate both the background rate of crime and the excitation effect of a crime event separately. The models generated also accounted not only for the patrol dosage in a given area at a particular time but also the recent (past 24 hours) patrol history of a location. The findings were in agreement with many previous studies. Models of burglary offences found that an area was at a heightened risk of being victimised again in the days following a burglary and that police patrols could have a suppression effect on this risk of repeat-victimisation. Models generated for a broader grouping of crimes found that police patrols over the preceding 24 hours could deter crimes from happening. In contrast to previous research, the models did not find a near-repeat contagion effect from crime events. A potential explanation for this was put forward: that the hot-spot police patrol initiative in place at the time was designed around disrupting repeat and near-repeat victimisations, and thus if such an initiative were effective that might account for why near-repeat victimisations were not evident within the models.

The models of MOPAC7 crime types also found that police patrols significantly increased the likelihood that a crime would be reported during the hour of the patrol. To the author’s knowledge, whilst such a phenomenon has previously been hypothesised (Andresen & Hodgkinson, 2018), it has not previously been quantitatively estimated. This may be at least partly due to the temporal resolution of the analysis being much higher within this analysis than in previous evaluations of police patrol. It could be argued that, given the expected effects of police patrol, the analysis conducted in Chapter 4 has been performed at a more appropriate unit of analysis than much of the previous research into police patrol dosage. Even recent studies which
have used GPS data to measure police patrol dosage have aggregated their analysis to weeks (Williams & Coupe, 2017) or months (Ariel et al., 2016). The models of MOPAC7 crime types also suggested that recent police patrol dosage had a deterrent effect on crime. This effect was anticipated by Sherman (1990) and described as a ‘residual deterrence’ as offenders are unsure when a patrol might return to the area and thus the offender is less likely to offend.

The evaluations conducted in Chapters 3 and 4 assumed that the police recorded crime data and GPS data were accurate and complete. However, Chapters 5 and 6 explicitly investigated the veracity of this assumption. In Chapter 5, witness statements were analysed in order to assess the geocoding accuracy of the police recorded crime data that was used in Chapter 4. Due to the limited number of crime events for which witness statements were available, the quantitative findings must be treated with caution. However, several conclusions could be made from the data that were available. For instance, the MPS uses an automated system to geocode each crime location, and the system assigns each crime a geocoding accuracy classification. Errors in this geocoding system were discovered even for crimes which had been classified as the highest accuracy level. Historically, these errors have been less important. However, with the growing interest in micro-place based policing strategies, and evaluations that use (appropriately) more precise spatial and temporal units of analysis these errors are more likely to effect the results of such evaluations.

Several qualitative themes also emerged. The findings suggest that greater nuance is required in the designation of which areas are most at risk of crime. Many of the crimes analysed were not, as originally assumed, likely to be susceptible to a deterrent effect from police patrols. This was because they did not require an offender to pass through a public area in order to reach their victim and thus the ‘capable guardianship’ that police patrols are expected to provide can not be delivered. In particular, many of the violence and criminal damage offences were committed in private locations. As such, police patrol strategies that target these crimes at micro-places would be less effective than they would otherwise be if these
8.1. Summary

non-susceptible crimes were removed from the risk calculation.

Chapter 6 used GPS data collected from foot patrol officers in London UK to assess the similarity of assumed officer paths based on different GPS ping refresh rates. The findings suggest that the median difference between a path interpolated from a 30 second ping rate and the path interpolated from a 5 minute ping rate is about 60 metres. However, the distribution created from comparing paths created from 30 second pings to paths created from 5 minute pings showed that the distance between the assumed paths could range up to more than 250 metres. This distribution may represent a conservative estimate of the difference due to the constraints placed upon what was counted as a path. To the author’s knowledge, no other research has examined the accuracy or similarity of assumed police patrol paths using GPS data. Studies which have used GPS data to measure patrol (Ariel et al., 2016; Williams & Coupe, 2017) have not discussed the accuracy of their data. Though neither the study by Ariel et al. (2016) nor the study by Williams and Coupe (2017) interpolated the officer’s path between GPS pings.

The findings from Chapter 6 were then used in Chapter 7 to re-evaluate the effects of the patrol strategy covered in Chapter 7. The assumed paths that officers took between GPS pings was perturbed to simulate a potential path they may have taken, rather than assuming they walked in a straight line between pings. Multiple path perturbations were calculated and each was used to generate models in the same way as the analysis in Chapter 4. When the original direct path between GPS pings was used to model the effect of police patrol on crime, the models produced similar results to the earlier analysis - that during a police patrol, there was an increase in the likelihood that a crime would be reported but that patrol dosage over the previous 24 hours reduced the likelihood of a crime being reported. However, the results across all 50 simulations that were generated were less consistent. Whilst more than half of the simulations produced a significant effect estimate for recent police patrol - that it suppressed crime - there was little evidence that police patrol changed the likelihood that a crime would be reported in the same hour. These findings suggest that, whilst evaluations of police patrols at micro-places are a welcome
and necessary addition to the literature, the quality of the data being used in these evaluations can have a significant impact on the estimated effects of police patrol.

8.2 Limitations and further work

This thesis has provided a timely analysis of some of the issues that exist when evaluating micro-place based police interventions. However, several assumptions had to be made during the analyses and the research conducted has not been without its limitations. This section discusses these and also presents avenues for future work.

Fundamental to the entire thesis was the measurement of police officer movement. Both Chapters 3 and 4 used GPS data from officer-worn radios to measure police patrols. A key assumption made in calculating patrol dosage was that one patrolling officer would provide the same deterrent effect as multiple officers at the same location. As discussed, this was motivated by research (Kleck & Barnes, 2010) that has shown that patrols involving two officers offer no crime reduction advantage over those involving one. However, it is possible that when multiple officers patrol an area independently they would have a greater cumulative deterrent effect than would two officers walking side-by-side. This might be explored in future work, as might the association between the total geographical area covered by (single or) multiple officers, as opposed to the duration they are present. Similarly, consideration might be given to lines of sight, and how visible police patrols might be to passers-by, given the configuration of the street network. For example, police officers may be more visible to the public along straight roads than they are on short street segments that are part of an intricate section of the street network.

Related to the above point regarding the number of officers on patrol, the perturbed paths calculated in Chapter 7 were simulated for each officer individually. This means that where two officers were patrolling side-by-side, their original paths would likely be quite similar if their five-minute pings are close together. However, their perturbed paths may diverge as the method used to perturb a path did not account for other officers nearby. It was beyond the scope of this study to try and
define when patrols were ‘solo’ or ‘paired’ and to try and incorporate this into either
the initial evaluation conducted in Chapter 4 or the simulated perturbed paths used
in Chapter 7. Future research into how the number of officers impacts on crime in a
given area is certainly warranted.

This leads onto another boundary issue inherent in the use of the CAD grid
for aggregating patrol dosage, namely that the grid system is an artificial construct
designed to be easily understood by police officers, but which ignores entirely the
underlying environment (R. B. Taylor, 2010). Recent research has demonstrated
that there are gains to be made by using street-network based predictive algorithms
that account for this rather than grid-based ones that do not (Rosser et al., 2016).
Moreover, the analysis presented does not account for potential spill over effects
whereby what happens in one cell affects what happens in those nearby. Taking
account of such dependencies might produce different results to those presented
here. For instance, Baudains et al. (2019) did incorporate neighbouring regions
into their analysis of insurgent and police interactions in India. The model was a
better fit for their data than the same model without a neighbour-area interaction
term suggesting that a natural extension of the research in this thesis would be to
incorporate nearby patrol dosage into the models generated.

The objective of the police patrol strategies covered here was the reduction of
crime. Police recorded crime was the most suitable dependent variable for these
analyses given this objective, however there are other dependent variables which
could have been used. For instance, previous research has also investigated the
effect of police patrol on citizens’ fear of being victimised (Kelling et al., 1974)
or their confidence in the police. Research into public perceptions of the police
by Bradford et al. (2009) found that positive encounters with police officers and
visibility of officers improved confidence amongst the general public.

Confidence in the police is also linked to perceptions of fairness and efficacy.
One of the major limitations of using GPS data to measure police patrols is that they
do not account for what the police were doing. As discussed in Section 3.6, a possi-
ble future approach to studying the effects of police patrol would be to incorporate
additional data sources such as police logs, briefing information, or officer interview data to augment the ‘where’ provided by GPS data with an understanding of ‘what’ they were doing. Furthermore, to augment the accuracy of the locational data, other data can be used to determine an officer’s location. For instance, mobile phones use nearby telephony masts and WiFi receivers in order to more precisely locate an individual and a similar method could be used to triangulate officer locations and movements.

Finally on the issue of measuring police patrol, this thesis has only considered police officer foot patrol. This is of course not the only means of transport for police officers; officers routinely travel by motor vehicle and increasingly by pedal bicycle. Furthermore, PCSOs as well as other private security personnel are not incorporated in any of the analysis in this thesis. Ariel et al. (2016) found that PCSOs can also have a deterrent effect on crime even though they possess very few powers compared to a police officer and so inclusion of these other ‘capable guardians’ in the determination of dosage would be a sensible extension of the approach taken in this thesis.

Chapter 5 represented an attempt to understand the error and uncertainty of police recorded crime data but the ability to validate that error was hindered by the small proportion of crimes for which victim statements were available. Given that police recorded crime data are already known to be imperfect (Her Majesty’s Inspectorate of Constabulary, 2014b; Her Majesty’s Inspectorate of Constabulary and Fire & Rescue Services, 2018), this area of research warrants further attention. The original intention of this thesis had been to perform a similar analysis to the one outlined in Chapter 7 with crime data as well as patrol data. Such an analysis would still be relevant to perform if the locational information contained within the crime data could be more robustly quantified. Crime data are also used to determine where hot-spots are designated, and as such verifying the accuracy of crime data geocoding is relevant not just to evaluations of police patrol but to determining if patrol locations were correctly defined in the first place.

Bringing together the issue of defining a hot-spot as a some-what artificial
8.2. Limitations and further work

‘box’ area and the difficulty in measuring police patrols, a future area for consideration (both from an academic and practitioner point of view) is to provide officers with optimal patrol paths rather than simply ‘areas to get to’. Ariel et al. (2016) highlighted that officers felt they were ‘chasing the clock’ the reach their intended patrol locations. It may be that officers are more receptive to a patrol strategy which provides them with an approximate route to take on their shift rather than specific areas they are expected to reach. A system that schedules randomised patrols has previously been deployed in Los Angeles, USA (Yin et al., 2012). Although that system was for the much more spatially confined paths of the LA Metro Rail system, it provides an example of patrol path planning which is reactive to changes in patrol resource availability.

To conclude, this thesis has highlighted how data quality can impact the estimated efficacy of police patrols in micro-places. Whilst it may not be possible to mitigate all the data quality issues these should at the very least be discussed and acknowledged within evaluations conducted at high spatial and temporal resolutions. Given that police patrol strategies are (rightfully) becoming more focussed in time and space it is only appropriate that evaluations of these strategies are conducted at similar units of analysis.
Bibliography


HMSO. (1862). General Regulations, instructions, and orders for the government and guidance of the Metropolitan Police Force.


Shaw, C. R., & McKay, H. D. (1942). *Juvenile Delinquency and Urban Areas*. Retrieved from http://www.amazon.com/delinquency-delinquents-differential-characteristics-communities/dp/B004OTU11O/ref=sr%7B%5C_%7D1%7B%5C%7D%7B%5C%7D4?s=books%7B%5C%7D%7B%5C%7D%7B%5C%7Dqid=1407177860%7B%5C%7D%7B%5C%7Dsr=1-4%7B%5C%7D%7B%5C%7Dkeywords=juvenile+Delinquency+in+Urban+Areas+1942


