Thesis submitted for the Degree of Doctor of Philosophy

Reconceptualising risky facilities: exploring identification, patterns and features using a mixed-methods approach

Melanie Flynn

Department of Security and Crime Science
University College London

January 2018

Declaration

I, Melanie Flynn, 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 thesis.

Signed:

Date: 31st January 2018

Acknowledgements

This thesis has been a long time in the making. From when I first registered for a PhD I have moved jobs (losing direct access to the data I needed for the research in the process), moved house four times, got divorced, remarried and had beautiful twin boys. Not long after starting the PhD, I secured a full-time senior lectureship. I then completed a Postgraduate Diploma in Continuing Professional Development (Higher Education Practice), and grappled with the roles of firstly admissions tutor and latterly course leader, which has absolutely drained me and taken up so much of my time. I have been involved in course re-validations, new course validations and tried to explain why my course has not met its (in my opinion highly over-ambitious) targets for student achievement. As a lecturer I still believe education is not (just) about the grade that someone achieves, but the process and experience of being at university. I believe a degree is so valuable because it teaches you how to learn; how to enquire; how to view the world with a critical eye; and in the current socio-political climate, I believe that nothing is more valuable than the ability to challenge and to evaluate the evidence-base. Many of my students have faced extreme personal difficulties and challenges, and some of them have been the first in their family to go to university. When you walk across that stage to graduate I am proud of every single one of you and I know the worth of your achievements cannot just be measured by your final degree classification. Thank you to you all, for helping me become the academic that I am today.

I embarked on this PhD, expecting it to be intellectually difficult, but I viewed it as ‘just’ a large research project, and not that different to things I had done before. I was completely wrong. It has been intellectually challenging, but the most difficult part of completing the PhD has been trying to manage it alongside full-time work and (more recently) family life. However, during the long period I have been registered as a PhD student, I have also developed significantly. I know my subject area so much better than I did previously and this has really helped my understanding of the material that has gone into my thesis. I have also carved out a reputation for myself as something of an expert in the field of wildlife crime, within the wider area of green criminology. I have successfully published in this area, captured research money, produced work for the WWF-UK and flown out to Malaysia to advise the judiciary of Sabah and Sarawak on the introduction of sentencing guidelines for wildlife offences. All of these things have helped establish my credibility as an academic, and develop my confidence. For that I am hugely grateful to the University of Huddersfield and my colleagues and various line managers, who have supported me in these endeavours and recognised my research output (allowing me something of a reduction in teaching hours). At the same time they have nagged me incessantly to GET THE PHD COMPLETED. Less time invested in the course I lead, my students, and my other research commitments might have helped me put this to bed sooner, but my integrity and

commitment to my paid work and my students prevented this. For that, I cannot apologise.

So my job has made completing this PhD extremely difficult, but at the same time I am hugely proud of the things I have achieved in my career so far. I have also met many supportive colleagues, many of whom I now class as some of my closest friends; people I can turn to who will understand and empathise (and never judge!). So to my colleagues past and present, thank you. In particular, thanks to Jason Roach who alerted me to the job at Huddersfield in the first place and has been mentor, adviser, office-mate, and friend. I even copied him and had twins! To Grainne McMahon, my favourite feisty, Irish feminist and the person who shared with me the pain of being 'poisoned' by Sambuca. Thank you for being my confidante when I needed it the most - I would not have got through those things without you to talk to. To Kate Clegg, who started out as my student and then became one of my best friends. We've shared laughs and tears and I know you are always there for me (just as I hope you know I am for you). Our busy lives have kept us (and our children) apart far more than they should, but hopefully we will both find some time soon to do something about that. Your girls (and the dogs) are incredible and you should be very proud of what you have achieved in your life so far. Love you.

To Carla Reeves, who started working at Huddersfield not long after me and is now my line manager. I wouldn't have it any other way. You've always had my back, even when you have to deal with my 'enthusiastic responses' to all the (often ridiculous) diktats that come our way. I don't envy what you have to deal with (but, hey, you get paid more than me!). Without wanting to sound patronising, I'm also really proud of everything you have achieved. You deserve every happiness, my friend. To all the administrative staff at the University of Huddersfield, whose work is invaluable in supporting the academic staff, and particularly to Christine Bower, our long-time course assistant (enjoy your retirement) and to Julie Stakes who taught me how to deal with admissions when I first started at Huddersfield, and talked with me about her grandchildren and my boys, until she lost her battle with cancer last year. Julie, you are much missed.

Finally, to old colleagues - and friends - Dan Birks and Susan Donkin. Working with you both as part of 'CSI:Brum' was honestly the best part of my career so far. I have missed you incredibly whilst you have been living in Australia (and haven't been the tiniest bit jealous that you completed your doctorates way before me!). I still remember not having a clue that you two were seeing each other, but I was so happy for you when I found out. Now you have an amazing daughter and I am really looking forward to having the chance to spend more time again with you when you come 'home'. Play-dates to be booked asap.

To all my other colleagues, too many to list here, I know many of you probably think I am too

dictatorial for my own good (and certainly opinionated) so you may not realise this, but you have all influenced the academic I am today. I have been really lucky to have worked with such a diverse, and honestly lovely bunch of people. That said, there are a couple of people I could well have done without having to work with, and there's nothing like having to respond to a spurious (later withdrawn) grievance when you are in the last few months of your PhD - so, yeah, thanks for that, too.

I also thank University College London and the people who I have crossed paths with as colleague or fellow students of the Jill Dando Institute. Naturally my thanks also go to my supervisors Kate Bowers and Shane Johnson, who have been with me throughout the long slog. Sorry! Your guidance has always been practical and to the point - just what I have needed. Kate, you have tried so hard to get me across the finish line (and much sooner than this). Thank you for not losing patience with me (or at least not letting on that you had!). I also thank you so much for the trust you have placed in me as I have finished much of this final write-up with little time for you to provide me with feedback. Any glaring omissions or things that just don't make sense are, then, entirely down to me (and you've probably not even seen them yet), so thank you for keeping the faith.

Of course, I am indebted as well to the American police department that provided me with the secondary data upon which this thesis is based, and facilitated my access to officers and the study area in general. In particular to the Analyst who heard about my data problems and offered to help, ultimately arranging my access to the crime and offender data, as well as posting me a DVD containing geographic base data. Without this, I could not have continued, so you truly saved my PhD. Everyone that I spoke to from the PD and all those who I met when I undertook my field trip were unreservedly helpful and welcoming. Whilst the secondary data made this research possible, the incorporation of various forms of qualitative data and insight, including my own experiences and observations, has significantly added to the findings and their interpretation and this would not have been possible without the time and assistance of those who agreed to be interviewed, the then Chief, and the officers who drove me around the study area whilst I did some of my observations. The field trip was supported by UCL and by the University of Huddersfield, who also paid some of the PhD fees I incurred, so thanks also for this financial support.

Colleagues and organisations considered, I turn to my other friends and to the most important people in my life - my family. A PhD journey as long as this one has inevitably seen some people fall by the wayside, but I could not have done it without all of you. To Dan, we had a good run and were genuinely happy. We have shared some amazing experiences and supported each other in all our endeavours. Unfortunately, this eventually meant we grew apart and started down separate paths. I am so glad that you have found Caroline and are truly happy and thank you so much for the support you gave me as a new academic and a new PhD student. I am also grateful that we had such good people in our lives that we can both still call friends, especially you Keith; if only our plans for world domination had been put into effect!

To my Mum and Dad, you have given me everything. You have made sacrifices throughout your lives to provide me with the opportunities that you didn't have. I could write an entire thesis on what you have done, how you have supported me, how much I appreciate it (even if you don't always realise this) and how very much I love you both, but I think I should keep this short. All I will say then is THANK YOU from the bottom of my heart, and I hope I have made you proud.

Finally to my darling husband, Chris. You turned my life upside down and I've loved every minute of it. No regrets, ever. There is no-one for me except you. I fell head-over-heels in love with you and nearly seven years later I still feel the same. You make me incredibly happy and you have been my rock when things have been so difficult I have nearly given up. You also gave me the most precious thing possible: our twins, Brandon and Connor. Granted, this has made finishing the damn thing even harder, almost impossibly so (why won't they just sleep?!), but I would not be without them for the world. I am so lucky to know what it is like to be Mummy, and without you I would never have had that. Thank you with all my heart. Now I cannot wait to truly enjoy our lives together without the spectre of this doctorate hanging over us and demanding so much of my time.

Not really an acknowledgement, but also love to the various animals that have shared my life: Sherry, Pixie, Maxine, I miss you. Geordie and Trafford, my fur babies, your cuddles and purrs have brought happiness and calm. To all other animals, I will continue to fight for you and try to be the voice you do not have.

The last few years, including trying to complete this PhD, have nearly broken me. But I am still here, so ultimately from this experience, I will take the following:

I am strong. No matter what, I will not give up.

Melanie Flynn, January 2018

Abstract

This research concerns the spatial concentration of crimes at the micro-level, with a particular focus on the concept of risky facilities. This is done through an exploration and reconceptualization of crime associated with facilities and attributable to specific premises, or addresses. I utilise a mixed-methods approach that draws heavily on quantitative, secondary analysis of police recorded crime data, but also makes use of interviews with serving police officers, annotated maps and observations that are analysed qualitatively. Mixed methods research is underused in this field of study, thus this provides an additional, methodological contribution.

The research aims to fill a gap that was formed after the initial conceptualisation of risky facilities and the subsequent research that has tended to focus on how to explain the existence of these ‘problematic’ premises and, therefore, what to do about them. Thus the research is predicated on my contention that the concept of risky facilities has not been sufficiently empirically tested, nor defined. As such, the broad original contribution of this research is that it provides an empirical and conceptual exploration of the existence, nature and definition of risky facilities.

The aims of the study, therefore, are to consider whether crime concentrates in facilities, and how; to consider how this concentration manifests and can be identified; to explore key features associated with risky facilities; to critically discuss the concept of facility concentrations and consider the appropriateness of definitions and terminology within this field; to add to, and where necessary challenge, existing knowledge and its application in the field of spatial crime concentrations, particularly with respect to crime in facilities; and to make recommendations for policy, practice and further academic research.

Statement of research impact

The research reported here has advanced the use of mixed methods research in a field dominated by quantitative and quasi-experimental approaches. It also proposes an empirical definition for the spatial crime concentration of ‘risky facilities’ that can be applied to studies of this phenomenon anywhere in the world, to facilitate greater replicability and comparison. It has further added to the evidence on the ubiquity of this phenomenon, but has done so utilising an under-used method of crime attribution, and shown that despite the possible limitations of such an approach, this is a viable and more conceptually defensible approach. The research has also explored the distribution of risky facilities across a city and the journeys taken to offend there.

Contents

List of tables

List of Figures

Copyright and licenses

Dissemination of research findings

Publications

The following publications are in the final stages of preparation and are based on the research generated in this thesis:

Flynn, M. & Bowers, K. (in preparation). Exploring crime concentrations: The ubiquity of risky facilities. [Chapter 7]

Flynn, M. (in preparation). Identifying premises that are 'risky': A proposal for an empirical 'cut-off' for labelling premises as risky facilities (or not). [Chapter 8]

Presentations

The following presentations were based on research contained in this thesis:

Flynn, M. (2017). Police perceptions of risky facilities. Presentation to the European Society of Criminology Conference, September 2017, Cardiff, UK.

Wellsmith, M. (2012). Identifying and assessing risky facilities. Presentation to the American Society of Criminology Conference, November 2012, Chicago, Il, USA.

Wellsmith, M. (2010). Exploring within facility crime concentrations: Some early findings. Presentation to the Environmental Criminology and Crime Analysis Symposium, July 2010, Brisbane, QLD, Australia.

Wellsmith, M. (2009). The Routines of Crime: Using micro-analysis for theory testing in environmental criminology. Presentation to the University of Huddersfield, School of Human and Health Sciences Research Day, March 2009, Huddersfield, UK.

Wellsmith, M. (2008). Risky Premises: Analysing crime attractors, generators and risky facilities. Presentation to the European Society of Criminology Conference, September 2008, Edinburgh, UK.

Publication plans

The following papers are also in various stages of preparation, and will be completed and submitted after this thesis has been examined (note that these are working titles only):

Flynn, M. & Bowers, K. (in progress). Journeys to risky facilities. [Chapter 9]

Flynn, M. (in progress). Police officers' perceptions of risky facilities (and other high crime areas). [Chapters 7, 8 & 9]

Flynn, M. (to prepare). Using qualitative methods to explore spatial crime concentrations: Notes on a risky facilities field study. [Chapters 6, 7, 8 & 9]

Flynn, M. (to prepare). Crime attractors, generators, repeat victimisation and risky facilities: Are they one and the same? [Chapter 10, with reference to various findings]

Chapter 1: Introduction

It is now well established that crime concentrates across a number of dimensions, including space. This appears to occur at all levels of aggregation, although most research has been carried out at the macro (town, region) or meso (problem estate, hotspot, neighbourhood) level. There are far fewer (albeit a growing number of) studies of micro level crime concentrations. Further, the research on crime attractors and generators is theoretically and empirically questionable. Recent work on ‘‘risky facilities’ (Eck, Clarke & Guerette, 2007) begins to address the former issue, but raises two further questions. Firstly, what is the relationship between crime attractors, generators and risky facilities? Secondly, what do we actually know about the concept, and existence, of risky facilities? To what extent do they exist across a range of premises, crime types and times? What is the nature of such patterns of concentration? Indeed, how should we even define something as a risky facility? This important concept, once identified and (somewhat loosely) defined, has been applied and explanations have been sought for why some premises are risky when others are not. Whilst this is an extremely important pursuit, it seems to be a not uncommon example of ‘putting the cart before the horse’. It is my contention that the concept of risky facilities has not been sufficiently empirically tested, nor defined and that before we can embark on efforts to explain this phenomenon, we should first test the extent of its occurrence, the breadth of its existence and determine a clear, broadly applicable, empirical definition. In doing so, we therefore justify further research endeavours. By establishing an empirical definition, we strengthen further research (and enforcement-based analysis) and make it more comparable: the criteria for determining if a facility is categorised as risky or not becomes more consistent and defensible.

Thus, in broad terms, my original contribution is an empirical and conceptual exploration of the existence and nature of risky facilities, that results in further knowledge about the extent and breadth of this type of spatial crime concentration, provides an empirical definition that results in a universally applicable ‘cut off’ between risky and non-risky facilities and situates our understanding alongside the related concepts of crime attractors, generators and repeat victimisation.

Drawing on environmental criminology and opportunity theories, this study aims to pursue these issues. The thesis proceeds through a review of the theoretical literature, research findings relating to spatial concentrations and then a critique of the approaches used to identify and measure spatial crime concentrations, specifically those that might be referred to as risky facilities. The methodological approach and data collection and preparation are then presented.

The research consisted of three studies. The next three chapters present the data preparation, findings and discussion of the material from each of these. Finally, there is an overarching discussion and summary of key findings, that contains also recommendations and suggestions for further research.

Chapter Two: Theoretical underpinnings

This study considers spatial (and temporal) concentrations in crime across a range of public and quasi-public establishments. In the subsequent chapter I consider the evidence for the existence of these concentrations and how they might be explained. Firstly, though, it is necessary to consider how place has been theorised and incorporated into criminological study. Thus, this chapter briefly presents early explorations of crime and place before focusing on the approach and theories that underpin my study; broadly environmental criminology and opportunity theories.

2.1 Introduction

As certain places (and times) experience more offences than others, and such patterns are not random, it is reasonable to ask if there are things about these places at these times that make crime more likely. Dispositional (whether internal, sociological or environmental) and functionalist theories of criminology have notably failed to explain, or even address, such questions, focusing instead on why it is that certain individuals or groups appear to commit crime or be criminalised more than others.

A collection of theories and approaches that we might call ‘environmental criminology’ have developed, in part as a response to this disregard of the role of ‘place’ in crime and the phenomena of crime concentrations. These are discussed further, below; however, it is important to first briefly consider the history of place-based explanations of crime.

2.2 Early Explanations

The first notable consideration of place in relation to crime was that of the early criminological cartographers; the first to note a relationship between crime, place and offenders. Working in the 19th century against a backdrop of social dislocation resulting from the agrarian and industrial revolutions, these moral statisticians and ethnographers were also influenced by the scientific drive towards systematic collection of facts about the social world, with a desire to match these to spatial patterns in crime (Smith, 1986). Much of this early work established that crime is not randomly or evenly distributed through space nor, as Guerry showed, are spatial concentrations the same for different crime types (Morris, 1957). This research was not limited to mapping social statistics, either. Mayhew (1862), for example, mapped interview data, revealing the existence of London rookeries: areas of high concentrations of offender residences. However, although the criminological cartographers were the first to identify spatial concentrations in the distribution of crimes, victims and offender residences, the explanations given were situated firmly in the social world.

The most significant movement in criminology to next consider spatial patterning was part of the famous Chicago School and is credited as playing a major role in shaping thinking about crime and place, despite many of the original ideas now being considered outdated. An academic approach closely related to social intervention, much of the innovative and comprehensive research carried out, and the model developed, is now viewed as something of a product of its time and the specific circumstances that had developed in the ever-growing metropolis of Chicago. This was a new, urbanising and rapidly expanding city, that had grown from just 200 inhabitants in the 1830s (being incorporated as a city in 1837) to around 3 million inhabitants by the 1930s. The rate of expansion was additionally fuelled by restoration and rebuilding after the Great Chicago Fire of 1871 (Bellair, 2017; City of Chicago, 2018). It was against this backdrop of urban growth and the influx of immigrants seeking work in the factories and packing plants that served this important trading centre, that Park, Burgess and McKenzie (1925) published their famous work *The City*.

The underlying approach taken can be seen as one of human ecology (Grove and Burch, 1997); a competitive struggle for resources and space resulting in a kind of Darwinian 'survival of the fittest'. It was proposed that this led to radial expansion of cities, resulting in concentric rings or 'zones' that each had a different social and economic make-up (Burgess, 1925; McKenzie, 1925). Additionally, the speed with which such growth and expansion occurred, and the mix of newly arriving rural American and European immigrants, was seen to result in an anomic society, lacking in established support networks and social norms. Thus the approach was also influenced by the work of Durkheim (1892; 1897). This evocation of human ecology was one that centralised the deterministic and functionalist nature of society (communities) (Grove and Burch, 1997; Bellair, 2017).

Building upon these conceptions, further research using social statistics established that offenders were far more likely to live just outside the central business district in what was termed the 'Zone in Transition' (Zone II) and ethnographic study revealed more about the nature of this distribution and the different 'societies' that were formed within the different zones. This led to Shaw and McKay (1931; 1942) proposing social disorganisation as the explanation for the high concentration of offences occurring in the Zone in Transition. This 'disorganisation' resulted from the large number of immigrants, deteriorated housing, abandoned buildings, and the lack of an established population (as people moved out to more attractive zones when they were economically and socially able). Shaw and McKay (latterly supported by the research of Lander (1954)) also described the negative effects of racial or ethnic heterogeneity. All of these features were seen to prevent the development of an organised, stable community with established *shared* values and crucially, therefore, an ability to exert informal control on its inhabitants (Bellair, 2017; Bottoms & Wiles, 2002; Shaw & McKay, 1942).

This idea of social disorganisation as a cause of crime is a form of sociological positivism. In other words, it is the nature of society that causes some of its inhabitants to offend. In this case, it is the experience of living in an unsettled environment with few established norms or social control. Later research in geographically disparate areas was generally unable to replicate the concentric city zones seen in Chicago, an issue compounded by the difficulty of (and lack of agreement regarding) operationalising the concept of social disorganisation and of assuming that informal control must lead to low crime, rather than being a response to high crime, as was found in some research. In addition, the environmental explanations given for the distribution of offenders across a city were criticised for ignoring the effect of regimes of power and dominance, for being overly reductionist and deterministic and for assuming that shared-values mean non-criminal values (Bellair, 2017).

The approach seemed to have ‘had its day’ as urban growth stabilised and then declined following the post-war boom in the US. Poverty, socio-economic status (SES) and (relative) deprivation continued to receive attention, but now concerns were about urban decline, not transition and instability through growth and movement (Bellair, 2017; Wilson, 1987). After something of a hiatus period, however, social disorganisation was somewhat reconceptualised to account for the (less systemic) effects of this decline on even socially integrated neighbourhoods that were unable (or unwilling) to exert prosocial informal control, as a result of social exclusion or what may be thought of as an absence of stake in conformity (Bellair, 2017; Toby, 1957; Wilson, 1996).

More recently, this approach to spatial crime concentrations has been revitalised by the application of new research and analysis methods (Raudenbush & Sampson, 1999), and development of the more ‘dynamic’ conception of collective efficacy (Bellair, 2017; Sampson and Groves, 1989; Sampson, Raudenbush & Earls, 1997). This has highlighted the importance of social cohesion (particularly in the form of shared values and trust) *combined with* a willingness to intervene in protecting against crime within communities (Bursik, 1988; Kubrin & Weitzer, 2003; Sampson & Groves, 1989). Research has demonstrated that social cohesion need not be dependent on stability and networks, and that collective efficacy measures are associated with lower levels of neighbourhood violence. These findings are generally consistent across a number of different locations and countries (Pratt & Cullen, 2005; Sampson et al., 1997; Sampson, 2012), including the UK (Hirschfield & Bowers, 1997), though Bellair (2017) points out the seeming exception of the Netherlands (see, for example, Bruinsma, Pauwels, Weerman & Bernasco, 2013; Steenbeek & Hipp, 2011).

Whilst such an explanation for area crime rates might hold significant worth, there remain operationalisation and measurement issues (Bellair, 2017) particularly when seeking to obtain suitable data for small geographic units (Sherman, Gartin & Buerger, 1989). Despite this, there have been attempts to apply social disorganisation at the micro-level, with proponents arguing that even small areas such as street segments can be recognised as having their own ‘communities’ and social life (Weisburd, Bushway, Lum & Yang, 2004; Weisburd, Groff & Yang, 2012; Weisburd, Shay, Amram & Zamir, 2018). This research has struggled to obtain data that is appropriate at this level, though the most recent published study has tested the concept in Israel, where such data are more readily available, with promising outcomes (Weisburd, et al., 2018). Even considering this, though, because of the focus on street segments it remains questionable how applicable such explanations are to crime concentrations at premises level, as will be discussed in later chapters.

It can be seen, therefore, that early explanations of crime placement focused predominantly on the spatial distribution of offenders or the sociological and developmental impact of particular environments on individuals. As such, the role of place in crime remained less important than the role of people. Here ‘place’, to criminologists, was where offenders (or victims) lived. At most, the general environmental diaspora was seen to affect the disposition of those exposed to it, making them more likely to carry out deviant acts. These theories do not, therefore, directly consider the importance of place in determining *what* crime is committed, *where* and against *which targets*.

It could be said, therefore, that they fall foul of two major criticisms that may be made of dispositional theories of crime generally. Firstly, they operate on the seemingly erroneous assumption that there are ‘criminals’ and ‘non-criminals’ who are distinguishable and in some way pre-determined or fixed. Secondly, such approaches suppose that these people identified as criminals are driven to offend and, once pre-disposed to do so, be it as a result of some faulty biology or psychology, their experience of the society in which they live or the reaction of society to them, will do so regardless of the suitability of the environment in which they find themselves.

However, contemporary interest in crime places developed differently with a focus on offender decision-making, physical environments and movements patterns through these.

2.3 Environmental Criminology and Opportunity Theories

Early proponents of environmental criminology, Brantingham and Brantingham (1981), pointed out that a criminal offence has four features, or requirements: a law (proscribing such behaviour); an offender (the preoccupation of most criminologists until this point); a victim (albeit this could be debated for certain types of offence, such as narcotic possession); and a place. Quite rightly, they called for all four features to be the focus of criminological enquiry, whilst recognising that place had been notably ignored by much criminological endeavour. So it was that perspectives offering place-based explanations for crime or the patterning of offences came to be known as environmental criminology. As such, environmental criminology is an umbrella term that may be applied to a range of theories and approaches interested not only in the people of crime, but also the environments, or places, in which it occurs. It should be noted, however, that in recent years, this term has also been co-opted by some within the 'green criminology' field. It has been suggested (see, for example, White, 2008) that the alternative name 'place-based criminology'' be used to avoid confusion. Whilst there is a good argument for this change, 'place-based' may in fact be too narrow when applied to the range of theories and approaches currently associated with environmental criminology. As a result, I will use the original term in the absence of a more appropriate, and widely-recognised, alternative.

As will be considered in chapter 3, crime concentrates across a number of dimensions and entities. Of particular interest for this research are concentrations in space. Early place-based explanations of crime have been shown to be unsatisfactory for considerations beyond the offender or their residence. By continuing to focus on the offender and not considering the location of their offending, these approaches are incapable of explaining why certain areas or facilities experience higher crime than other, often similar, places. Therefore, the key theories that I will be drawing on in my consideration of place-based concentrations in crime are routine activity theory, the rational choice perspective, and crime pattern theory.

These are opportunity theories as their focus is upon what is a suitable crime opportunity (a particular set of conditions or circumstances), how these opportunities are formed, how they are 'discovered' by potential offenders and why they are (or are not) acted upon. All of these are important questions for considering why it is that crime concentrates spatially and temporally, as will be discussed later. These theories are now introduced.

2.4 Routine activity theory

Routine activity theory states that for a criminal offence to occur a motivated offender must come together in time and space with a suitable target in the absence of a capable guardian (Cohen & Felson, 1979). This theory can be used to explain changing patterns of crime observed at different levels of granularity. For example, Cohen and Felson (1979) believed that an increase in domestic burglaries at the macro (country) level was the result of more single-family households and increasing numbers of women going out to work (more empty homes, less guardianship), as well as the proliferation of lightweight consumer goods (suitable targets) (see also, Felson, 1987). Although the meso level was not explicitly considered in the original paper (Felson, 2008), at this level routine activity theory can help to explain why crime is not randomly, nor evenly distributed through space. Finally, at the micro level, routine activity theory can explain individual victimisation. Crime patterns are (very simply) explained by routine activity as being those places that provide the best opportunities. As not all places will provide such suitable opportunities, crime will not be evenly distributed.

Felson (1995) has stated that the motivated offender in this crime equation was originally taken as given. In other words, routine activity theory assumed a ready supply of people willing and able to take advantage of crime opportunities as they were presented. This was to change when the concept of the intimate handler was introduced, which is discussed below.

The first protective element introduced was the capable guardian. Even if a motivated offender came across a suitable target, it was postulated that an offence would not take place if that target was being guarded. The capable guardian need not be somebody formally responsible for the target (such as an owner or security guard for a theft target), they may simply be a passer-by (although the effectiveness of varying responsibilities is argued; compare, for example, Newman (1972) and Jacobs (1961) with Clarke (1992) and Mayhew (1991), as discussed by Felson (1995)). Today, technology can also provide guardianship although, as with people, it may not always be 'capable'.

Through the linking of routine activity theory to Hirschi's (1969) control theory, Felson (1986) added the protective mechanism of the handler, who acts on the potential offender. Briefly, socialisation and the production of social bonds create a handle on the potential offender that can be 'grabbed', most easily by family members and the like who are intimately involved with, and proximate to, the individual; but also by the community if sufficient ties are present (Felson, 1995). Commonly referred to as the intimate handler, Felson (1995), therefore, prefers to use the wider term ‘personal handler’. Perhaps addressing some of the criticisms that were to face opportunity theories, this inclusion extends routine activity theory beyond its original focus on the immediate circumstances of offending. However, this element of ‘those who discourage crime’ (Felson, 1995) has received the least attention within environmental criminology discourse.

The final protective mechanism to be introduced concentrates not on the original routine activity elements, but on the circumstances of their convergence in time and space. This is the place manager who, by their actions, can affect how and whether all the other elements come together (Eck, 1994). The role of place managers is considered particularly important to those who study place-based crime patterns (Eck & Weisburd, 1995), particularly those focusing on crime within establishments (Eck et al., 2007), as managers have at least some control over layout, image/reputation, security provision, internal conditions (e.g. temperature, lighting), services and goods provided, routine practices, recruitment (of staff), access, and acceptable conduct (Madensen, 2007: Madensen & Eck, 2008). They may also influence the architectural design and even the location or zone within which the premises is situated (although this is less likely).

With these additions, the routine activity explanation evolved into a crime triangle of offender, target and place (the ingredients of crime) inside a supervisor triangle of handlers, guardians and managers (Felson, 2008), all of whom can operate at varying degrees of responsibility, from those most invested in discouraging crime (family members, property owners) through to those with no personal ties or occupational responsibility (Clarke, 1992). In terms of place, Felson (1995) adapts Clarke’s (1992) four levels of responsibility to propose: personal responsibility by owners or their intimates; assigned responsibility of employees (with a specified role to look after the premises); diffuse job responsibility of other employees not specifically tasked with crime prevention; and general responsibility of “any bystander or visitor whose presence discourages crime or who notes an illegal activity that is or might be occurring there” (Felson, 1995:56). As noted above, there is some debate as to the effectiveness of crime prevention at different levels of responsibility. This is raised again below, and further research is required before we can draw definitive conclusions (see for example, Reynald, 2010; 2011).

But how does the motivated offender come across these targets, and what of those that are not stationary? How might we explain why a supervisor is or is not present? In other words, what are the routine activities from which this theory got its name?

As Felson has consistently stated, the routine activity approach is a theory of *crime and everyday life* (Felson, 2002). Routine activity theory explains spatial patterning of crime as resulting from the way individuals move about their environments according to their routine activities; those regular patterns of work, leisure and home commitments and the paths between them. These ''routines' are a part of the ecology of everyday life (Felson & Cohen, 1980; Felson, 2008), not some specialised, pre-planned crime 'trip', but the things that we all do on a regular basis: go to work/school, the gym, visit friends, go out for a meal and spend time in our homes. Some of these will be obligatory activities (e.g. the need to supply income) and some will be discretionary (e.g. leisure) (Chapin, 1974). Our routines are spatially constrained by the locations of these stable anchor points and the journeys we make between them, as well as temporally constrained by the duration of our journeys, activities and commitments to others (Chapin, 1974; Cullen & Godson, 1975; Ratcliffe, 2006).

As proposed by the crime triangle, then, it is at points where the routine activities of motivated offenders and suitable targets *intersect* that crime opportunities are presented. If these intersections do not occur on the routine activity pathways of capable guardians, if personal handlers have not exerted effective control and if the place of intersection is not appropriately managed, then such opportunities will be exploited. At the meso level it is where many offences are committed that is of interest (in producing, for example, a crime hotspot). Excluding the more remote supervision offered by personal handlers, such areas can be explained as those locations that are ineffectually managed and where the routine activities of *many* motivated offenders and *many* suitable targets overlap, but those of capable guardians do not.

Routine activity theory, then, explains how crime and crime patterns are a product of people's routines; their everyday lives. What it does not address is how these routines manifest as movement within a given environment. How do people move from one activity to another? Further, how do potential offenders recognise opportunities (or guardians) and what forms of place management are conducive to crime prevention? Theories being developed independently at around the same time (Felson, 2008) would start to provide answers to these questions and together form the basis of 'opportunity theories of crime'. Crime pattern theory would address movement and environments, whilst the rational choice perspective would seek to explain how offending decisions are made. It is to the former, that I now turn.

2.5 Crime pattern theory

As already stated, environmental criminology was really the first approach within the discipline to consider the place of crime other than as a sociological entity. In doing so, Brantingham and Brantingham (1981; 2008) highlighted the role of the 'environmental backcloth' against which crime takes place. Woven through this backcloth are opportunities and constraints, boundaries and temptations. A key aim of environmental criminology is to try and understand the role this backcloth plays in both the formation of crime opportunities and their spatial and temporal patterning. But these are complex phenomena, for which simple, flexible rules are required (Brantingham & Brantingham, 2008), thus ‘crime pattern theory’ was proposed (Brantingham & Brantingham, 1984; 1993a).

The basic tenet of this approach is that “Crime is an event that occurs when an individual with some criminal readiness level encounters a suitable target in a situation sufficient to activate that readiness potential” (Brantingham & Brantingham, 1993a: 266). This simple statement, of course, becomes much more complex in practice and Brantingham & Brantingham (2008) now present crime pattern theory as consisting of eight rules, which can only be briefly considered here. All of these rules are premised on the idea that crime patterns are the result of aggregations of individual activities. For many people these will be non-crime routines, and even for those who do offend, much of their everyday life will be spent on non-crime activities, but all of these trips will be affected by the environment in which they occur (Brantingham & Brantingham, 1993a; 2008).

Rule one states that as individuals carry out activities they make decisions and that when these activities (thus decisions) are repeated frequently, such decision processes become regularised (Cullen & Godson, 1975); thus an “abstract guiding template” is created. For crime decisions, these are referred to as crime templates (Brantingham & Brantingham, 2008: 80). This rule can be exemplified by the concept of driving to work ‘on auto-pilot’.

Rule two recognises that most individuals are part of a network of family, friends and associates all of whom will impact each other’s decisions (i.e. the processes in rule one) (Brantingham & Brantingham, 2008). This rule allows for the effects of peer pressure and informal control on decision making as well as introducing opportunities and imposing constraints on the activities available to the individual. Rule three allows for the aggregation of individual patterns into ‘‘typical’ patterns of decision processes and crime templates (Brantingham & Brantingham, 2003a).

Rule four states that crimes are committed when there is a ‘triggering event’ and a process locating a target or victim that fits the individual’s (or network’’s) crime template. When crimes are committed, the experience adds to the bank of knowledge, thus affecting future actions. This suggests that negative crime experiences will lead to some form of adaptation, which may include engagement in non-crime activities (Brantingham & Brantingham, 2008).

Rules five and six are key to crime pattern theory, and are the most often cited elements of it. Rule five states that the range of daily routine activities undertaken by an individual occur in activity *nodes* (e.g. home, work, shops) and are reached by taking the normal *pathways* between them (Brantingham & Brantingham, 1993b). Within these larger movement patterns, are micro-activities and movements, such as the way we move about in our homes or work places (Brantingham & Brantingham, 2008). These micro movements are of particular importance to place-based research, especially that focusing on small spaces, such as street blocks or individual premises. The nodes visited (at whatever scale of resolution) and the paths taken between them are also regularised: these are the routine activities of everyday life that Cohen & Felson's (1979) approach was constructed around, and it is the verbalisation of these through pattern theory that make these two approaches complimentary. Aggregating these routine activities produces an individual's activity space, around which is their awareness space. This is the space about which we are cognitively aware because of the mental maps we hold from our routine activities. Such spaces are usually within visual range of activity spaces (Brantingham & Brantingham, 2008).

In terms of crime, rule six adds that those who commit crime have *normal* spatio-temporal patterns of movement and crimes, therefore, will be committed as part of their normal routine activities, that is to say near their normal activity and awareness spaces where they feel most comfortable (Brantingham & Brantingham, 1993a; 1993b; 2008; Rossmo & Rombouts, 2008). This can also be related back to routine activity theory when we predict that crimes will occur where and when the activity spaces of potential offenders overlap with those of suitable targets (including the activity spaces of suitable victims). Crimes will be most likely to occur where and when these spaces do not overlap with the activity spaces of capable guardians. Further, areas of high crime concentrations, such as hotspots or crime generators (discussed in chapter 3), will be those places where the activity spaces of large numbers of potential offenders and suitable targets overlap.

Rule seven recognises the importance of this overlap in creating offence opportunities and completes the theory by explaining that these opportunities will be exploited when the "potential offender's willingness to commit a crime has been triggered and when the potential target or victim fits the offender's crime template." (Brantingham & Brantingham, 2008: 87).

Finally, and of importance to the current study, rule eight states that all of the other rules operate against an environmental backcloth (Brantingham & Brantingham 1993a; 1995; 2008). The nodes, paths, activity and awareness spaces will all be influenced by the built environment within which they are produced or situated. Thus, we cannot understand or predict how people will use and move about the space they inhabit, or when and where crime opportunities will be formed and potentially exploited, unless we study the urban (or suburban or rural) form that encompasses them and the opportunities and constraints it creates.

Much research has been carried out attempting to test the rules of crime pattern theory. For example, journey-to-crime research can be seen as driven by the above propositions and from this we now have much greater knowledge of offender movement patterns and the areas they feel most comfortable offending within (see, for example, Chainey & Ratcliffe, 2005; Rengert, 1992; Rossmo, 2000, cf. Smith, Bond & Townsley, 2009). Further consideration of such research is beyond the scope of this chapter, but is discussed later in relation to spatial patterning of offences.

Crime pattern theory clearly attempts to deal with offender's criminal behaviour, as well as the places where this is played out. However, it still fails to fully deal with the process of decision making, without drawing on a further approach: the rational choice perspective.

2.6 Rational choice perspective

All approaches within environmental criminology draw on explanations of crime as the end point of decision making processes. These may be conscious, subconscious or both, but they are not random and are therefore predictable (Brantingham & Brantingham, 1993a). In other words, why is a crime opportunity exploited or not? Rational choice perspective attempts to address this from a neo-classical perspective, stating that any person, when recognising a crime opportunity, will make a cost-benefit calculation that results in either taking the opportunity (rewards outweigh risks and effort) or not (risk or effort outweigh the benefits) (Cornish & Clarke, 1986).

Within the rational choice perspective, crimes are not seen as the outcome of some fixed (possibly pre-determined) criminal motivation, but rather as the outcome of similar motivations, desires and preferences that are taken into account by all of us when we make decisions. Such decisions are also affected by the opportunities and constraints presented by a given situation at a given time (Cornish & Clarke, 2008). So offenders are essentially rational decision-makers but, again like the rest of us, these processes are affected by multiple choices over considerable time periods (Gottfredson & Hirschi, 1990), thus the rational choice perspective is a limited model of decision making focused on the criminal event (Cornish & Clarke, 1986 cf. Ekblom, 2000 who seeks to address this). Cornish & Clarke (2008) describe it as a heuristic device (not a complete criminological theory) that looks at offending as present-centred behaviour that is influenced by the immediate environment.

As would be expected, returning to a rational actor model of offending has resulted in significant criticism being levelled at this perspective (and indeed its proponents). The purpose of this chapter is not to revisit this criticism, but to lay the foundations for the forthcoming literature review. Therefore, I will only consider those critiques that have directly impacted upon the development of this perspective as I further outline the key principles associated with it.

The rational choice perspective is clearly a micro theory of crime as it deals with the decision-making processes of individuals. However, as with crime pattern theory, the summation of these decisions can produce 'typical responses'. But how rational are potential offenders when faced with the opportunity to commit crime? Cornish & Clarke (1986) view all human behaviour as purposive and crime is no exception. However, the benefit of the criminal act is not limited to monetary or tangible gain, but also includes psychological benefits such as enhanced status, gratification or excitement. Ultimately, this recognition leads to the added consideration of situational precipitators (Wortley, 2001) to that of situational opportunities. Wortley's critique of the focus on only "…half of the situational crime prevention story" (Wortley, 2001: 64) led to his proposals for four situational precipitators. These were prompts, pressures, permissions and provocations. In line with the 4x4 matrix of situational crime prevention (which focused specifically on environmental opportunities), these were presented as 16 cells, suggesting how such precipitators could be controlled or reduced. Controlling prompts, included such things as controlling triggers (for example, restricting pornography) and reducing inappropriate imitation (such as may be achieved through rapid repair of vandalism and removal of graffiti). Controlling pressures can be exemplified through encouraging compliance (such as by encouraging participation in rule-making) and reducing anonymity. Permissibility could be reduced through such approaches as rule setting (with clear codes and reminders about acceptable behaviour being posted) and clarification of consequences (of undesirable behaviours). Finally, reducing provocations could be achieved by reducing frustrations (for example, through efficiencies and fair practices), reducing crowding (such as found in licensed premises) and controlling environmental irritants (such as smoke and very loud music).

Ultimately, a number of these ideas were incorporated by Cornish and Clarke (2003) to extend the original situational crime prevention matrix. First conceptualised to propose methods that would increase risks and effort, whilst reducing rewards (a 12-cell matrix) this became a 4x4 classification when removing excuses was added (see Clarke, 1992; 1997). Responding to Wortley's (2001) criticism, this is now a 5x5 matrix that also considers the reduction of provocations (Cornish & Clarke, 2003).Therefore situational precipitators are now recognised as a legitimate part of opportunity-reduction, targeted at reducing the crimes of a rational offender, though not just as they respond to the quality of the opportunities to offend that are signalled by the environment, but also because such environments actively bring about this (rational, criminal) behaviour (Wortley, 2001).

Thus, crime may be seen as a 'rational' response not only to indicators that it is 'worth' committing, but also to frustrations, annoyances, perceived injustices and threats. The behaviour may also be rational*ised* through excuse-making (which may be seen as an incorporation of the work on techniques of neutralisation (Sykes & Matza, 1957).

Further, despite criticism levelled at this approach, Cornish & Clarke (1986) always saw rationality as 'limited' or 'bounded'. The perspective recognises that decision-making is not completely objective, that not all possible options are available, that decisions are constrained in many different, and changing ways, and, importantly, that often the information available to the decision-maker is limited. Thus decisions are subjectively rational to that individual, in a given set of circumstances that exist at a given time for the information available. As such, decisions may seem *ir*rational to an objective observer (particularly in hindsight), but this does not presuppose they actually were. The constraints and influences experienced by the decision-maker (our potential offender) may be immediate or distant, may be situational/environmental or social. Thus here, remote 'causes' and the concept of affordance also have their role to play in considering how the decision-maker understands, constructs, recognises and assesses available opportunities and whether or not to exploit them. (Ekblom, 2001; Ekblom & Sidebottom, 2007). It is no surprise that offenders often make what may seem like poor decisions, if the information they have about the situation is limited or it is variously interpretable. As Roach and Pease (2013: 83) summarise: "Affordance…is best thought of as the psychology of what actions come to mind in a certain setting." So some people may see an unattended bag as a security threat, some as a crime opportunity and others as lost property. When seeking to explain offence commission (and crime placement), therefore, we are interested in those for whom it affords a prospective theft and how they *then* decide whether to commit this or not. Again, affordance (and subsequent decision-making) will be affected by proximal and distal causes and circumstances. It is also affected by individual cumulative experiences, in other words, experiential learning (Matsueda, Kreager & Huizinga,2006).

Another element of the rational choice perspective is the role of 'crime scripts' (Cornish, 1994); the step-by-step procedures for committing a particular crime. Crime scripts can be seen as somewhat similar to the templates referred to by crime pattern theory, but scripts are much more detailed, consisting of the decisions, actions and resources required for each step towards committing a crime, such as the preparation, entry, target selection and so forth (Cornish & Clarke, 2008). As such, this approach gives yet further insight into how to *think thief* (Ekblom, 1997) and where and when we can interrupt this sequence to prevent the successful commission of a crime. For considering crime placement, understanding scripts will help explain why and *how* certain opportunities are taken.

Overall, then, rational choice perspective and the associated concepts introduced here, provide the ‘missing link’ for the two previously considered theories: how the offender makes the decision to exploit a crime opportunity. We can also see that this perspective offers considerable scope for crime prevention, which is not surprising given that it was formalised to underpin the contemporaneous development of situational crime prevention (Cornish & Clarke, 2008).

2.7 Conclusion

Taking the above theories and responses into account, then, it can be seen that crime opportunities are considered to be the result of the overlap of the activity spaces of potential offenders and victims, or the locations of suitable targets, as people go about their daily routines, constrained spatially and temporally by the environmental backcloth, their routine obligations, and their knowledge and (spatial) awareness. Interpretation that a set of circumstances affords a crime opportunity varies by individual and situation, and is affected by remote and proximal, chronic and acute ‘causes’. Suitable opportunities are then exploited when the offender is not appropriately controlled, targets are not guarded and places are not effectively managed, thus providing sufficient rewards for an acceptable degree of risk and effort, in the absence of neutralising excuses or as a result of provocations. In response, crime can be prevented by shifting the balance in this decision making process or preventing the overlap of activity spaces in the first place.

It should be clear, therefore, that if these theories are correct, we would not expect all places or times to provide equally attractive opportunities for crime, rather we would expect to see some degree of concentration. The following chapter explores the existence of crime concentrations across different entities (people, products, times and spaces) and at different resolutions. The phenomenon of spatial crime concentrations, notably at the micro-level, receives the most attention as the central concept for my research.

Chapter 3: Crime patterns and research findings

3.1 Introduction

It is now a well-accepted fact in criminology that crime is not distributed either randomly or evenly across a number of dimensions. Not only does crime concentrate on certain people, products, times and places, but these concentrations tend to be particularly intense in a small sub-section of the sample. This phenomenon is not peculiar to crime distribution and is often referred to as the 80-20 rule: where 20% of some things are responsible for 80% of the outcomes (Clarke & Eck, 2003). This pattern can be identified via ranked, cumulative percentage tables, although the proportions are rarely actually 80-20. This pattern is commonly referred to as a ‘J-curve’ (Eck et al., 2007) because of its resemblance to a reclining letter J when data are aggregated by the chosen unit of analysis and then ranked so that the unit with the highest number of offences is nearest the origin.

In this chapter I summarise the literature relating to crime concentrations, very briefly considering people (both offenders and victims), products or goods, and times (seasons/times of the year, days of the week and times of the day) before focusing on concentrations in space. Of most relevance, thus receiving the most attention, is the literature on concentrations at the micro-level, particularly that relating to the concept of risky facilities - the focus of my research. Crime attractors, crime generators, (place-based) repeat victimisation and the proposal of a ‘law of crime concentration’ at place are also considered, in the context of the risky facility phenomenon.

3.2 Crime Concentrations

Concentrations in the distribution of crime are often referred to generally as forms of repeat victimisation (RV). Farrell (2005) presents different types of RV and distinguish: target repeats (however defined: person, place, vehicle, etc.); tactical repeats (e.g. theft of hot products, use of similar MO); spatial repeats (when a particular place or space suffers disproportionate levels of crime); crime-type repeats (which could be seen as homogeneity of offending within a criminal career and will also likely lead to target repeats); temporal repeats (crimes committed in quick succession and related to spatial and tactical repeats); and offender repeats (those committed by the same offender). However, this categorisation contains quite significant overlap (e.g.

concentrations of crime in particular establishments could be target (place) or spatial repeats), whilst grouping many different areas of interest together as target repeats. Therefore, I instead consider concentrations of crime across different entities, namely: People; Products; Times; and Places, with the latter being of greatest relevance to this study. Repeat victimisation as a concept, and the findings of RV studies, are included and discussed where relevant.

3.2.1 People

Crime concentrations in 'people' apply to both offenders and victims. Since Wolfgang's famous Philadelphia cohort study identified that 6% of his sample accounted for more than half of all the crimes committed (Wolfgang, Figlio & Sellin, 1972), the concept of the chronic or prolific offender has been established. This highly disproportionate contribution of offences within a small sub-set of offenders has been seen in multiple studies, particularly in the field of developmental criminology and criminal careers research. Of particular interest, there are examples of this being demonstrated using the Gini coefficient (Fox & Tracy, 1988; Piquero, 2000; Piquero & Buka, 2002), a method similar to one that is employed as part of the current research.

When crime is over-represented amongst certain types of victim, we may refer to these as 'at-risk' groups. Such categorisations are usually carried out by gender, age, employment type, social background and so forth. More specifically, when a particular individual or household is offended against on multiple occasions this is referred to as repeat victimisation (RV) (Pease, 1998). The number of offences and the time period across which these must occur for the 'label' to be applied is open to debate and the extent of RV varies by crime type (Clarke & Eck, 2003; Farrell, Sousa & Weisel, 2002; Farrell & Pease, 2001; Farrell & Pease, 2008; Weisel, 2005), however research consistently shows that those individuals (or households) who have been victimised once are at greater risk of further victimisation (e.g. Farrell & Buckley, 1999; Forrester, Chatterton & Pease, 1988), that this risk continues to increase with each subsequent crime they suffer (Ellingworth, Farrell & Pease, 1995), and that it is greatest shortly after the event (e.g. Polvi, Looman, Humphries & Pease, 1991). Going beyond the scope of usual victimisation studies, Townsley and Farrell (2007) have even found the phenomenon holds for victimisation of prison inmates (by other inmates or staff). Interestingly, it has also recently been noted that over the period of the so-called 'crime drop' experienced across much of the 'western' world since the early to mid-1990s, the concentration of crime in a small proportion of repeat victims has actually increased (in England and Wales, at least) (Ignatans & Pease, 2015). In other words, as crime decreased significantly, the extent of concentration - of unequal distribution - increased. In fact, it appears that those who experienced the greatest amount of crime also experienced the greatest reduction over time, but this was not significant enough to re-balance the distribution more equitably, so that repeat victimisation, for both personal and property crime, is now even more concentrated than it was before ‘the drop’ (Ignatans & Pease, 2016).

Opportunity theories can be used to explain the phenomenon of RV as either boosts or flags (Everson & Pease, 2001) and this may also be applicable to other forms of crime concentration. Boost (or event dependent) offences occur when repeat offenders continue to predate against successful targets, or tell others about them (this could be linked more specifically to the rational choice perspective, as the perceived risks, rewards and efforts are based on greater information than a potential offender would otherwise have, and the balance is now assumed to be favourable). Flag (or risk heterogeneous) offences occur when a target is victimised by many different offenders because something about it suggests it is particularly attractive or vulnerable; as such this could be linked specifically to both rational choice (e.g. an insecure property is seen as a desirable target to multiple rational offenders) and to the routine activity theory, because the target is in close proximity to a large number of offender activity spaces. Further research into this explanation, and the temporally limited nature of increased RV risk, supports these two explanations, suggesting that event dependence may be the key explanatory factor (see, for example, Clarke, Perkins & Smith Jr, 2001; Everson and Pease, 2001; Kleemans, 2001; and Tseloni and Pease, 1997) but that this may vary by location, even between proximate areas (Morgan, 2001).

Later research into spatial RV suggests that victimisation can also increase the risk to nearby targets (e.g. Johnson & Bowers, 2004; Townsley, Homel, Chaseling, 2003), perhaps because similar types of suitable target tend to be concentrated in space and time (flag) or because a successful offence breeds further nearby attempts (boost) (Bowers & Johnson, 2004); or, in relation to shootings (and perhaps gang related activity), as retaliation (Ratcliffe & Rengert, 2008).

Of course the work on repeat (and near-repeat) victimisation of individuals and households is much more extensive than has been summarised here, but further consideration in relation to people and offences in residential settings is beyond the scope of this review.

3.2.2 Products

Crime is also concentrated in products, with certain items being stolen much more frequently than others. Initially Cohen and Felson (1979) used the acronym VIVA to highlight the value, inertia, visibility and access that would identify desirable targets for predatory crime. This was later replaced by Clarke's (1999) term 'hot products' with their CRAVED attributes of concealability, removability, availability, value, enjoyableness and disposability, in relation to targets of acquisitive crime. Commonly stolen items that fit this definition include cash, mobile telephones (see, e.g. Mailley, Whitehead & Farrell 2006; Whitehead, Mailley, Storer, McCardle, Torrens & Farrell, 2007) and consumer electricals (e.g. Armitage & Pease, 2008). Research by Wellsmith and Burrell (2005) has shown that those items most often stolen in domestic burglaries neatly fit this acronym, whilst Sidebottom and Bowers (2010) conclude the same about items stolen through bag 'dipping'. More diversely, this concept has also been applied to species at risk of poaching and trafficking, notably parrots (Pires & Clarke, 2012) and livestock theft in Malawi (Sidebottom, 2013).

This concept, like the other dimensions of crime concentration under discussion, can also be explained through the application of opportunity theories. The rational offender will make target choices based on maximising rewards (whilst limiting effort and risk) and some products will satisfy these requirements more than others. Further, routine activities at micro, meso and macro-levels will affect the availability and accessibility of certain products differently to others, such that these are encountered more often, are considered of greater value, or are less well protected for example, again increasing the likelihood they will become targets of crime. This also relates to the role of 'market forces', such that products at an optimum level of availability and value (mass market) will likely be more desirable (see Wellsmith & Burrell (2005) for a test of this proposal). More recently, a further acronym – AT CUT PRICES – has been developed to identify (and explain, with reference to opportunity concepts) 'fast-moving' consumer goods that are more likely to be theft targets (Gill & Clarke, 2012). Here, key features of such goods are that they are affordable, transportable, concealable, untraceable, tradable, profitable, reputable, imperishable, consumable, evaluable and shiftable. What is important, then, is being able to steal and sell-on items with low risk, being able to assess the likely profit to be realised and knowing that these goods will sell.

The consideration of hot products also particularly highlights how crime concentrations in different dimensions are likely to overlap, or can even be explained by one another. For example, people who suffer repeat acquisitive crimes may well be those people who are most likely to own (or be least able to protect) hot products. In the same way, spatial concentrations of acquisitive crime (discussed below) may well reflect, at least in part, the spatial distribution of CRAVED/AT CUT PRICES goods across the shops that sell them, the residences that house them and the facilities that use them (e.g. schools may be targeted for burglary as they often contain large numbers of computers (Dedel, 2005)).

3.2.3 Temporal concentrations

Temporal concentrations of crime can occur at a number of different resolutions. Clearly the crime rate (at any geographic unit) changes over time and this is of much concern to governments, policy-makers and the media. Indeed, as was noted in chapter 2, it was the increase in domestic burglary across the USA in the 1960s that Cohen and Felson (1979) used to exemplify their ‘new’ routine activity approach to crime explanation. However, there are also significant, cyclical concentrations across seasons, days of the week and times of the day.

Much research into crime seasonality focuses on *temperature* and other weather conditions (e.g. Cotton, 1986; Rock, Greenber & Hallmayer, 2003, *cf* Cheatwood, 1995; Rotton & Cohn, 2000; 2001). Some researchers have attempted to combine such findings with opportunity-based explanations for offending, considering the impact of different weather conditions on issues such as visibility, guardianship and social interaction (Cohn, 1990; Landau & Fridman, 1993; Tompson & Bowers, 2015).

Perhaps linked to considerations of temperature, evidence has been presented of seasonal offending patterns for homicide (McCleary & Chew, 2002) but findings are contradictory (Cheatwood, 1995). Offences requiring a ‘get away’’ in public are more likely to take place when there are more hours of darkness, thus reduced visibility (all other opportunity factors being equal). For example, more bank robberies occur in the winter than at other times of the year (Weisel, 2007; van Koppen & Jansen, 1999), whilst Bentley & Kerr (2008) found a summer peak in ‘sneak’ burglaries resulting from windows and doors being left open during warmer weather, significantly reducing the effort required on the part of the offender.

The days and times during which offending takes place are related to opportunity and these patterns appear to be strongly linked to routine activity and crime pattern theories as the conjunction of offender activity spaces with suitable opportunities will only occur at certain times, due to various environmental, operational and individual constraints (Ratcliffe, 2006).

The days on which offences take place are often, although not always, known. A number of studies have shown temporal concentrations of offending by day of the week. For example, around one quarter of bank robberies in the US take place on a Friday (Bruce, 2004; Weisel, 2007). This pattern is likely related to longer opening hours on Fridays, ‘pay day’’ quantities of cash, or a desire for money for the weekend (Weisel, 2007). Reflecting changing opportunities, weekend robberies have increased as more branches have opened on these days (Bruce, 2004).

Taking a completely different type of crime, Kposowa and Breault (1998) found that homicides in the US (compared to deaths from other causes) were highest on Saturdays, followed by Fridays and Sundays, explaining this pattern with reference to drinking behaviours. Similar patterns for assaults have been shown across the literature on alcohol-related violence (e.g. Burrell & Erol, 2006).

It is also proposed that all crimes have different peaks and troughs at different times of the day (Felson & Poulsen, 2003). Analysing concentrations by the hour of offending is more difficult than for other temporal patterns, as such accurate information is often unknown (much police recorded crime data contains ‘time from’ and ‘time to’ fields as the exact moment of offending is not known). Methods of dealing with such ranges have been proposed (Ratcliffe, 2002), but the published research remains limited.

Bank robberies are an example of the type of offence for which the exact time is generally known. Research from various countries has shown the morning to midday period to be the peak time for such crimes (see Weisel, 2007). Possible explanations for this are professional robbers targeting times which are least busy or when the greatest amount of money is on the premises (Gill & Matthews, 1994; Van Koppen and Jansen, 1999), whilst opportunist thieves may prefer more customers present to deflect attention away from them (Weisel, 2007). An Australian business crime survey has shown service stations to be most vulnerable overnight (and in the late evening) (Taylor, 2002) and convenience store robberies are reported to occur most frequently during similar periods (but limited by opening hours) (Altizio & York, 2007). As a final example, Smith, Bowers & Johnson (2006) found temporal concentrations for bag thefts in a pub chain in Westminster, London, with self-reported thefts peaking around 6pm (as reflected by police recorded crime data), possibly reflective of after-work drinking when premises are not only busy, but those present are more likely to have bags, briefcases, laptops, and so forth, with them (compared to those who have returned home and then gone out later in the evening).

Temporal concentrations in crime, then, appear highly related to both routine activities and rational decisions about the ‘best’ time to offend. As Brantingham and Brantingham (1993a; 1995) state, crime generators and crime attractors (discussed below) may not hold this status at all times or for all crime types. As such, any consideration of spatial concentrations must also take into account this further dimension. Although my research focuses on spatial concentrations, therefore, temporality will also be considered.

3.3 Spatial concentrations

As with the other dimensions of crime concentration considered above, spatial concentrations of crime are apparent at all levels of resolution, from international to regional, within cities to particular Basic Command Units (BCUs), from street corners to specific premises (Brantingham, Dyreson & Brantingham, 1976; Sherman et al., 1989). Opportunity theories again clearly predict and help explain the existence of spatial concentrations of crime. In particular, variations in the environmental backcloth, the way people use and move about within the environment, particularly their awareness spaces, and the routine activities that people engage in, will all result in a topographical map with peaks and troughs of (discovered) opportunities, to be exploited by the rational offender. Those locations with the most suitable opportunities and/or those where opportunities are most likely to be come across (and afforded) will, we would expect, have the most crime.

In addition, we must consider the 'supply' of potential offenders. The literature suggests that, generally, crime trips are short, at least in the aggregate, such that there appears to be a distance decay function, with more crimes occurring closer to an offender's or several offenders' anchor points (which are usually assumed to be their homes (Block et al., 2007; Rengert, Piquero & Jones, 1999; Wiles & Costello, 2000), though with some evidence of a 'buffer' of non-offending around this anchor point (Rossmo, 2000; Rossmo & Rombouts, 2008 cf. Block, Galary & Brice, 2007; Canter & Gregory, 1994). This knowledge is important when trying to explain spatial (and temporal) patterning of offences, as even the most suitable targets may not experience high offending rates if they are long distances away from any offenders (and therefore their activity spaces). In other words, as well as inherent attractiveness, a target's suitability will also be affected by its proximity to a supply of offenders. Thus when seeking to explain the existence of crime concentrations, researchers often also consider the wider environment, including areas of high concentration of offender residences, or other locations that might supply offenders, including schools or even other high crime locales. Further research on journey-to-crime, and the importance of this to spatial concentrations, is considered further (albeit necessarily briefly) below.

So opportunity and the environmental backcloth seem to be able to suitably explain, at least on a superficial level, why crime concentrates spatially. Of course, the challenge that remains is in determining what makes something a suitable and discovered opportunity and how people move about the environment. With such knowledge we can then further test our theories and make suggestions for reductive interventions.

First, however, we need to fully explore and understand these concentrations and associated phenomena. In an early comparison of concentrations of crime amongst people (offenders and victims) with concentration amongst places, Spelman and Eck (1989) concluded spatial concentration was more likely. Further, as noted earlier, Brantingham and Brantingham (1981) called for a greater focus on place as a neglected dimension in the study of crime and victimisation. Patterns relating to where crime concentrates are an important issue to explore, and as such, a growing body of work - often referred to as place-based crime/criminology or crime and place - has been established (Weisburd & Eck, 2018). Analysis in this field has consistently demonstrated that crime concentrates spatially, so that a small proportion of places contribute a large proportion of crime. The phenomenon is held to be almost ubiquitous, with Wilcox and Eck (2011) referring to it as the 'Iron Law of Crime Concentration'. Indeed, most recently a systematic review and meta-analysis of spatial crime concentration (based on 44 studies) has provided further support to claims that crime is spatially concentrated in a relatively small proportion of places, although the extent of this concentration, or at least the extent to which it manifests, seems to vary by unit of analysis, crime type and location (as well as over time) (Lee, Eck, SooHyun & Martinez, 2017). Also recently, Weisburd (2015) has called for further research in the field of 'the criminology of place', arguing that it remains under-explored and provides fertile ground, particularly for early career researchers. In particular, he advocates studying micro-places (which he refers to as microgeographic hotspots) and testing his proposed concept of the 'law of crime concentration at place': that a narrow bandwidth of places (equating to a small percentage of the total, but to as yet be determined) accounts *universally* for a particular proportion of crime (he proposes looking at 25% and 50%).

This next section, then, presents the extant literature from the field of crime and place, with a particular focus on smaller areas as advocated by Weisburd (2015) and as most relevant to my research. In particular, I further consider the proposed law of crime concentration at place and, firstly, concepts associated with spatial concentrations of crime: hotspots, crime attractors and generators, and risky facilities.

3.3.1 Hotspots

Simply put, hotspots are geographic clusters of crime (Braga, 2008). Their practical importance for policing, specifically crime analysis and thus resource targetting and problem-solving, has increased since the advent of affordable, user-friendly mapping or GIS (Geographic Information System) software. This makes their identification quick and easy, with enforcement and prevention resources targeted accordingly (Eck, Chainey, Cameron, Leitner & Wilson, 2005). As Anselin, Griffiths and Tita (2008) point out, criminologists tend to define the hotspot as an area of proportionally greater crime incidents than other similarly sized areas in a city, and that these areas are usually smaller than neighbourhoods (or police 'beats'). As with the other concentrations considered here, hotspots within a city can account for a disproportionate amount of crime. For example, Sherman et al. (1989) famously found that just over 3 per cent of street addresses and intersections contributed over 50 per cent of police calls for service in Minneapolis. Knowledge of where such hotspots exist, can be used to better target police resources, including patrols (Sherman & Weisburd, 1995), as well as to identify areas for further analysis as part of the problem-oriented policing approach (Goldstein, 1990; see also, amongst others, Spelman & Eck, 1987 or for a more recent view see Eck & Gallagher, 2016).

Hotspots remain a useful concept, albeit we could debate exactly how they are or should be defined, as well as the most appropriate ways in which to identify and display them (for different issues and purposes, as various techniques for identifying hotspots or crime clusters can be employed; see, for example, Block, 1995; Canter, 1997; Eck, et al., 2005; Grubesic, 2006; Hirschfield, Yarwood & Bowers, 2001; also the collected chapters in Goldsmith, McGuire, Mollenkopf & Ross, 2000 and in Hirschfield & Bowers, 2001). However, once we have established that crime concentrates in this way, thus added to our growing evidence about spatial concentration, I believe that *identification* of hotspots remains of most use to practitioners. What is 'going on' within hotspot areas is of most interest to criminologists: the features, premises, activities, and so forth, that result in their existence. Theoretically, hotspots (and the times at which they are 'hot') may be caused by instances of repeat victimisation (of people) or because they are, or they contain, areas where many activity spaces overlap and there is a ready supply of suitable targets (and possibly a lack of guardianship or effective place management) (Brantingham & Brantingham, 1999). Thus the environment plays an important role in producing these concentrations, by providing the nodes and paths that create such spaces and holding the targets against which crimes are committed. Of particular importance for my study is the proposed impact of facilities on the production of hotspots (as has been noted since the earliest studies on hotspots, such as Sherman, et al., 1989). As will be seen, below, the presence of hotspots has often been attributed to the crime generating (or attracting) nature of facilities, such as schools or licensed premises, situated within them. It is to crime in premises, therefore, that I now turn.

3.3.2 Attractors and Generators

The idea of a crime attractor or generator has grown out of crime pattern theory and relates to a particular type of spatial crime concentration. Brantingham and Brantingham (1995) define crime generators as areas that produce crime as a result of attracting large numbers of people for reasons unrelated to criminal motivation. In other words, people visit such places as part of their non-criminal routine activities, but because some of these people will be potential offenders and some of them may be suitable victims (or the places themselves may offer suitable targets) and because the environment is conducive to crime (e.g. through poor place management or a lack of capable guardianship as discussed in chapter 2), crime opportunities are presented and taken. Although Brantingham and Brantingham (1995) do not state that such places need to be 'facilities', much research into crime generators interprets them as such, perhaps because particular buildings or complexes are more easily definable than public open spaces (although the term is now used so widely that almost any area seen as problematic may be labelled as a 'crime generator', as is apparent from a search on Google).

Certain types of facility have received much more research attention, perhaps because this area is strongly theoretically driven (the distinction of the crime generator from any other form of crime cluster is an entirely theoretical one). Such research tends to either (a) propose that a particular type of premises (e.g. a school) is a crime generator and then gather data to test the influence of such premises on local crime rates (e.g. Roman, 2003; Roncek & Bell, 1981); or (b) compare area crime rates and then attempt to explain these by reference to premises that may be (theoretically) considered as crime generators (e.g. Brantingham & Brantingham, 1995). The sophistication of the methodologies employed in doing this varies, but none of the published research shows that the presence of a particular facility *causes* higher crime rates, nor does it test the contribution of *all* facilities within in an area (only those selected by the researcher). One problem of using these methods, for example, is that there is no systematic inclusion of, for example, libraries or churches as crime generators because it is assumed that they are not, but if this has not been proven, then how do we know our assumptions are correct? I consider this issue more below. Taking a somewhat different approach (looking at changes in crime over time), Taylor (2002) presents Australian recorded crime data for robberies that show large percentage increases in the number of offences in service stations. However, the categories of facility presented are too crude to establish the degree to which such establishments are truly crime generators. In the current study I seek to address some (albeit not all) of these issues and further discuss approaches to identifying crime generators.

Brantingham and Brantingham (1995) define crime attractors as those locations to which potential offenders are drawn because they are known to be good places to commit crime. Often cited examples of such locations are local drug markets, red light districts and violent bars. However, our understanding of opportunity theories suggests that anywhere that is known to have suitable targets and/or weak place management and guardianship may attract offenders, thus shops, for example, may act as both generators or attractors (Brantingham & Brantingham,

1995) for different or even the same populations. Therefore, added to the conceptual and research issues discussed within this review, if we wish to distinguish between generators and attractors we would also need to determine offender motivation.

Research in this area has tended to concentrate much more on crime generators than crime attractors. This may be because identifying an attractor, thus the inherent offender motivation, is more difficult, but it may also be because of misuse of the terms. When Brantingham and Brantingham (1995) developed the terms crime attractor and crime generator, they gave them different definitions. The outcome was the same (high crime within facilities and, possibly as a result, surrounding areas), but the theoretical mechanism by which it occurred was different (McCord, Ratcliffe, Garcia & Taylor, 2007). Whilst it is widely acknowledged that a location can be both an attractor and a generator (either at different or the same times), this does not mean they are the same thing and because the mechanism that produces them is different, the method of reducing crime at such locations may also need to be different.

Throughout the literature, albeit with notable exceptions, these terms are confused or even used interchangeably (as if the style of the piece – not using the same word repeatedly – is more important than accuracy). Examples of this commutation appear in Roman's (2003) research into schools as generators of crime and Tilley, Pease, Hough and Brown's (1999) consideration of the different forms of generator that may produce higher rates of domestic burglary. Specifically, the list of 'offender related generators' includes: "Offenders travelling into an area specifically to burgle because of its reputation for rich pickings" (p8). Later, they also (pp29-30) refer to a football stadium consecutively as an attractor and a generator. This lack of distinction is by no means limited to these examples.

In addition to the above criticisms, most of the published research limits consideration to 'all recorded crime' or certain types of crime. I have come across very little that discusses and compares generators or attractors across a range of crimes (to identify what crime they generate or attract), a range of times/days or longitudinally. Therefore, the consistency and persistency of these types of concentration is, generally, also unknown. Research into risky facilities - which we might consider to be more specific a concept than attractors and generators, and which is discussed below, has started to address this, but a systematic approach to crime in all types of premises across the city has still not been taken.

Taking the methodological and terminological issues into consideration, however, there remains an extensive body of literature on the impact of different land uses or specific facilities on area crime rates, which will now be summarised, although the findings (in terms of referring to such facilities as generators) remain tempered by the aforementioned concerns.

As stated, certain types of facility are clearly recognised theoretically as potential crime generators and these have received much more research attention. Specifically, there is literature relating to schools as crime generators (For example, LaGrange, 1999; Roman, 2003; Roncek, 2000; Roncek & Faggiani, 1985; Roncek & Lobosco, 1983), which all use variations of the ‘impact on local area rates’ method. Generally, (state) schools have been shown to increase crime rates in their immediate area, but the distance of this effect appears relatively short.

Perhaps receiving more attention than any other type of facility are on-licensed premises (pubs, taverns, bars, and so forth). Research has consistently shown that areas in close proximity to such premises have higher crime rates (Donnelly, Poynton, Weatherburn, Bamford & Nottage., 2006; Frisbie, Fishbine, Hintz, Joelson & Nutter, 1977; Kumar & Waylor, 2002; Roncek & Bell, 1981; Roncek & Maier, 1991; Roncek & Pravatiner, 1989; Zhu, Gorman & Horel, 2004), although whether these crimes are ‘contained’ within bars and their immediate environs or spill into local neighbourhoods was not explicitly considered, and research by Block and Block (1995) has shown that the most dense clusters of licensed premises related crime did not always occur in the same spaces as dense clusters of licensed premises. More recent research on ‘problematic licensed premises’ has tended to fall under the umbrella of risky facilities and is discussed further, below.

There is also some evidence that abandoned buildings (Spelman, 1993), transit stops or systems (Kooi, 2007; Newton, 2004 cf. Loukaitou-Sideris, Liggett & Iseki, 2002), sports stadia (Kirk, 2008), casinos (Piscitelli & Albanese, 2000 cf. Stitt, Nichols & Giacopassi, 2003) and shopping malls (LaGrange, 1999) may generate or attract crime.

An alternative strategy for research into crime generators and attractors has been to analyse area crime rates in relation to multiple types of facility and not necessarily distinguish between the types of premises selected, other than to designate them as ‘mixed land use’, ‘‘non-residential’ land use/premises, ‘commercial’ land use or ‘crime attractors’ and ‘crime generators’. Again the findings suggest that the presence of theoretically identified attractors and generators may be related to areas with higher crime rates. However, this research usually fails to disaggregate these categories to determine the effects of different types of facility, and to consider that such facilities may be placed into high crime areas, as opposed to causing them.

Not only this, but such research has also displayed mixed results when considering the effect of/on area crime rates. For example, Wilcox, Quisenberry, Cabrera and Jones (2004) found that those residents who lived near more non-residential land-use (in this case business premises) perceived higher amounts of incivility within their neighbourhoods. In direct contrast Sampson & Raudenbush (2004) found no evidence of an increase in perceived incivilities in the presence of non-residential land use, but they did find an influence for the grouping of bars and liquor stores, signs of commercial building security, presence of alcohol or tobacco advertising.

McCord et al. (2007) point out there may be methodological flaws in both conclusions and develop an alternative method to test the generative and attractive effects of specified land uses. In their study, generators were represented by schools, subway stops and expressway off ramps, whilst attractors consisted of cheque-cashing stores, pawn shops, drug treatment centres, halfway houses, homeless shelters and certain types of licensed premises. The results of their rigorous analysis showed that those people living closer to a ‘crime generator’ perceived higher levels of both crime and incivility. The same was found for those living close to a crime attractor. However, in terms of victimisation *against* commercial premises, research suggests that convenience stores (Hunter, 1999), service stations and pharmacies (Taylor, 2002) are less likely to suffer robberies if they are situated in concentrated commercial areas. A flaw of the McCord et al. (2007) study is the use of perceived levels of crime and incivility. Whilst this may be important for informing intervention requirements, it does not actually show that the incidence of crime was high. Indeed, it may be that respondents perceived high crime levels because they lived near establishments with a reputation for being a ‘crime problem’ (regardless of whether or not this was unfounded).

As the McCord et al. (2007) study suggests, the idea that a facility can be a problem generator is not limited to issues of crime. Applying the same logic, that generators are facilities that attract large numbers of people, then we may anticipate disorder/incivility generators, fear generators, hate generators and, as La Vigne (2007) discusses, traffic congestion generators (possibly resulting in associated tensions). However, there is limited literature on these topics.

Crime attractors and crime generators may be useful concepts for explaining the existence of hotspots, and it may indeed be the case that certain facilities have negative effects on crime rates within their vicinity. However, I contend that the research in this field has tended to be too concerned with looking at area effects of individual, types or groups of facilities at the expense of considering crimes at those facilities themselves. It is highly likely, as suggested by the existing research, that certain types of facility or land use impact on their local areas and the recorded crime rates therein. There are also interesting and important questions raised within this field of study regarding whether such effects may be additive or multiplicative when the number or type of generative/attractive facilities is increased (Bowers, 2014), whether there are ‘toxic’ combinations of facilities or land use types and whether some facilities or land use types may have crime-neutral or, most importantly, crime-inhibiting effects, possibly in specific combinations. However, this does not negate the criticism that attributing area crime to the existence of particular types of premises (or even to specific addresses in a given location) is problematic (I discuss this further when I consider approaches to measuring premises-related crime concentrations in chapter 4, below). Nor does it negate that the terminology of crime generators (or attractors) is often, as I see it, wrongly used when referring to actual premises (as opposed to a class of facility). To some extent the more recent conceptualisation of risky facilities further highlights this latter issue, but also provides an alternative label that can be used more satisfactorily when considering high crime at/associated with particular addresses, though studies still fall foul of the former criticism. As will be explained, because of the critique presented in this section, my study initially considers the concept of crime attractors/generators, but then focuses on risky facilities as the main conceptual entity. It also takes a different approach to determining the crime associated with premises, as will be discussed below. First, however, the literature relating to risky facilities will be presented.

3.3.3 Risky Facilities

The most recent iteration of the J-curve concentration is the idea of the risky facility (Clarke & Eck, 2003; Eck et al., 2007). In this case, it is a small number of specific premises (addresses) within a homogenous set of facilities that are experiencing the disproportionate amount of crime. In other words, within a chosen area, we may identify that licensed premises tend to experience a large proportion of crime, but when the contribution from each individual establishment is considered we find that just a small number of these contribute most of the crime, whilst many of them contribute relatively little. Unlike crime attractors and generators, risky facilities by their definition require a consideration of crime within premises (as opposed to their effect on the wider area). This concept is, therefore, of particular interest to my study. Being aware of risky facilities is important for a number of related reasons. Firstly, a small number of premises may be contributing a disproportionate amount of crime (appearing in aggregated crime records). Secondly, risky facilities may adversely affect the reputation of a whole area (and other non-problematic premises nearby). Thirdly, identifying risky facilities presents a further way to focus crime prevention interventions, with the biggest gains likely to be achieved through tackling the biggest problems. Finally, risky facilities can be compared with non-risky facilities of the same class in order to test theoretical assumptions (Eck et al., 2007; Wellsmith, Birks & Donkin, 2007).

Risky facilities can be seen as a form of repeat victimisation, only in this case the ‘victim’ is the location in which the repeat offences take place. We could argue, though, that a building cannot be victimised. Indeed, most of the offences in question will also involve a victim who is a person (even if this is some remote business owner who suffers only economically), therefore the discrete concept of risky facilities remains useful.

Farrell (2005: 145) would appear to disagree when he states that “[r]estricting the term to victims (people) is inappropriate” as other targets (including business premises) can be repeatedly targeted. He goes on to suggest that the unit of analysis for determining repeats should be that which is most appropriate for crime prevention. In practice, recognising, analysing and responding to any type of crime concentration is of course important, however this does not mean the term repeat victimisation needs to be used. In fact, I would argue that the risky facility distinction is crucial. Whilst it is true that in some instances the owner of the establishment *is* a repeat victim (such as shopkeepers and shop theft or licensees being assaulted), the notion of a risky facility opens our minds to the idea that the premises *hosts* the crime (thus the risk), as opposed to *experiencing* it. Whilst this may seem like semantics, it means that the terminology is more in keeping with the key focus of environmental criminology and its tenets: that the *places* of crime are important and that these places (through their design, the routine practices and activities associated with them, the way they are managed or even the people who use them) can *create the opportunities* that allow crime to flourish. For this reason, I argue that not only should the term risky facilities (as opposed to repeat premises victimisation) be firmly embedded in the discourses of environmental criminology, but also that research needs to continue testing the concept across a range of facilities and geographies and needs to be applied to explaining its occurrence.

In terms of the extant knowledge about risky facilities, this term has only been used since 2003 (Clarke & Eck, 2003), and did not really appear in the academic literature until 2007 (Eck et al., 2007) so there are limited studies specifically referring to ‘risky facilities’. Instead one has to inspect the research involving the previously discussed notions of ‘crime attractors’, ‘crime generators’ and repeat ‘victimisation’, as well as unpicking studies involving particular types of premises. A number of such studies that show evidence of uneven distribution of offending within a set of facilities are summarised in the table in Appendix 2. Also worthy of note is the fact that this type of concentration has been found even when only those facilities experiencing at least one crime are analysed (Wellsmith et al., 2007), within a sample chosen as possessing characteristics likely to represent problematic premises (Hope, 1986) and when crime rates (as opposed to incidence) were analysed (Sidebottom & Bowers, 2010).

From considering this research, it is clear that a wide range of establishments appear to include risky facilities. However, very few of these studies were carried out with the specific purpose of identifying such concentrations. In most cases, this pattern is a descriptive aside to some other purpose. This has changed more recently, particularly through the work of Ashby & Bowers (2015), Bowers (2014), Madensen (2007), Madensen & Eck (2008), To this end, my study will consider all classes of facilities for which there are sufficient data to test for risky premises, as well as investigating whether such premises are risky for different crime types and at different times.

Linking the concept of risky facilities back to theory, there are a number of explanations as to why this phenomenon may be apparent in recorded crime data, which are summarised by Eck et al. (2007) as: random variation; reporting processes; targets; offenders; and place management. Wellsmith et al. (2007) add the external environment and proximity to other premises. Random variation suggests that the distribution has occurred by chance. In other words, there is nothing about the premises or the environment that makes one place riskier than another and if the analysis was repeated at some other time there would be no consistency in ranking. This explanation highlights the importance of using sufficient and appropriate data and, where possible, carrying out time series analysis (Eck et al., 2007).

Another explanation that may result in such a distribution is reporting differences. It may be that those premises showing the highest amounts of crime actually report a much greater proportion of incidents. If this were the case, the distribution would not necessarily be reflective of problematic premises, but in fact could be the opposite: those premises that did not report incidents might be the ones with the bad reputations and the worst relationships with the police (Eck et al., 2007; Wellsmith et al., 2007).

Linking back to opportunity theories, these concentrations may be explained by the distribution of targets. It may be that risky facilities are the ones with the most (perhaps the largest premises), or best, suitable targets (Eck et al., 2007). For this reason, it is important not only to consider the volume of crime at each location, but also the rate of crime (e.g. Sidebottom & Bowers, 2010) using an appropriate denominator.

In terms of offenders, it may be possible that risky premises are more attractive to potential offenders than others in their class. There are a number of reasons why this might be the case. The premises may be situated close to a concentration of offender residences. Alternatively, offenders may be attracted to certain establishments because of poor place management or the availability and suitability of targets.

Place management is theoretically extremely important in risky facility formation. As discussed in chapter 2, the layout, practices, management and many other place characteristics may all determine the level of crime experienced within a premises (Eck et al., 2007) as is now well established in relation to on-licensed premises (e.g. Graham, Bernards, Osgood, Homel & Purcell, 2005; Homel, Carvolth, Hauritz, McIlwain & Teague, 2004; Macintyre & Homel, 1997; Madensen, 2007; Madensen & Eck, 2008). The concept of risky facilities also proves useful in helping determine positive and negative design and place management, by allowing comparison of the features of risky and non-risky premises.

In addition to those environmental features that place managers may control, the external environment may also impact upon risky facilities. This may be related back to the idea of area effects discussed in relation to crime attractors and generators, as well as to the proximity of offender residences or 'bad' neighbours. Additionally, the recent research on social disorganisation and collective efficacy at the micro-level, suggests that these 'community' explanations may also have a role to play in the emergence of risky facilities. If there is social cohesion and a willingness to intervene in response to unacceptable behaviour, then this is just as likely to influence the extent of crime at facilities situated within those 'communities'. However, this would require the facility to be seen as part of that society, with shared norms and values, and for people to feel the same desire to intervene and protect when a business is being 'victimised' as they do when it is a neighbour. Although this is possible, it is perhaps less likely, and has not yet been explicitly tested. It also seems less likely for large facilities (such as big box stores and hospitals). Further, environmental and community explanations would seem to suggest that facilities in close proximity (i.e. within the same neighbourhood) would be either all risky or all non-risky. To date, research on the clustering of risky facilities, whilst limited, suggests that this is not the case (Block & Block, 1995; Madensen & Eck, 2008). Finally, in relation to area and environmental features, it must be recognised that risky facilities may be affected by the types and quantities of other premises situated near them (of the same or different classes). As with crime generators, other facilities may have a summative, multiplicative or neutralising effect, but this is as yet significantly under-researched (though there is some evidence that risky facilities may have a 'radiating' effect on area crime (Bowers, 2014)).

Madensen & Eck (2008) reduced the explanations for risky facilities to four possibilities. These are (1) risky facilities as a reflection of their situation in high crime ('bad') neighbourhoods; (2) place management explanations; (3) patrons, whereby risky facilities *attract* offenders because they are seen as good places to offend; and (4) behaviour settings, which relates to the norms or behaviour in a given environment.

These four hypotheses are tested for bars by Madensen & Eck (2008) using data from

Cincinnati, Ohio (and findings from previous studies). They found little evidence to support the first hypothesis that high crime bars are produced only in high crime neighbourhoods, however, as bar locations were also highly skewed (with 29% being situated in only two out of 53 neighbourhoods) the authors were not able to carry out statistical analysis to test this hypothesis, instead relying on mapping of high and low crime bars. However, as high and low crime bars were often in close proximity to one another, they concluded that any neighbourhood effects were unlikely to be significant predictors of risky bars, which was also consistent with the earlier work by Block and Block (1995). Of particular interest to my study, they state: "This finding encourages the notion that bars function as relatively autonomous microenvironments that are at least partially insulated from external neighborhood-level effects." (Madensen and Eck, 2008: 116-7). This suggests that social disorganisation, socio-economic status, racial heterogeneity, and arguably local informal control, are unlikely to be the 'cause' of some facility addresses (in this case bars) experiencing more crime than others. The authors go on to propose a model that is heavily dependent upon the decisions and actions of place managers in determining the degree of violent crime experienced. This model also suggests that the remaining two hypotheses (patron and behaviour setting) are in fact also functions of place management.

A survey study of frequent bar patrons by Quigley, Leonard and Collins (2003) also found that bar characteristics were the strongest predictors of violence. Having categorised establishments into violent and non-violent bars (of which the former were much greater in number, though this separation did not reference risky facilities), they also found, however, that individuals with certain personality characteristics were *attracted* to bar environments that promoted violence. Thus the existing research on 'risky' licensed premises favours place management explanations for the disproportionate crime contribution.

3.3.4 The universal law of crime concentration at place

Although not focused on facilities, a recent development within place-based criminology requires some consideration. Similar to Wilcox and Eck (2011) describing the seemingly ubiquitous concentration of crime in places as an 'iron law', Weisburd (2015) has recently suggested that there might be a universal law of crime concentration at place, whereby a (yet to be established) narrow bandwidth of places (the focus has been on street segments, but this could also include risky facilities) contributes a disproportionately high level of crime; which he has 'set' at 25% and 50%. He believes this pattern will hold regardless of the location and researchers have embarked on testing if this is the case, with generally confirmatory results (Gill, Wooditch & Weisburd, 2017; Haberman, Sorg & Ratcliffe, 2017; Levin, Rosenfeld & Deckard, 2017). This is discussed more in the next chapter, that looks at issues regarding identification of spatial crime concentrations and unequal distributions.

3.3.5 Journeys-to-crime

There is a large body of work on the journeys taken to commit crime, which establishes patterns such as the tendency for trips to be short, as introduced above, as well as critiquing them. The tendency of offenders to take certain types of journey may be an extremely important influence on where crime concentrates, at various levels of resolution. Thus the study of journey-to-crime is conceivably extremely important for considering, amongst other things, the emergence and location of risky facilities, yet to my knowledge, there is no published literature exploring this. The current research aims (as discussed in chapter 5) to begin to rectify this. Therefore it is necessary to rehearse here some of the major findings and critiques.

As already noted, the abiding lesson from journey-to-crime research is that crime trips tend to be short (see, amongst others, Chainey & Ratcliffe, 2005; Wiles & Costello, 2000), but this finding alone highlights and raises a number of issues and critiques. Firstly, it focuses on the distance of crime trips, as does the vast majority of the literature (Townsley & Sidebottom, 2010). Yet as Rengert (2004) has pointed out, a crime journey is made up of three elements: the origin, direction and distance. Direction is a neglected feature in the journey-to-crime research (Rengert, 2004), even though this could tell us a lot about the impact of the environmental backcloth and opportunity clustering (see, for example, Van Daele & Vander Beken, 2011).

The origin is usually assumed to be the offender's home address (and for analysis purposes tends to be the police listed residence for that offender at the time of committing the offence or arrest). However, there are any number of other locations that could be the 'origin', including a family or friend's home, a place of work or leisure, school, and so on (Wiles & Costello, 2000). In fact, the origin could arguably be anything that Brantingham and Brantingham (1981) have referred to as a node, depending on when we perceive the crime 'trip' to begin. Though beyond the scope of this study, there clearly needs to be more research carried out to consider "…an offender's *entire* node network to understand journey-to-crime patterns" (Townsley & Sidebottom, 2010: 901, orig. emphasis).

A further element that has been identified (but rarely studied) is the trip taken after the offence has been committed, that is to say the journey *after* crime (see Lu, 2003). Albeit an extremely interesting (yet difficult) concept to consider, this has not been incorporated into the current study.

Finally, in relation to the features of the journey-to-crime, the destination has not yet been discussed. I would also add this to Rengert's (2004) list of elements, as without a destination there is no journey (and no crime). This has likely been excluded from consideration because it is 'a given', but it seems only to be used as a place to draw the arrowhead of a journey, rather than an element in its own right. There is also arguably a great deal of research into this; it is the focus of the majority of most modern studies concerning crime and place. However, studying the locations of crime separately, rather than studying them as part of the journey-to-crime, one ignores the relationship between all these nodes and the crimes committed 'en route' between them (Mago, Frank, Reid & Dabbaghian, 2014). In particular, one also fails to consider how the distance and direction of journeys are affected by the relative positions of origins and destinations, as well as how likely distance and direction to be travelled from the origin might actually *determine* destination (thus crime location and concentration).

Returning to the most researched issue of distance, it has been established that mean distances tend to fall within a range of approximately 500m to 6km (Vandeviver, 2013), depending upon the study, methods used and, crucially, crime type (White, 1932; see also Bichler, Christie-Merral & Sechrest; Capone & Nichols, 1976; Rhodes & Conly, 1981; and Wiles & Costello, 2000, amongst others). It is also proposed that the frequency of offences decreases with distance away from the origin (invariably the offender's recorded residence), and that this distribution can be described by a distance-decay function (Capone & Nichols, 1975; Rengert et al., 1999; Rhodes & Conly, 1981; Rossmo, 2000). This is consistent with Zipf's (1949) least effort principle, and in keeping with opportunity theories (that it is rational to offend in areas that are known, close-by and take little effort to get to). However, these seemingly well-established findings have been challenged, due to a number of potential methodological flaws.

Firstly, as has been pointed out, when longer (particularly very long) journeys do exist, these tend to be excluded from the analysis (Vandeviver, 2013). This may be deliberate (for ease of analysis and coding), because such trips are less likely to appear in local crime records. This is possible if 'external' offenders are less likely to be detected, or if a local 'origin' has been recorded instead of the offender's residence (which may be more accurate, but introduces an inconsistency into the comparison), It might also be because long journeys do not involve offending at typical city locations (Van Daele & Vander Beken, 2011). Additionally, because the focus of journey-to-crime research tends to be on aggregation and means, the trips at the extremes, particularly those that are less frequent (which research suggests are those which are longer), and those of offenders that fall outside the usual patterns, tend to be 'lost' as noise or outliers, rather than being the focus of further investigation (though see Van Daele & Vander Beken, 2011 and Vandeviver, Van Daele & Vander Beken, 2015).

Secondly, it has been argued that the patterns found by much research may be affected by aggregation (van Koppen & De Keijser, 1997) and nesting (Townsley & Sidebottom, 2010). Aggregation effects relate to the biasing of aggregated outcomes (notably distance decay) by the more frequent patterns of individuals (such that the overall pattern does not represent the distribution of all the individuals that make it up, a type of ecological fallacy). This is problematic if aggregated data are used to make inferences about individual journeys. Nesting effects are the impacts of repeatedly sampling from the same group, such that a relatively small proportion of prolific offenders (or other features, such as age or experience) might skew the overall patterns obtained. These problems have been proposed and tested and the results raise concerns that need to be recognised (Andresen, Frank & Felson, 2014; Smith et al., 2009; Townsley & Sidebottom, 2010; van Koppen & De Keijser, 1997). However, there is also subsequent literature responding to these critiques that continues to lend support, in particular, to the ubiquity of the distance-decay pattern at aggregated *and* individual level (for example Bichler, Christie-Merrall & Sechrest, 2011; Levine and Associates, 2003; Rengert et al., 1999).

This necessarily brief overview of journey-to-crime patterns and research concerns, highlights, as an aside, the importance of critically considering methodologies and the validity of the conclusion drawn from any analysis. More specifically, it sets out my arguments for considering the relationship of key spatial phenomena (such as journeys-to-crime and risky facilities) and their influence on one another.

3.4 Conclusion

Crime concentrations are now well established phenomena. Such distributions are apparent across a number of dimensions including space. In recent years, there has been a general move towards analysis at the micro-level, which is a natural progression for environmental criminology, helping to develop and support more rigorous theories of crime and place (Taylor, 1998). At the facility level, this has led to the development of concepts such as crime attractors, crime generators and risky facilities. There is a significant amount of literature establishing that high crime areas or hot spots often include (or are perhaps caused by) high-crime facilities. Brantingham and Brantingham (1995) theorised that certain types of premises could act as crime attractors or generators as a result of the exploitation of many good opportunities by many potential offenders in these overlapping activity spaces. Such entities are now well established, as a result of a reasonably large body of research evidence. This identifies (and to some extent explains) why shopping malls and bars have higher incidences of crime than libraries or churches, but it does not (explicitly) consider the variation of crime within any one type of facility. The concept of risky facilities has filled that gap, and the research shows that crime is not evenly distributed across addresses within the same type of facility, with a small proportion of premises accounting for a large proportion of crime. The pattern has been observed in multiple jurisdictions, and for various different types of facility. However, there has been little systematic research carried out looking at risky facilities, and the methods used to identify and compare across studies (thus across facility types) have not been critically considered. Indeed, it could be argued that risky facilities were conceptualised and then almost immediately research endeavours turned to *explaining* the concept, rather than first thoroughly *exploring it*.

Key issues relating to identifying and measuring spatial crime concentrations and related concepts are discussed in the following chapter. However, it is important to note now, why this under-exploration and vagueness is a problem. As it stands, risky facilities are deemed those premises that contribute a disproportionate amount of crime (for facilities of that type, in a given area). However, because there is no consensus over when this pattern exists or how to identify it, it is not possible to compare across research studies (for example to determine if one area or time has greater concentration or inequality than another), to compare across facilities (to see if one type of premises experiences more disproportionality than another), or to compare across other features of the offence or environment to identify patterns (for example, considering concentration within premises by crime type). Further, because there is no agreed cut-off point for when something is a risky facility (as opposed to a non-risky one), it is not possible to know if the patterns found in one study are supported, or challenged, by those in another. Applying different empirical definitions might have resulted in different findings. This means that the body of evidence that is being built up regarding this important form of micro-spatial concentration is less rigorous and persuasive then it could be, and leaves it open to greater critique. Weisburd's (2015) law of crime concentration could be applied to risky facilities, allowing for some comparison, but the percentage crime contributions selected remain (as will be discussed more in the next chapter) problematically arbitrary. This research seeks, in part, to address these issues, as will be further explained in chapter five.

A further observation to be made regarding the existing research, is that some of the most commonly explored concepts within the criminology of place have not been considered with respect to risky facility patterns. For example, journey-to-crime and risky facilities. As has been shown above, there is some research considering the spatial patterning of risky facilities, but this is relatively limited and there is only one study (Bowers, 2014) that considers the possible impact of risky facilities (as opposed to crime generators/attractors) on the wider area in which they are situated, which may include other facilities (though this itself has not yet been explored). In addition, despite the purported desire to test theory and propose interventions by exploring risky facilities, there still remains very little research that even compares risky with non-risky premises. To consider all of these issues is beyond the scope of this research, therefore instead, I make a call for such work to be carried out in the future, and commence this endeavour by considering the distribution of crimes amongst different facility categories and the premises within them, and comparing risky and non-risky facilities in relation to some of the spatial features discussed in this chapter, most notably co-location (with one another, and with offender residences) and journeys-to-crime. It is hoped that this will incentivise others to adopt similar approaches and to carry out further related research.

A more detailed consideration of the research questions is set out in chapter five. Firstly, however specific issues regarding methods and approaches to studying spatial crime concentrations are explored in the following chapter (four).

Chapter 4: Researching spatial crime concentrations

There are a number of issues that are particularly relevant when researching and analysing spatial crime concentrations. Those that are particularly pertinent to this research study - units of analysis, attribution of crime to particular facilities, identification of spatial concentrations, and cut-off points, cateorisation and labelling - are now considered in more detail.

First, I note that the majority of the research presented so far relies on quantitative research methods. My critiques of this are discussed further in chapter 6, as I justify adopting a mixed methods approach for this study.

4.1 Units of analysis and micro-geography

It is my contention that as crime concentrations are in fact the aggregation of individual routine activities and offending behaviours, much more attention should be paid to the analysis of micro level patterns and the individual routines that produce them. This is echoed by Steenbeek and Weisburd (2016) and Groff, Weisburd and Morris (2009) who, although referring mainly to street blocks, identify a number of benefits to micro analysis. Firstly, it can be conceived as "constituting a single behavior setting bounded by time and space" (p64). Secondly, as there is less aggregation, the risk of ecological fallacy is reduced. Thirdly, as any variation identified results from a single premises, it is more likely to be resolvable by the police or other intervention, as well as being more manageable in size. Fourthly the problems of spatial heterogeneity are reduced. In considering trajectories of juvenile delinquency at the street block level, Groff et al. (2009) conclude that there was significant variation at the micro level that would not have been identified had a higher level of aggregation been employed. Micro-level analysis is becoming more common in research on the universal law of crime concentration, on work relating to crime trajectories, and is also apparent within the literature on attractors, generators and (in some cases) hotspots (Bernasco & Block, 2010; Braga, Hureau, Papachristos, 2010; Brantingham, Brantingham, Vajihollahi & Wuschke, 2009; Groff, Weisburd & Yang, 2010; McCord & Ratcliffe, 2007; Nelson, Bromley & Thomas, 2001; Oberwittler & Wikström, 2009; Weisburd et al., 2006). However, much of this research, still focuses on units of analysis greater than individual addresses, usually blocks, block faces and street segments. Address-point analysis is more common when considering residential burglary, but in the context of the wider body of research on crime and place, study of nonresidential address point data remains relatively rare (with the notable exceptions considered in the current review of literature). The study of risky facilities, however, requires us to consider individual premises, and this may raise interesting questions regarding the relationship between crime patterns at premises, block, neighbourhood and city-level and the methods used to investigate them (Bernasco, 2010; Brantingham et al., 2009; Rengert & Lockwood, 2009; Tita & Greenbaum, 2009; Weisburd, Bruinsma & Bernasco, 2009). Future research will also need to consider the appropriateness of macro, meso and micro level explanations, their applicability at different levels of analysis and the inter-play between them.

4.2 Attributing crime to facilities and the argument for studies using 'within facility' crime data

As is apparent from a consideration of the preceding literature, there are particular approaches that tend to be employed when considering spatial concentrations of crime, particularly at the micro-level. The vast majority of studies use one of two main methodologies to attribute crimes to particular places of interest: street segments (with or without buffers) or address points with buffers. There are a number of arguments in favour of this approach, which will be considered below. However, there are also significant flaws, which seem to have been accepted on the grounds that there appear to be few alternatives. However, I intend to show that, though also flawed, there are benefits to adopting an alternative approach as a way of adding a further dimension to our understanding of crime concentrations within facilities. The original risky facilities article (Eck, et al. 2007) discusses the concept in relation to crime committed *at* premises, and this is the approach taken by Madensen (2007) and the earlier research by Block & Block (1995). However, generally study of crime concentrations involves identifying hotspot areas and looking at what facilities are within them (as already critiqued), or identifying types of facility and drawing buffers around them to capture crime.

There are three key problems with using buffers to capture attributable crime. Firstly, determining the size of the buffer. Secondly, how to deal with premises of interest that are close by and the (potentially regular) occurrence of overlapping buffers. Thirdly, the flawed assumption of attribution or causality (which can also be applied to unbuffered street segments when used as the unit of analysis for studying facility crime). I will now consider each of these in turn.

Determining the size of the buffer to place on either a street segment or an address point could employ an arbitrary approach, a theoretically driven determination and/or the use of quite sophisticated statistical analyses to determine the extent of effect, such as changepoint regression (Ratcliffe, 2012). Alternatively, it is possible to weight crimes by the distance they are from the ‘source’. Though a combination of theoretical and statistical methods is likely the most defensible, this approach is still based on problematic assumptions relating to causality and, possibly, to still arbitrary decisions relating to appropriate cut-off points or degree of weighting. When using buffers, there is also a decision to be made regarding the shape of these. Often they will reflect the shape of the unit of analysis: a circle for an address point or a ‘sausage’ around a street segment; however, some studies recognise the problem of applying this type of buffer to a non-uniform environmental backcloth and street networks. In these cases, researchers have instead produced buffers that are more appropriate to capture the real-world movements of people and instances of crime.

Therefore, it can be seen that if buffers are to be used, their determination needs to be thoughtful and justified, and more advanced techniques must be employed. Although it may be possible, given sufficient research (which does not yet exist), to establish universally applicable (by facility type) buffer limits, it is more likely that such limits will vary by study area, at least when these comprise different types of location. For example, suburban buffers might be larger than CBD buffers, those in cosmopolitan cities might be different to more rural towns. Certainly, there is likely to be variation from country to country, particularly if there are different routines, travel patterns, zoning and planning conventions, and street networks/layouts. Indeed, one of the benefits of the work presented by Ratcliffe (2012) is that it was able to distinguish between appropriate buffer sizes for individual premises. In principle this makes it a viable approach for identifying risky facilities (other issues with buffers as discussed below, aside), but there remain disadvantages. Firstly, it is affected by the accuracy of the geocoding employed in the area under study, and this is recognised as being something of a weakness, particularly for US police data (Mazeika & Summerton, 2017; Ratcliffe, 2012). Secondly, although academic researchers may be in a position to apply this technique, for many practitioners this may be a difficult, and relatively time consuming, task therefore it is not surprising that often a more simple, but less defensible approach is employed. Further research looking at the appropriateness of different types and diameter/lengths of buffers, along the lines of that carried out by Groff (2011) and Ratcliffe (2012), is needed in order to determine how appropriate such methods are. In particular, this needs to be expanded beyond the typical suspects of licensed premises, to consider a wider variety of facilities, that may less obviously ‘spill’ crime into the surrounding streets.

However, sophisticated the approach is, the use of buffers raises further practical and conceptual issues. In practical terms, it is highly likely that the more simplistic buffers drawn around units of analyses will overlap, as the premises or streets themselves will be in close proximity. This then raises the methodological issue of how to attribute and count crimes that fall within multiple buffers; the main options being to count them multiple times or to split/weight them across the buffers in which they fall. So, for example, under the first method a crime falling in two buffer zones is counted once for each buffer (thus its associated unit of analysis) and under the second method it would count for 0.5 crimes in each. Alternatives could be to discount such crimes (though this approach is not used as there is no theoretical justification for doing so) or possibly to weight or attribute according to their distance from the origin. This problem again highlights that there is no single, clear empirical approach taken and I contend this is because the use of buffers itself is fundamentally flawed.

The reason I make this claim moves us to the third critique: that of attribution or causality. In essence, the reason for using buffers or street segments is to recognise that crimes associated with particular premises do not just occur inside those premises, but also spill out into the spaces around them. Aggregated areas of influence (buffers, street segments or census blocks) also seek to account for the fact - usually the reason for the research being undertaken - that at least some of these crimes would not have occurred, or would not have occurred where they did, if it were not for the presence of the premises in question. For example, it is widely accepted that much violence and disorder occurs in the so-called night-time economy because of the presence of drinking establishments. The exact reason(s) for such crime can be debated, drawing on a very broad range of explanations including psychopharmacological effects, peer-group and social norms, (sub)cultural and masculinity theories, permissive environments and the opportunity and design theories set out previously (for summaries of these approaches to alcohol and night-time economy related violence see, in particular, Graham and Homel, 2008). Whichever approach is taken, however, there is a clearly established connection between drinking and/or partaking in the night-time economy and engaging in, or being the victim of, violent crime. Thus if the individuals involved had not visited the particular licensed premises they did, this crime would either not have occurred at all, or would have occurred at some other time and space in proximity to the premises they *did* visit.

It is laudable, therefore, to seek to capture these crimes and associate them with premises that resulted in their commission and placement. Indeed, this could be seen as an attempt to identify and measure the effects of crime generators and attractors (though whether they can be isolated from one another when in close proximity is debatable) and is a more appropriate way to do this than to identify hotspots and then look at what premises are in them (and assume this is the reason why they have formed). However, the simple existence of a crime in the vicinity of any particular premises clearly does not mean that the crime occurred *because* of those premises. So to attribute that crime to those premises purely based on proximity seems inappropriate. The flaws of this approach are yet more apparent in areas with numerous premises in close proximity. If considering land use more generally (e.g. does crime occur more near commercial areas) this may be reasonable, however, if seeking to consider patterns by facility type or even address (as the risky facility approach seeks to do) then how are their effects separated? To what extent does their co-siting increase or even decrease crime occurrence (Bowers, 2014)? Why is a crime attributed to one premises and not another, or is counted twice and attributed to both, or even more theoretically bizarrely, why is half a crime attributed to each?

To summarise, I contend that the use of street segments and/or buffers (or hotspot mapping, though that is not being discussed specifically here) to identify crimes that can be attributed to specific premises is fundamentally flawed, yet it is almost exclusively the approach that is taken in this type of micro-spatial analysis focusing on crime attractors, generators and risky facilities. It is flawed as identifying suitably sized and shaped locations or buffers is complex and even the most sophisticated approaches fail to justify why it is appropriate to attribute crimes to nearby premises in this way, as well as struggling to deal with how to count crimes in the vicinity of multiple premises.

I believe the reason this approach still dominates is because there is little alternative. It is a methodology that seems to take a ‘least bad’ philosophy. I propose, however, that if this is the case, there is another ‘least bad’ method to consider, one that I believe is actually conceptually more justifiable and is also empirically more defensible as it is more likely to under-count (all) premises’ contributions. This echoes the approach used by Madensen (2007) but applies it to all crime across a city.

To this end, the current research is based upon what I refer to as ‘facility crime’ (because my focus is on facilities specifically, but the same method could be used for any premises). It is the same approach that is likely to be employed when analysing (residential) burglary - an offence that can only occur at (not in the vicinity of) a specific address. That is, crime attributed to the facility is only that which is recorded as occurring at the facility’s specific address. No buffers are used and, in theory, a crime can only occur at one particular premises, so there are no concerns about double-counting. How I have approached this in practice is dealt with in the methodology chapter (6), where I also note specific errors and limitations in relation to coding and the dataset used. In general, however, there are two particular weaknesses associated with this. Firstly, it is extremely dependent upon the quality of the data available. This is notably in relation to the accuracy of address-recording but also relates to identifying the type of facility in question. Thus it is affected by the accuracy and completeness with which the original dataset (usually police recorded crime data) identifies addresses as facilities and (if required) the type of facility this is. Other data sources may wish to be additionally used to help identify premises, but this is in turn affected by the completeness of such datasets and the ability to match them on address or (possibly inaccurate) geocodes.

Secondly, when research is not limited (as mine is not) to offences that only occur at specific addresses (such as burglary), this approach will under-count both crimes that could be considered directly associated with the address in question (for example a fight that has spilled out of a bar, but is recorded as occurring in the street) and that which may have resulted because of the existence of that facility in that place (which is of course what the dominant approach seeks to do). However, as quantitative research seems arguably incapable of identifying the latter, anyway, I contend that under-counting is more defensible. At the very least, the findings from such an approach need to be considered, to determine if the resulting patterns of offending and support for opportunity theories and spatial phenomena such as crime generators and risky facilities are consistent with the extant research.

In relation to facility-focused research, the work reported in this thesis is, to my knowledge, the only city-wide, multi-facility type research to adopt this approach. This is certainly the case when recorded crime data are taken as the starting point (as opposed to identifying a facility type, such as a school or licensed premises, and then obtaining all crime recorded as associated with those premises (e.g. Madensen, 2007). This original contribution, therefore, provides an alternative viewpoint on crime concentrations in facilities, thus contributing to current knowledge on risky facilities, as well as to the growing body of work testing Weisburd’s (2015) proposed law of crime concentration at places.

4.3 Identifying crime concentrations at places

We know that crime concentrates in any dimension, including space, so that distribution is not even. We suspect that this is always the case regardless of the unit of analysis/cone of resolution, and so forth. Weisburd (2015) proposes there could be ‘universal rules’ of crime and place as introduced in chapter 3, and uses 25% and 50% crime contribution as a way of considering and comparing the (anticipatedly small) proportion of the location (in this case, risky facilities, but usually street segments) that accounts for this amount of the crime. Clarke & Eck (2007), on the other hand refer to the risky facility concentration with reference to the 80-20 rule, though recognise that these figures are only approximate. How then, do we actually determine if the distribution of crime across facilities is sufficiently unequal to warrant further attention? How do we know if the risky facility pattern exists across studies? Do we look for an 80-20 pattern?

Do we consider the proportion of premises contributing 25% (or should it be 50%)? Do we look for a J-curve (and how do we know when we have one)? In the risky facility literature there is also some mention that this distribution might be seen as a form of 'power-law' (Eck, et al. 2007).

4.3.1 The power-law distribution

A power-law is one of a number of heavy-tailed distributions (meaning that the 'tail', which can be right, left or both, is heavier than an exponential distribution). Simply, a power-law is a relationship between two variables, where the frequency of an event varies as a power of some attribute (here crime count) and can be summarised by the general function

$$f(x) = ax^k$$

with the complementary cumulative distribution function displaying as a straight line on a log-log plot of occurrences (crime count) versus number of events (premises) with that count or more (Clauset, Shalizi & Newman, 2009; Mayhew, 2015 Newman, 2017). In honour of one of the first people to study this distribution phenomenon (in perhaps its most famous application: income inequality), this form is often referred to as a Pareto distribution, and the 'law' is sometimes known as Zipf's law (Pareto, 1896 cited by Newman, 2017; Newman, 2017; Zipf, 1949)

As Gillespie (2015) and Clauset, et al. (2009) (among many others) recognise, there are a large (and seemingly growing) number of power-law distributions reported across many fields of study. Indeed, Gillespie refers to this as an "apparent ubiquity" (20015: 2) but points to work by Stumpf and Porter (2012) that questions whether many of these are really power-law distributions, or just something that might look like them. The 80-20 rule stems from this type of (Pareto) distribution (Gillespie, 2015) and as the J-curve can be considered a 'heavy-tailed distribution' (Gillespie, 2015) it is reasonable to seek to fit this to a power-law distribution. As most phenomena do not fit to a power-law distribution for all values of x (Clauset et al, 2009), it is usually fitted to the 'tail'. As these are the values of most interest in terms of risky facilities, this seems appropriate. However, it is noted that there are other distributions that might (better) fit, so it is worth testing these claims. There is little point referring to the risky facility or J-curve pattern as a power-law distribution (even if casually), if this is not, in fact, the case. Indeed, future researchers may wish to test the appropriateness of fitting this to a power-law, or seek to use it as a way of testing whether their data conform to 'the' risky facility distribution. Therefore, it is important to test whether this is in fact appropriate, and an undertaking worth pursuing.

4.3.2 The Gini coefficient

The Gini coefficient (Gini, 1912; 1921; see also Dalton, 1920 and Alker & Russett, 1964), and the resulting index, are used heavily in economics but have also been employed in many other fields (Deltas, 2003; Xu, 2004). The measure quantifies the dispersion of a distribution, such as income or wealth, across a (sub)-population. The Gini can vary between 1 (absolute inequality) and 0 (absolute equality). In addition, because it is a relative measure that controls for the total underlying prevalence of a characteristic, when calculated for several populations (e.g. countries) the coefficients produced can be used to create an index, which thus compares inequality across these units of analysis (Alker & Russett, 1964; Xu, 2004).

The Gini coefficient has been used in criminology, most notably within criminal careers research to consider and compare the distribution of offending in cohort studies (e.g. Fox & Tracy, 1988; Piquero, 2000; Piquero & Buka, 2002). It has also seen some limited consideration in relation to the distribution of offending across places, such as the locations of metal thefts from railways (Ashby, Bowers, Borrion & Fujiyama, 2014) and of crimes in Swedish (Stockholm) and British (Merseyside) schools (Lindström, 1997). In essence, this latter use of the Gini coefficient summarises the extent of concentration of problems into particular places.

Fox and Tracy (1988), the first to use this measure (which they coined α) in relation to delinquency, and Piquero (2000) state that the ability to compare across groups and times is the real value of the Gini coefficient. Thus this approach allows one to conclude which of a number of populations exhibits the most unequal distribution (assuming these are selected appropriately) (Deltas, 2003; Fox & Tracy, 1988; Piquero, 2000; Xu, 2004). In the current study, this may be considered as the greatest degree of concentration within addresses. A further favourable asset of the Gini index approach, compared to the bandwidth approach proposed by Weisburd (2015), or the more prosaic '80-20' rule, for example, is that it does not rely on (usually arbitrary) cut points (e.g. 25%) (Fox & Tracy, 1988; Piquero, 2000). It should be noted, however, that the Gini coefficient does not indicate the nature of the distribution (i.e. the type of distribution curve; in this case it has been posited risky facilities follow a J-curve or power-law distribution – calculating G will not identify if this is the case). Nor does it demonstrate where the inequality (concentration) lies (Alker & Russett, 1964; Fox & Tracy, 1988).

4.4 Cut-off points and labelling

It is important, particularly in practice, to appropriately use terminology, particularly if this could be viewed negatively. We must be justified, for example, saying that a named facility is 'risky', with the connotations this brings and the attention the premises - and likely the place manager - will receive. It would be hugely problematic to use definitions that ignore the absolute extent of crime at a premises, and call it risky because it is the topped ranked premises, even though the actual amount of crime it experiences is low, particularly with regards to the rest of the study area. This is also a reason why the testing of the ubiquity of the risky facility concept is important, for if not all types of facility experience this kind of disproportionate concentration of offending in a few addresses, then there is much less justification for ordering premises and then targeting those that are most highly ranked. This also raises the issue of choosing appropriate methods of categorisation. If we consider all facilities (or even all nonresidental land use) together, then the premises identified as problematic will be different to if we consider separate categories (types) of premises. Further, the decision of how fine this categorisation should be, is also likely to impact on our findings. In addition, there is a balance to be achieved between reducing within-group variation and increasing across-group variation, whilst maintaining suitable sample sizes. Ultimately, it is proposed this should be driven by both theoretical (categorisation by the business taking place at each type of premises and the routine activities associated with this) and methodological considerations.

Whether there is a large difference in the outputs of analysing categorised (or not) data will also be affected by the method of identifying (labelling) something as risky. If, say, a 'top ten' method was to be used, then there can only ever be ten facilities identified per selected category. Therefore, if all facility crime is considered together, no matter how much crime there is, or how many problematic premises there might be, the number deemed to be risky will be ten. If facility types are instead split into ten categories, with the top ten of each identified as risky, the same area would have 100 risky facilities. Even when using a formalised cut off point, then, the labelling of premises as risky - and possibly requiring intervention - will be affected by these types of decision.

The issue of categorisation (and accuracy and completeness of data) notwithstanding, it is important that the application of crime concentration labels (such as risky facilities, hotspots, repeat victims, recidivists, and so forth) is done in a systematic and consistent way. With regards to their academic application, definitions of recidivism, repeat victimisation and hotspot are generally well-accepted and reliably used. This is particularly well exemplified by the work of Ratcliffe & McCullagh (1999; and associated 2001). In their study of hotspots they use a two stage method of identification, drawing on LISA statistics, notably G_i^*. In very basic terms, this approach identifies areas that have significantly higher levels of crime than in the surrounding study area (based on the mean and variance) and is based on the work of Getis and Ord (1992;

see also the rewritten version of Ord & Getis, 1995). A significance level is selected and those that reach this level and have higher crime are then deemed hotspots. This means it is "…both significant and robust in a statistical sense…" (Ratcliffe & McCullogh, 2001: 332). Although they recognise that decisions still have to be made regarding a number of parameters to apply (such as the search radii and cell size), the approach used to identify hotspots in this study is an example of efforts to use transparent, replicable and justifiable methods to determine the focus of policy/intervention, as well as to apply a label, in this case 'hotspot'.

Unfortunately, methods such as this, particularly outside of research on hotspots and mapping, are rare and even within hotspot research, as already noted, different methods continue to be employed, and there are variations in the statistical techniques and parameters used (Chainey, Reid & Stuart, 2002; Getis & Ord, 1996; Hart & Zandbergen, 2014). In addition, even for these established concepts, it is not always clear what should be considered particularly *chronic* levels of concentration. It is questionable to what extent more sophisticated measures of identification are used in practice, and as terminology such as hotspots has come into more common usage, it is likely that these are referred to more casually, and when, perhaps, this is not warranted.

My main contention, however, is that identification and empirical definition *is* a significant problem for risky facilities, for both policy and within the academic study of crime and place. Across the research literature on risky facilities there is no agreed upon, widely used method of applying a 'cut off'. However, for much of this research, some distinction has to be made between what is risky, high crime or otherwise problematic and what is not, which means several different definitions are used. This is problematic because it affects the comparability of the findings from one study to another. This means that if the same patterns are not identified, it is not known if this is because there are actual differences or because risky facilities have been defined differently, with a different cut off. In the same way, if explanations are not supported we do not know if this is because the proposal was flawed, or what has been deemed risky (or not) is different. Additionally, therefore, I believe this means that the body of knowledge being built up is less convincing due to the lack of a systematically applied empirical determination of what ought to be categorised as a risky facility, and what ought not.

Even when a cut-off method does edge towards paradigmatic (such as the 80-20 rule or, more recently, Weisburd's universal law referring to a contribution of 25% and 50% of crime), it is not used by all researchers, nor can it often be regarded as anything other than arbitrary. Considering risky facilities specifically, there is certainly no evidence of consensus across the literature regarding a cut-off point. Some studies base their cut-off on the 80-20 rule, but then as acknowledged in the early literature, the 80-20 split is rarely an accurate representation of concentrations in real life, rather it is a conceptual nod to Pareto. Even if this were to be used, should the cut-off be the top 20% of premises (when ordered so that those contributing the most crime are listed first), regardless of the amount of crime they cumulatively contribute, or those premises that cumulatively contribute 80% of crime (regardless of what proportion of addresses this is)? Following this line of thinking, even if there was agreement on adopting one particular approach, the research to-date suggests the chosen cut-off would likely be arbitrary, with no reference to real-world empirical data on concentrations. Therefore, other than being a 'round number', why use '25% or 50%' as the defining crime proportion that will be contributed by a 'narrow bandwidth' of places. Why not 10%, 15%, 30%? Obviously any decision relating to this would include some degree of subjectivity, preference, or personal choice, at least in the first instance, but if this can be minimised and a method of empirical and/or theoretical justification presented to support this choice, then the selected definition will be more defensible.

It is acknowledged that much research does not use a cut-off but instead identifies the risky facility distribution and then carries out explanatory analysis on the whole dataset. However, this still does not then address the issue of which is a risky facility and which is not, that is required to truly test explanations. Otherwise, the research is looking at what is associated with higher incidences of crime, not what might be responsible for the heavy-tail of the distribution.

4.5 Conclusions and moving forward

In light of the critical consideration of the approaches to quantifying and measuring spatial crime concentrations, set out above, it is proposed that there are gaps in the research regarding the use of micro-units of analysis other than street segments, such that this study will focus on crime only occurring *at* facilities, not in the regions around them. I also contend that the concept of 'risky facilities' remains under-explored and I propose that the following questions need to be considered:

1) Is there an unequal distribution of crime? This should be identified using the Gini coefficient to allow for comparison across studies and units of analysis (facilities).
2) What is the nature of this distribution? This should be considered by visual inspection (looking for 'J-curves'), goodness of fit tests (for example, to the power-law) and may also incorporate the bandwidth approach, though further work may need to be done to establish a convention for this.
3) Which facilities are 'risky'? This requires the development and application of a method

for empirically determining what facilities within a given class (and accounting for the underlying crime incidence) should be defined as risky.

Therefore, a large part of this study is devoted to considering all three of these elements, by determining the nature of the distribution of crime across a number of different facility types and by proposing (and then applying) an empirical definition of risky facilities.

The next chapter explains in greater detail the overarching aims of this research project, and all of the research questions that have been operationalised from the issues presented through the review of the literature.

Chapter 5: Introducing the research aims and questions

Throughout the literature, there is evidence of an interest in land use as a determining factor in crime placement. There is also a move, particularly more recently, towards smaller units of analysis, including street segments and individual addresses. Nonresidential places have garnered attention, particularly with respect to the work on crime attractors and generators, and for certain types of crime, such as (alcohol-related) violence. However, there is a great deal of attention paid to residential crime, perhaps because it can be clearly defined as having occurred at a specific address. My interest, and where I believe there is still scope for significantly more research, is on nonresidential, public (or quasi-public) places. This study, focuses on *nonresidential* facility crime, hereinafter facilities and facility crime. The working definition of facility used is set out in chapter 6.

Given the extant literature on spatial crime concentrations presented thus far, it is my contention that not only is more research required at the facility level, but that we need to further consider how we attribute crime to facilities in order to test these patterns, that further consideration is needed not just of why these risky facility concentrations exist, but of what this pattern actually is, as well as exploring how the risky facility pattern should be defined and identified. I also propose that to further understand the emergence of risky facilities we need to study their relationship (in comparison to non-risky facilities) with the environment in which they are situated, including other facilities, the layout of the city, and the supply of offenders. The following set of aims and related research questions seek to address these issues, and hope to provide the impetus for further research, seeking to test and expand our existing knowledge of the risky facility concept and its place within the broader study of crime and place.

5.1 Aims and research questions

Considering the specifically relevant existing literature, the gaps in knowledge identified and my critique of the approaches taken to the study of crime concentrations in places, the current research has the following aims:

Aim 1: to consider whether crime concentrates in facilities, and how this concentration manifests.
Aim 2: to consider how risky facilities can be identified.
Aim 3: to explore key locational features associated with risky facilities.

Aim 4: to critically discuss the concept of facility concentrations and consider the appropriateness of definitions and terminology within this field.
Aim 5: to add to, and where necessary challenge, existing knowledge and its application in the field of spatial crime concentrations, particularly with respect to crime in facilities.
Aim 6: to make recommendations for policy, practice and further academic research.

Aims 4, 5 and 6 are addressed through synthesis of the research findings and a critical discussion of the extant literature. Aims 1 to 3 are directly addressed through empirical analysis and have been operationalised into the following set of research questions and sub-questions.

Aim 1: to consider whether crime concentrates in facilities, and how this concentration manifests.

RQ1: How 'ubiquitous' is the concept of risky facilities?
RQ1.1: Is crime in facilities concentrated in some premises more than others? More specifically, do different classes of facilities in a given location show evidence of unequal distribution of crime, which may be considered consistent with the concept of 'risky facilities'?
RQ1.2: Is there variability in the concentration of crimes within facilities by: (a) facility type; (b) crime type; (c) time of offence?

Research question one seeks to further expand the risky facility literature by not only considering whether the pattern exists in an unresearched location, but most importantly to consider how ubiquitous this pattern is, across different types of facility within a city, as well as for different crimes and different times of the day. RQ1.1 first establishes, using a number of methods, but most notably the Gini coefficient, whether the pattern of risky facilities is present in the study area for all facility crime. RQ1.2 then tests whether this pattern exists comparatively, and how crime concentration varies according to different features as set out. Thus, these questions seek to respond to my critique that *concept* of risky facilities has not yet been sufficiently researched and established, particularly for a range of different facility types (and through application of an under-used method of attribution: using only crime recorded as occurring at the premises in question).

RQ2: Are risky facilities stable?
RQ2.1: Is the distribution pattern persistent?
RQ2.2: Within each class of facilities, is the ordering of premises consistent?

Following on from the above, research question 2 then considers risky facilities over time, a further under-researched area. RQ2.1 involves studying the distribution of crime over a ten year period to ascertain if the risky facility pattern persists, whilst RQ2.2 addresses the consistency of this distribution, by looking at whether the same addresses tend to be risky across the whole period studied or not. Again, this responds to a gap in the literature in terms of longitudinal analysis of risky facilities, and suggestions are made as to how this could be expanded upon in the future.

RQ3: How is the concentration of crime within facilities distributed?
RQ3.1: For each class of facility, does the concentration of crime best fit a proposed power-law distribution?

In response to some claims in the literature, research question 3 involves testing the goodness of fit of a power-law distribution to the risky facility pattern (J-curve) found in the current study. This is important as if this distribution is rejected, then not only should authors refrain from referring to it as such, but it also eliminates the power-law distribution as a potential method of identification, both for the current and future studies.

Aim 2: to consider how risky facilities can be identified.

RQ4: Can risky facilities be defined empirically?
RQ4.1 What quantitative methods could be used to empirically define risky facilities and what are the implications of using these?
RQ4.2: Which quantitative method is most appropriate?
RQ4.3: Where do operational police officers perceive risky facilities to occur and how does this compare to quantitative identification using recorded crime data?

Research question 4 seeks to deal with issues relating to identifying what premises are, or should be, labelled as ‘risky’ and responds to the gaps in the literature regarding the appropriate use of the risky facility label and the need to develop and apply a consistent ‘cut-off’. Research question 4.1 asks what quantitative methods could be used to empirically define risky facilities and what are the implications of using these? A series of parameters are established for what methods should be included for consideration, a number are selected and their possible advantages and disadvantages set out. These are then applied to the study data and the empirical implications considered. Following this, research question 4.2 asks which quantitative method is most appropriate? This is answered with reference to the issues considered for question 4.1, as well as practical implications.

Research question 4.3 introduces the opportunity to analyse qualitatively collected data to consider police officer perceptions of what premises are ‘risky’ and to compare these with the facilities identified as set out above. This is important, as police recorded crime data do not always necessarily reflect those locations that place the greatest demands on police time and resources, nor those premises that are considered the most problematic regardless of the number of offences that are actually reported and formally recorded for that address. Additionally, the analysis of interview data, my own observations and informal conversations with other officers, provide the opportunity to explore some of the reasons why such premises might be recorded and/or perceived as risky and what is ‘going on’ at them. This would not be possible from analysing quantitative data alone and yet it is an approach that is notably neglected in spatial concentration research and even to some extent in contemporary crime and place study more broadly.

Aim 3: to explore key locational features associated with risky facilities.

RQ5: Where are risky facilities located and what types of crime journeys are made to them?

RQ5.1: How are risky facilities distributed across the study area?

RQ5.2: What distances of journeys-to-crime are associated with risky facilities (compared to non-risky facilities)?

RQ5.3: Do distances of crime journeys differ by (a) facility type; (b) crime type; (c) time of offence?

RQ5.4: What are operational police officers’ perceptions of journeys-to-crime at risky facilities?

Research question five begins to consider some important features in relation to location and the environmental backcloth against which risky facilities are situated. This study explicitly refrains from seeking to *explain* risky facilities, instead taking a step back to further explore their existence, and compare this type of distribution across all facilities within a city (and over time). However, as has already been noted, there has been very little research on the locational attributes of risky (versus non-risky) facilities, beyond some consideration of siting/co-location. Here, then, I begin to rectify this, by considering both qualitatively and quantitatively where risky facilities (and non-risky facilities) occur within the study area generally, and with respect to offender residences (RQ5.1). I then consider journey-to-crime in relation to risky facilities, by analysing and comparing the distances travelled to risky and non-risky facilities (RQ5.2), and comparing these by facility type, crime type and time of day (RQ5.3). Finally, in RQ5.4, this is also considered in the context of qualitatively gathered data regarding the trips taken to offend at such premises (as perceived by police officers).

Though journey-to-crime is only one element of the locational and environmental features that have been neglected in risky facility research, it is such an integral part of the crime and place discourse that I have chosen this as the most appropriate place to start expanding the research agenda.

Chapter 6: Overall Methodology

6.1 Introduction

In order to address the research questions that have been presented, a pragmatic mixed methods approach was taken. This utilised both quantitative and qualitative approaches, most notably secondary analysis of police recorded crime data and content and thematic analysis of interviews with serving police officers. Supplementary data were also collected through asking interviewees to annotate a map of the study area with 'problematic' locations, and through field notes compiled by the researcher whilst undertaking independent and accompanied observations. Illustrative (and clarificatory) data were also obtained using Google Maps and Google Streetview.

This chapter sets out the methodological considerations relevant to the study, the methods employed and overarching or study-wide issues relating to data preparation and analysis (with more specific method and analysis considerations dealt with in each of the subsequent study chapters). The methodology adopted is considered by moving through the different levels of research design, presenting my 'worldview', setting out the research strategy and approach taken and introducing the data gathering tools. In this section a number of practical considerations that impacted upon the research are also presented, and there is discussion of the approach taken to ethical issues.

The chapter then moves on to critically establish the methods of the study, covering data collection and data preparation. Although a mixed methods approach is adopted, the quantitative and qualitative data were collected and prepared separately, therefore this section is also presented in this way. After consideration of each approach, the data in question are briefly explored, presenting descriptive statistics to summarise the recorded crime data, and providing an overview of the qualitative data, including the themes identified through the thematic analysis.

Throughout, limitations and implications that are directly related to methodological concerns and the data preparation and analysis are considered, but the chapter concludes with a final overview of the key limitations of and justifications for the approach taken as well as summarising how the different elements of the research are used to address the various research questions.

6.2 Research Design

6.2.1 Worldview and orientation

In terms of my ontological position I would classify myself as predominantly realist, believing that phenomena have their own reality and are not (necessarily) dependent on the social actors involved (Bryman, 2015). However, I do recognise that people bring their own subjective meanings and that their reality could be construed as 'constructed'. I also believe that *crime* is a social construction, in the sense that a 'crime' only exists if it is recognised as such (through someone being aware it has taken place and, more importantly, because those with the power to do so have labelled this behaviour as criminal) (Becker, 1963; Henry, 2007). However, accepting this is the case, once that crime has then been recorded, I believe this becomes a social reality, thus when looking at the concentration of recorded crime (as opposed to deviancy, risk, safety, and so forth) it is appropriate to adopt a realist ontology.

This approach is common within environmental criminology; constructivist ontologies are rarely explicitly seen or discussed when considering spatial concentrations of crime, though there are some examples where this may be implicit (e.g. McCord et al., 2007). However, there *is* recognition of the validity of the constructivist approach in both theory and practice, as rational choice and situational crime prevention both refer to the 'perceptions' of the potential offender (thus how they understand and construct the reality around them) and in the concept of 'thinking thief': where the researcher or practitioner is called to put themselves in the shoes of a potential offender, to see the world, thus the crime opportunities, as they may do (Ekblom, 1997).

Therefore, although in a realist sense, I rely on recorded crime as a measure of social reality, I also recognise that this might not be the *only* reality thus to some extent I seek to consider how realities are constructed from more than one viewpoint. This is most notable when considering the interview data (as *perceptions* of the police) and, of course, in my own observations.

Epistemologically, it could be argued that this research is predominantly post-positivist. Positivism in the social sciences, in simple terms, relates to applying the rules and approaches of natural sciences to the study of the social world (Bryman, 2015). Thus it involves actions such as measuring and testing (and is most associated with quantitative approaches). Positivism has faced much criticism regarding the ability of any research carried out in a 'real world' environment to be objective, neutral and value-free and to identify 'the truth' (as summarised, for example, by Bryman, 2015 and Johnson & Onwuegbuzie, 2004) and it has been 'tainted' by the use of the term, as exemplified in criminology through the work of Lombroso *et al*, by particular schools of thought built around extremely reductionist (and misappropriated) views (Robson & McCartan, 2016). Therefore, contemporary positivism is sometimes referred to as post-positivist and takes a somewhat softer view, whilst remaining paradigmatically distinct from interpretivist approaches (Robson & McCartain).

The majority of the research reported here, sought to apply, as much as was possible, the scientific method and the quantitative approach dominated (as discussed below). However, qualitative methodologies were also used, and whilst I sought to limit the impact of my own values and experiences, I recognise that these will have influenced the interpretation of my observations. In addition, I have actively sought to analyse the perceptions of serving police officers, in order to produce a more well-rounded and complete analysis of risky facilities than can be achieved through a strictly post-positivist approach. These approaches are not, despite much literature to the contrary, incompatible (Bryman, 2015) and qualitative approaches are not 'unscientific' by default (Robson & McCartan, 2016).

With regards to the reasoning applied, a deductive approach was taken throughout (Bryman, 2015). The research was viewed through a particular theoretical lens (the collection of approaches known as opportunity theories), and this directly influenced the research questions that were developed, the methods of data collection and the subsequent analysis and interpretation of results. Some parts of the study also specifically developed hypotheses to be tested. On the other hand, in the thematic analysis I was certainly open to 'discovering' new or competing explanations (features of inductive reasoning).

Taking all of this into consideration, my rejection of a binary approach to research (characterised as realist/empiricist positivist research using quantitative methods versus constructivist interpretivist qualitative research) my worldview, and the research reported here, are most appropriately labelled pragmatic (Robson & McCartan, 2016; Creswell & Creswell, 2018; Johnson & Onwuegbuzie, 2004), or perhaps more accurately *realist pragmatic*.

To summarise, then, the research orientation of this study is most appropriately categorised as hypothetico-deductive pragmatist. This research philosophy is particularly suited to the mixed methods research approach, which is now considered.

6.2.2 Approach

The pragmatic approach suggests that researchers use what is best for the question or issue under exploration. This means they might take different stances with regards to world view and that, arguably, philosophical concerns of epistemology and ontology become less important (or at least certainly less entrenched) (Robson & McCartan, 2016). This is well aligned with the mixed methods approach, that uses more than one form of data collection, source, or analytical approach to address a research question or questions. This approach does not seek to bridge the gap between quantitative and qualitative approaches (as this itself can reinforce the dichotomy), but rather it should aim, at least in part, to 'mix' the data, or at least use it to address the same or closely related questions (Bryman, 2015). As Bazeley (1999: 285-6) says: "'Good' data analysis and 'good' theory, regardless of whether the evidence is from qualitative or quantitative sources, requires a sensitivity to what the data are saying to be able to ask sensible and intuitive questions of the data in the first place." There is an extensive literature on the history, development, acceptance and critique of mixed methods approaches, which is well beyond the scope of this chapter. Instead here I briefly consider models of mixed methods approaches, relate this to triangulation, and explain how the current study fits this approach to research, though first I briefly rehearse the relative strengths and weaknesses of quantitative and qualitative approaches.

6.2.2.1 Quantitative 'versus' qualitative research

Often viewed as a dualism, there is a long history of tension between quantitative and qualitative research approaches, which is particularly vociferous amongst those who see them as closely tied to particular epistemological orientations (Bryman, 2015). Those who take a pragmatic view often employ mixed or multiple research methods so that the weaknesses or limitations of each are counteracted by the strengths of the other (Johnson, Onwuegbuzie & Turner, 2007). Quantitative approaches facilitate larger-scale, large sample projects that can collect substantial amounts of data in a relatively short period of time. In this case, the quantitative data took the form of 'official' statistics[1]. As discussed below, these were difficult to secure access to, but once this was in place, the data were quickly obtained, with no costs to the researcher. It would not have been possible to test the 'ubiquity' of the risky facility distribution without data of this extent. Quantitative research also tends (though this will be discussed in more detail below) to score more highly in terms of reliability and replicability, as well as external validity/generalisability, though this does depend on the nature of the source data. Here, for example, the findings are not generalisable outside of the source country.

Quantitative data, by its scope and nature, also tends to be more amenable to statistical

[1] The label here is loosely applied as meaning data collected in an official capacity for some reason other than research, usually by governments or government agencies, as an adaptation

analysis, in particular hypothesis testing (Bryman, 2015). This was important for this study, as I needed sufficiently large sample sizes, and numerical data - or data that could be reduced to nominal categories - in order to investigate patterns in crime counts and distributions, as well as to test hypotheses relating to the extent and nature of concentration or unequal distribution of crime. Such data, and the associated quantitative analysis methods, are therefore particularly suited to identifying significant differences and correlations in data, but they are not as well-suited to considering meaning, and arguably, although they may be able to test proposed explanations, they are not *in themselves* particularly useful for identifying what these possible explanations might be. For this, we need to turn to theory, but can also draw on qualitative research findings.

Qualitative research tends to use smaller samples, and involve the collection of data in the form of words, although this can also be pictures, media or other visual representations. One of the main criticisms of quantitative approaches is that they are too reductionist taking the view that social phenomena can be converted into numbers, which in turn can be subjected to statistical manipulation and analysis to provide 'answers' to complex social concerns (in this case to questions relating to the existence, nature and features of some problematic premises to be labelled as risky facilities). On the other hand, words and discourse hold greater meaning and can better encapsulate the conflict, complexity and nuance of social life. Quantitative research is also criticised from an interpretivist perspective for believing the numbers it uses, such as crime counts, are a realistic measure of the concepts of interest, rather than *one* socially constructed reality, out of many (Bryman, 2015).

Qualitative approaches, in contrast, seek to interpret perceptions, and to 'translate' the views, opinions and understandings of those who participate in the research (which may include the researcher themselves) to an interested audience (academia, policy-makers, and so forth). In this study, qualitative research of police officer perceptions of high crime areas and problematic premises is juxtaposed against recorded crime, in recognition of the fact that those behaviours that end up in 'official statistics' do not represent all criminal (or antisocial) behaviour that actually occurs, or necessarily represent the issues of most concern to patrol officers, or the general public. However, as only police officers were included, this still represents a very narrow, and likely institutionalised, view of crime and resource demands. In an effort to somewhat ameliorate this, researcher observation (a limited quasi-ethnography) was employed as a further qualitative technique, to supplement the major research methods. This introduced a different perspective (although one that was still influenced by my experience working in, and with, police organisations, albeit only in the UK), but also further possible biases resulting from

---

of the definition applied by Bryman (2015) subjective interpretation. Additionally, the insights of these observations were limited to issues regarding the environment and locations of risky facilities, or the buildings themselves.

Ideally, other parties affected by the existence of risky facilities would also have been included in this research, notably place managers and local residents. Unfortunately, access and scope constraints made this unrealistic, for this project. However, given the benefits that diverse perspectives could bring (in terms of identification and developing explanations), I suggest there is a need for specific qualitative research of this type as part of the study of the risky facility phenomena, developing on the work undertaken by Madensen (2007) with place managers of bars, but with an expansion of the qualitative data collection and analysis (see, for example, Dymne, 2017).

6.2.2.2 Using a mixed methods approach

Mixed methods research is often used synonymously with multi-method research, but, as stated, a truly mixed methods approach seeks to mix the data or to use more than one source or approach to address the same research question (Bryman, 2015). There are examples of multi-method and, arguably, mixed methods research within place-based criminology (a notable example is Weisburd et al., 2006). There are also a small number of qualitative placed-based criminology research studies (Dymne, 2017; Kooi, 2015; St Jean, 2007). There is also evidence of a growing recognition of the benefits of such qualitative and/or mixed methods approaches, in a field that has been dominated by quantitative approaches (for example, Greene, 2014; Telep, 2018). However, quasi-experimental designs and multivariate statistical analysis remain the 'go to' standard in this field. These techniques, particularly coupled with spatial statistics, are clearly important tools to use, given the nature of the entities and phenomena being studied. However, I contend that they are not capable *alone* of helping us truly make sense of crime interactions, whether our unit of analysis is spatially defined or not. Instead, we need to be able to develop and test (and report on) a number of contextually situated mechanisms (Pawson & Tilley, 1997; see also Sidebottom & Tilley, 2012) that might explain the existence of facilities as risky; and this is most likely to be achieved using a mixed-methods approach (Robson & McCartan, 2016).

6.2.2.3 Models of mixed methods research

There are different iterations of mixed methods research, which vary by the weight and order given to the quantitative and qualitative elements and the purposes that these serve (Bryman, 2015; Robson & McCartan, 2016; Teddie & Tashakkori, 2009). In this study the data collection was completed sequentially (quantitative data were obtained first) but much of the analysis was carried out concurrently. Where different methods have been used to address a research question, the results are first presented for the quantitative analysis, then for the qualitative, this second section referring back to the quantitative findings and incorporating the discussion. This is mainly for clarity, rather than to suggest one approach is favoured, however, it is acknowledged that the quantitative and statistical analyses do hold something of a position of dominance. This may be, in part, because of the influence of environmental criminology's prevailing epistemological stance. In summary, I have used a mixed methods approach, endeavouring to employ a convergent parallel design (Bryman, 2015), with the data from both quantitative and qualitative elements being compared to produce the findings. This design is closely associated with triangulation, which is discussed below. However, I would argue this approach can overlap with, rather than be distinct from (Bryman, 2015), an explanatory sequential design (where the qualitative data have been used to explain or elaborate on the quantitative findings). Realistically my approach lay somewhere between the two.

6.2.2.4 Triangulation

Mixed methods research can encompass, or facilitate, triangulation. There are four main types of triangulation, and this research study used two: methodological triangulation and data triangulation (Arksey & Knight, 1999; Robson & McCartan, 2016; Teddie & Tashakkori, 2009). Bryman (2015) argues that effective triangulation, where the results gained from two different research strategies are cross-checked against each other, allows us to have greater confidence in the findings. When the findings do not correspond 'perfectly' (as is likely to be the case) then this can be seen as an enhancement of the findings (Bryman, 2015). Where the difference is greater, Bryman (2015) argues that the researcher needs to give greater weight to one approach, or to re-examine the data. Alternatively, Arksey & Knight (1999) believe that such differences can be embraced (as they recognise neither will be all encompassing, so triangulation is being used for the reason of completeness) or can be used to check for errors or flaws in approaches or research instruments. I would expand this to say that differences could be explained by reference back to the traditionally associated ontological and epistemological positions (for example, one set of data is presenting the pattern in crime that has been reported and recorded, whilst the other set of data shows how the police perceive and interpret crime in their locale). I would also argue that differences may suggest the need to gather further data, possibly using an alternative strategy. In interpreting the results of this study (where mixed methods have been used) I have endeavoured to interpret any discrepancies across the findings from these different perspectives, as appropriate.

6.2.3 Design frame

The experimental approach is often viewed by positivist researchers as the ‘gold standard’, of hypothetico-deductive research strategies. In experimental approaches variables are manipulated and control groups are used to objectively measure the effects of these changes. When applied to social science research, this approach is usually referred to as a quasi-experimental approach, as it is very rare to achieve such high degrees of control, and exclusion of confounding variables, in natural settings. Whilst my research design was not experimental, such design highlights the importance of comparison (Bryman, 2015), which was an important feature of the current research. Indeed, it is based very heavily on seeking to compare across types of facility, across facilities, across locations, between risky and non-risky facilities and even across data sources.

I characterise the research design employed, therefore, as cross-sectional, with an emphasis on comparison. It is cross-sectional as it, generally, used a snapshot of recorded crime, police perceptions and researcher observations at one moment in time and the variables used were non-manipulable (a key feature of cross-sectional research (Bryman, 2015)). However, I do also consider changes over time (or rather persistency and consistency of patterns), therefore there are elements of longitudinal research included in the study as well. As will be seen, below, this was facilitated by the provision of crime data that covered a ten-year period.

6.3 Data collection

It is important, particularly when adopting a pragmatic stance, carrying out ‘real world’ research and/or utilising a mixed methods approach, that the ‘right’ (most appropriate) data needs to be gathered using the ‘right’ tools (Bryman, 2015). Determining what is appropriate may be affected by a number of different influences, including research orientation (less of an issue here), theory, custom (within the field of study), and the confidence and experience of the researcher (Bryman, 2015; Boyatzis, 1998; Robson & McCartan, 2016). With regards the latter point, being “…methodologically multilingual…” (Robson & McCartan, 2016: 30) I strove to select the tools that would best allow me to meet the study aims. As with all research, however, practical concerns also came into play, particularly regarding access to appropriate data (these are further discussed below).

Ultimately, the data collection used in the research was (1) the collection of ‘source’ data (police recorded crime) for the purpose of secondary analysis; (2) interviews, with serving police officers (plus map annotation); (3) supplementary data, gathered through a field study (using both structured and unstructured observation, as well as informal conversations).

The strengths and weaknesses of each of these approaches are considered in the relevant (quantitative or qualitative) methods section below. A table, summarising the research questions and the data and analysis techniques used to address these can be found at the end of this chapter.

6.3.1 Design-related issues

6.3.1.1 Quality criteria

Quality criteria in quantitative research relate to three key areas: reliability, replicability and validity. In qualitative research alternative terminology or interpretations are often used as these concepts do not always sit well with the techniques or aims of qualitative research. Bryman (2015) summarises an alternative criterion that could be applied, as authenticity and trustworthiness. This includes such concepts as credibility, transferability, dependability and confirmability. However, he goes on to say that in many cases a slight adaptation of the quantitative research criteria tends to be used. As I took a mixed methods approach, there are yet further suggestions for ensuring quality in research, include clearly articulating the aims and purposes of the different methods/sources used and their role in the overall research strategy (Bryman, 2015; Teddie & Tashakkori, 2009). I have endeavoured to do this, throughout. Borrowing from more traditional qualitative methodologies I have also introduced an element of reflexivity, as well as a degree of self-awareness relating to the role I played in the data collection and interpretation of the findings, particularly with regards observer bias, as I selected *what* to observe and interpreted these observations from a particular frame of reference and set of experiences and knowledge (Foster, 2006). Beyond this, the usual quantitative quality criteria apply.

Reliability relates to the consistency of the measures involved, in other words the ability of a research instrument to produce the same results if repeated (Boyatzis, 1998; Bryman, 2015). Ideally test-retest (for quantitative) or inter-rater reliability measures (for qualitative) are employed to assess the reliability of research. This was not possible here, although as will be seen in chapter 7, the measuring of persistency and consistency with respect to the risky facility patterns over time, could be argued to demonstrate that both the concept (risky facilities), and this research, demonstrate a good degree of reliability.

Validity is deconstructed in many different ways. Key forms of validity to be briefly considered are construct validity, internal validity, and external validity (here, predominantly ecological validity) (Bryman, 2015), to which I also add descriptive validity. Construct validity relates to the extent to which a test is measuring what it purports to be (Farrington, 2003). In this study, it relates to the validity of the crime data in representing the actual experience of crime (which has been discussed above as a potential weakness of using recorded crime data). In the latter parts of the research, it also relates to the risky facility construct and the appropriateness of the selection method I employ (thus relevant issues are discussed in chapters 8 and 9). Internal validity concerns the appropriateness of the conclusions being drawn from the findings. In order to maximise internal validity, I discuss very carefully and explicitly the conclusions that I draw and the limitations and caveats applied.

External validity is taken to mean the extent to which the findings are applicable in other settings (Bryman, 2015). This is akin to generalisability. It is often the case that maximising internal validity through very specific (often qualitative) methods, has the effect of reducing the external validity and vice versa. This ‘either-or’ pay-off can arguably be reduced through the appropriate use of mixed methods research (Teddlie & Tashakkori, 2009), as I have sought to do in this study. As will be seen, the overarching findings of the current study are generally consistent with theory and extant knowledge, suggesting a greater potential generalisability. However, given the importance of context in explaining outcomes, and the particular environmental and demographic characteristics of the study area, the ecological validity is likely to be low outside the source country (the USA), and the results may also not be generalisable to rural or highly suburbanised areas, or indeed large cities with different facility provisions and distributions.

Though discussed predominantly in relation to evaluation research, a final dimension of validity is also applicable: descriptive validity. This relates to the reporting of key features of research (Farrington, 2003; Sidebottom & Tilley, 2012) and effort has been made to achieve high descriptive validity in this research, with presentation of source, data, coding, definitional, analytical and interpretive features, as well as explicitly considering a range of general and more specific limitations, throughout. This type of validity is also noted to be of particular importance in making research policy-relevant (Gill, 2011).

In terms of the presentation of research studies, there is some overlap between descriptive validity and replicability. Replicability of research requires that information about the way the research was carried out (such as the gathering of data, sampling techniques, inclusion criteria, and so forth) is provided, such that the research could be repeated by someone else (Bryman, 2015). Valid replicated research should produce the same (or at least similar) results. I have provided detailed information regarding the research processes followed, as well as taking care to articulate how I have interpreted results. However, the specific conditions of the qualitative elements of the research, and my (unstructured) observations in particular cannot truly be reinstated or repeated, therefore for the qualitative aspects, replicability is lower.

A final issue that needs to be mentioned is that of positionality. This relates to my reflections on my position within the research, the impact my presence has had on the data I gathered, and the role that my own experiences and knowledge might have had in biasing my interpretations or the weight given to the conclusions I have drawn (Boyatzis, 1998; Bryman, 2015). This is particularly relevant regarding the qualitative elements of the research, and is discussed alongside these.

6.3.1.2 Sampling

Related to issues of research quality, understanding the impacts of a chosen sampling approach is important to arguably all forms of validity. The sampling strategy adopted here is an opportunity sample, thus the external validity of the research in particular is reduced (Bryman, 2015). With regards the interview data, the participants were those officers who were available and willing to take part. There were no attempts to select either a random (probability) sample of officers, or to approach participants that might represent different levels of experience or specialism. This was a result of the management of the research project, and the need to limit my impact on the organisation. The sample size was also small, though for a mixed methods study of this kind, the data collected from these was reasonable. Other specific issues relating to the sample are discussed as they come up throughout the study.

Regarding the quantitative elements of the study, it could be argued that these are based on the whole population, rather than a sample, but this is not true if one considers that not all crime (or offenders) in the study area are represented, nor if one sees the study area as being the sample of the wider population (for all cities in the US), in which case this too is an opportunity sample, based on access constraints (discussed elsewhere).

6.3.2 Practical matters

Bryman (2015) (like many others) identifies a number of practical matters that impact upon research decisions, outputs and the conclusions that can be drawn from the findings. It is important to document these, for transparency and to seek quality in research. He particularly refers to access issues, an awareness of the politics of research, values in research, and ethical considerations. These are now discussed in turn.

6.3.2.1 Access

At the time I commenced this PhD, I worked in an environment that allowed direct access to police recorded crime data for one police service in England. As part of my job, I had been security-cleared by this organisation and was already undertaking analysis and publishing research reports and academic articles based on these data. However, before I reached the data collection and analysis stage of the project, I moved jobs. This left me in the position of needing to apply for police recorded crime data from any service or services who were willing and able to provide this to me, with no special right to access. I also required certain data fields that tend to be recorded in a way that would have made it harder for police services to provide me with completely anonymised data. The parameters and fields that were essential to address the research questions were:

- Crime committed in facilities (as defined, above) - I could extract this myself, but this created greater problems with anonymity as to do so would require access to a dataset that was more likely to have personal information (residential crime and that including inter-personal violence with named victims)
- Time and date
- Crime type
- Locational information - preferably a geocoded location, though it would have been possible to geocode this myself if it was necessary. I also needed the full address of the crime location (facility) for the analysis I wished to carry out (discussed below). The street segment, street, postcode or census output area were not specific enough.

It was the requirements relating to location that created the most problems regarding access. It became apparent that many of the police services I approached were concerned about anonymity issues, and it was also the case that it was considered quite a significant amount of work on their part to redact potentially identifiable information. One service that expressed a willingness to help, was unable to do so because the information on location type was incomplete for far too great a proportion of their records.

Eventually, access was secured to data from a police department (PD) in the US. They were able to provide me not only with city-wide recorded crime data that met my requirements, which was also for a ten-year period - far longer than I had anticipated obtaining, but also related offender data (with names removed) and both geographic and census data for the relevant area.

This allowed me to expand my research questions to cover comparison of journey-to-crime as well as to map the data against the city backdrop. Descriptions of the city, and the data provided, are given below.

I had originally intended that one of my contributions to extant knowledge would be a consideration of particular types of crime concentration at the micro-level in the UK, where such research is far more limited than in the US. This would have been an advantage of securing data from within my home country. Local (UK) source data would also have more readily facilitated the qualitative data collection. However, as such data appeared unobtainable, the US data were used as my sole quantitative dataset for the research. As well as the benefits of being able to expand my research questions, these data also allowed me to more directly compare my findings with those in the existing literature, as the type of data, type of location and type of society under study (in broad terms) were more similar to the majority of published findings on crime generators/ attractors and risky facilities, being situated in North America.

Finally, with regards access, I was able to undertake a ten-day research trip to the city in question for the purposes of clarifying data-related issues, gathering further general information about the location, policing procedures, and so forth and to collect the qualitative data (as well as to get a general 'feel' for the location, which was particularly important as a foreign researcher). I organised this myself, but was hosted by the PD, and I was also granted access to police staff for informal discussion and formal interviews, as well as being shown around the city and accompanied during some of my structured observations. The remainder of the time I undertook further structured and unstructured observation of the city, alone.

6.3.2.2 Politics

Using secondary analysis, particularly of official statistics or other data collected by agents of the state, could never be described as apoliticial. Politics and policy decisions impact access to the data, the data that are collected, why and how they are collected, and what might feasibly be interpreted from them. This might also be seen as influencing what is referred to as the 'dark figure of crime' (Maguire & McVie, 2017), as the relationship that citizens have with the police (and the views they hold about them) will affect reporting rates; attitudes of the police, institutional customs and direct policy orders will affect what reported crime they record and also what types of offence they focus their attentions on; policy and law-making will also determine what is construed as a crime. All of these issues and attendant decisions are affected by politics (and cultural contexts) and they all impact the data that is collected, particularly police recorded crime data, as used in this study.

Taking a pragmatic, mixed methods approach to research may help ameliorate some of the worst biases of official statistics, though in the current study, as the main qualitative approach involved interviewing police officers, some of these political and institutional influences are likely to remain. It is important that these be acknowledged when interpreting the findings and what can be concluded from them.

6.3.2.3 Values

Much like politics is an ever-present, no research can be value-free (Bryman, 2015). Even the most strictly formulated laboratory experiment involves potentially subjective decisions, such as which equipment to use, which brand of reagent, whether the lab-work ought to be carried out by postgraduate assistants, or highly experienced research staff, and so forth. Many of these decisions will be influenced by personal values and experiences. There is an important debate regarding the worth and role of values in research, from the extreme positivist empiricist end of the continuum, all the way through to the researcher-as-activist approach. However, this is beyond the scope of this thesis. I recognise that my values regarding the legitimacy of policing, the importance of analytical and evidence-led approaches to resource management and intervention and the ability of such agencies to impact crime rates have influenced my choice of research topic, questions and methods and, where necessary, I reflect on these values and experiences and their relevance to the research, as I progress.

6.3.2.4 Ethics

All research has to grapple with ethical considerations and ensure that all feasible steps are taken to minimise harm to participants, the researcher and the discipline (it might also be argued to society as well, though designing research to prevent the misappropriation and misuse of its findings, can be difficult). This research study was designed to comply with the ethical considerations set out by the British Society of Criminology and according to the requirements of University College London (UCL) as they stood at the time when the research was approved.[2]

There have been some changes in the university research landscape regarding ethics and research governance, that came about after the data had been collected. However, I refer to the more recent ethical considerations as set out by the British Society of Criminology (BSC) (2015), as my research met these requirements in any case, and because ethics should be “addressed

[2] The research was deemed exempt from Research Ethics Committee approval, confirmed by then UCL JDI Head of Department. A copy of this email confirmation is in the appendices

throughout the whole life of a research project and not just at the outset" (Abbot & Sapsford, 2006: 293). Given that no vulnerable individuals were involved in the research and that the participants are anonymised, however, the requirements (which are quite extensive) are summarised only briefly here.

As a researcher my work seeks to advance subject knowledge and I have ensured I am appropriately experienced and trained to make decisions about, and carry out, research in this field, using the range of methods employed. I have also ensured that I do not misrepresent myself, my research or my qualifications and that I have acted professionally throughout (BSC, 2015: 3). I have ensured that I have adhered to the requirements of the Data Protection and Human Rights Acts, taken caution to protect my data and kept identifiable information regarding respondents separate to the anonymised transcripts of their interviews (BSC, 2015: 3). It should be noted that no other identifiable personal data was used in the research, beyond addresses where crimes were committed (though no victim data was recorded) and addresses at which offenders lived (again no personal identifiers were provided).

As the data were provided by an American PD, I also sought my contact's advice on storage and use of the data, but no further restrictions (including of the Department's anonymity were placed on me). However, because I need to refer to specific premises/addresses and quantities of crime committed there, I have taken the decision not to name the city upon which the study is based. I have also taken steps to partially anonymise premises (e.g. using abbreviations in names or referring only to an associated street name). Therefore, when facilities are talked about, the name may be quasi-pseudonymous. Because of the need to refer to locations, street names (as a way of distinguishing multiple premises of the same chain) and present maps, it is recognised that the examination version of this thesis will contain material that could be used to identify the city, and thus the facilities within it. Although this is not strictly a breach of ethical protocols, and is in-line with the conditions of data provision, I have made the decision that the final, deposited version of the thesis will use redactions so that identification of the city is not possible (although it is recognised that with determined investigation, these efforts might not be foolproof).

Regarding responsibilities to research participants (BSC, 2015: 5-9), the consent form and cover sheet used for the interviews can be found in Appendix 1. This, plus my further explanation at the beginning of each interview, sought to ensure that participation was willing and voluntary and that any risk of harm to participants was minimised (in relation to both psychological harm - for example I did not ask questions about specific experiences of carrying out their role, and possible professional harm - for example I did not ask any questions that might have resulted in officers incriminating themselves). By clearly setting out what the research was looking at, the other approaches I had used, where the interviews 'sat' within the wider research strategy, and how the data would be used, it can be concluded that the respondents were giving informed consent. Participants were also given the right not to answer any questions they wished during interview, to withdraw consent at any time and to request to have their data removed and destroyed at any time up until the analysis had commenced. They were also advised that the data, including all copies of transcripts and audio recordings (for which separate consent was also obtained) would be destroyed after conferment of my award.

During my observations, I did not record information about, or take photographs of, specific, identifiable individuals. All observations were made in public places. Therefore, it was unnecessary to gain 'consent' from anyone who might have happened to be 'passing by'. It could be argued that as I have carried out research on named, specific premises, that permission ought to have be sought from the owners or place managers. However, the data owners and providers were the police as any information collected was about crimes, and although further information about places was provided during interviews, individuals were not named and there was nothing reported that could be said to be sensitive or personal information. Therefore, my judgement (Abbot & Sapsford, 2006) was that this would not cause harm, and no further consent was needed.

I was also sympathetic to the impact of my research on the organisation (BSC, 2015: 6) I ensured that I carried out the interviews at times when participants were already on the premises and were free to talk to me, and if they were still on duty, they were not asked to turn off their radios, so that they could respond to assistance calls if needed. Regarding the observations, some of these were accompanied, but this was at the offer of the PD; it was not something that I had specifically requested.

Finally, researchers have a duty to minimise the risks they are exposed to in undertaking research (BSC, 2015: 6). I carried out a research risk assessment for the field trip, including issues such as travel, health risks, crime and security threats and so forth and determined that the risks were minimal or could be limited by taking appropriate action (such as not carrying large amounts of cash). When undertaking formal observations on my own, thus visiting locations that experienced potentially high amounts of crime, these were mainly carried out from my hire car. I did enter some of the larger retail premises, but avoided leaving my car in many areas due to the risks of being victimised. I visited the areas that might be thought of as 'most dangerous' and where my presence would seem out of place, whilst accompanied by a police officer and on only one occasion, at the instigation of my contact who was also present at that time, did I leave the car to be shown around a high crime neighbourhood convenience store, but on this occasion we were accompanied by an armed, uniformed officer.

6.4 Methods

6.4.1 Quantitative data

As stated above, the quantitative data collected was ‘source’ data, or ‘official statistics’. Some of the strengths and weaknesses of this type of data, and the use of secondary analysis, have already been considered when discussing quantitative ‘versus’ qualitative approaches. In summary, using data such as police-recorded crime has a number of weaknesses and flaws (May, 2011). Most notably these relate to the completeness and accuracy of the data, both in terms of individual crime records, as well as the degree to which crime that is recorded truly represents the nature, extent and *crucially* patterns (Skogan, 1977) of illegal behaviours; the so-called ‘dark figure’ problem (Biderman & Reiss, 1967; Maguire & McVie, 2017). Criminal justice agencies, governments and researchers have, of course, developed many ways of collecting alternative or supplementary data in efforts to respond to this concern, but ultimately much research on crime phenomena - particularly larger studies - uses such sources of data, as the most comprehensive and consistent source of information about crime (May, 2011). That said, it is still important to note, that what is revealed through the analysis of these crime data may be more a reflection of police activity, targeting and decision-making (plus wider policy-decisions) than of the truly risky or dangerous places. The use of a mixed methods approach may have ameliorated this to some extent, but given that the other main source of data was serving police officers, this may still skew the patterns revealed in the findings. That said, if we interpret ‘problematic premises’ (as synonymous for risky facilities) as problematic *to the police*, then the data can be considered as much more representative. Where these issues particularly impact on the results or their interpretation, they are raised again.

Incompleteness is also a cause for concern when analysing recorded offender data. Without rehearsing all the arguments (again) such data have to be recognised as particularly unrepresentative of actual offending patterns and offender characteristics, as relatively few crimes (especially minor ones) are detected. In the US, in 2010,[3] the clearance rate for violent crimes was 47.2 percent and for property crimes was 18.3 percent (Federal Bureau of

[3] This data was selected as being more comparable with the crime data used in the study than more recent publications.

Investigation, 2011). Therefore, patterns related to offender addresses will be only partially representative, at best, as it is highly possible that those who are arrested and charged with offences are ‘different’ (and may reside in different locations) to those who are not processed for their offending (which we could refer to as being more ‘successful’).

Other concerns regarding secondary analysis more generally relate to the fact the data were not collected for the purposes of the research (Bryman, 2015), therefore although it may be the ‘best’ and most suitable data available, there are often likely to be flaws, omissions and peculiarities for the researcher to deal with (Bryman, 2015; Robson & McCartan, 2016). In addition, variables that *are* collected might have to be used as a ‘proxy’ for data that the researcher ideally wanted. These issues all impact on the reliability and validity of the findings and raise challenges that have to be addressed in the operationalisation of the research concepts and the preparation and analysis of the data. Some of these issues have already been considered in relation to quality issues (above), whilst others are discussed in terms of processes employed, below.

6.4.1.1 Describing the city, the quantitative data and the original variables

6.4.1.1.1 The city

The data collected relate to a somewhat typical, small city, including business, commercial and residential districts. They constitute crimes recorded by one specific police department (PD) as well as interviews with serving officers and observations across the city.

The city is situated in a southeastern US state, and was incorporated in the mid-nineteenth century. Being situated on a major railroad, the city played a role in transportation and distribution, as well as developing a number of local manufacturing industries. Nowadays the city hosts an annual industry-specific trade fair. This impacts upon what would otherwise be the city’s ‘downtown’ area, as for much of the year this consists mainly of empty exhibition spaces.

The population is recorded at around 100,000 to 110,000 and the index crime rate in 2007 was approximately 68 offences per thousand population. The city is also home to a large, comprehensive university. The 2000 census indicates an ethnic mix of around 55% white and 35% African American, with about 7% of the population (of any race) identifying as Hispanic or Latino. Just under half of households constituted a married couple, and about one third of households had children under 18s living at home. The median age of the population was 34 years old, just over one quarter of residents were under 18, and just under one-tenth were aged 18 to 24 (a key age in terms of offending). Slightly less than 15% of the population were below the poverty line (references redacted for anonymity; census 2000 available in parts online at https://www.census.gov/prod/www/decennial.html).

6.4.1.1.2 The quantitative data and variables

When supplied, the recorded crime data consisted of the following variables: incident ID (the unique identifier), offense (type), address, date occurred, hour occurred, day of the week, report area (beat), X and Y co-ordinates, description (location type), subdivision. Each row of data related to an individual incident, and these included events that would not be included in a sample of recorded crime data in the UK (where incident data, also known as calls for service, are recorded separately in an operational information system).

The offender data (a separate file), was made up of the following variables: charge description (what charged with), offence charged (categorical, e.g. driving, animal, child), date arrested, day arrested, month arrested, year arrested, address (which upon further investigation was the address where the individual was arrested), name (which was replaced by a unique number), age, sex, race, case ID (which equated to the incident ID in the crime data, therefore was the variable used for matching), UCR code, and X and Y co-ordinates (for the arrest address). In addition, as I intended to calculate journeys made to commit crime, I contacted the PD and they further provided me with a file containing the name identifier, current address and previous addresses of offenders, that I could use instead of the arrest address.

As stated, base geographic data were also supplied to support any mapping I wished to do. The crime and offender data were supplied in MS Excel format, whilst the base data were supplied in .mdb (MS Access) format. The crime data were converted into an SPSS file where most of the recoding took place. The offender data were more complex to join to the crime data, and this is set out below.

6.4.1.2 Cleaning and coding

In order for the data to be reduced into manageable categories consistent with the analytical aims of the research, many of the variables were recorded or new variables created based on the data provided. For example, offences were grouped into larger categories (initially at a number of different levels of specificity), dates were recoded into financial years, and so forth. The processes and resulting variables are briefly set out next. First, however, it is necessary to briefly mention categorisation.

In reducing data to a smaller number of categories, decisions had to made that would affect the rest of the research, the analysis that could be carried out and the findings that resulted. The aim was for categorisation to be driven by practical concerns (such as the number of categories that was feasible to manage, category counts), conceptual understandings (so that the resulting grouping 'made sense', such as categorising months into seasons (Dec, Jan, Feb = winter) rather than by position in the calendar year), theoretical models (such as gathering together locations associated with similar routine activities), and professional norms (such as using financial as opposed to calendar years, and separating acquisitive from violent crime).

Despite these endeavours, there was inevitable subjectivity in the categorisation process, and in hindsight, as will be discussed later, some of the resulting categories may have been a little too fine, particularly resulting in some small samples when separating offences by facility category and another feature (such as crime type). Overall, however, I felt that the choice of categorisation was defensible given the aims set out above. I would suggest, however, carrying out further research to explore the impact of these decisions, by repeating the analysis using broader categories, particularly in relation to facility type.

6.4.1.2.1 Crime data

The data provided included some records of things other than crime, including calls for service, arrests and citations, and even some lost/found property. As the study was about crime at facilities everything that did not appear to be a record of a criminal offence was removed. The original data covered the period 1/1/2000 to 28/3/2010. This equated to 123,300 records. There were 360 distinct 'offences'. These were recoded at three different levels of categorisation including a category for non-crime. Having removed these non-crime data there were 104,357 recorded crimes remaining. Throughout the analyses, all crime was used, except when specifically looking at crime type, when only the top-level coded offences of *acquisitive* and *violent* were used, as they were deemed the most distinct and of most interest in terms of previous research and policing activity. Having considered the findings of the qualitative analysis, in the future I would also like to explore drug offences (and possibly disorder) for this area.

The location type was also recoded, firstly to identify residential, non-residential and unknown, then to categorise non-residential into facility or otherwise, then to code type of facility (discussed below). Residential crime was removed as were offences not associated with an address (for example STREET).

The focus of the research was on facility-crime only, and as buffers were not being used (see the explanation for this elsewhere), the dataset needed to be reduced to only crime recorded as occurring at a premises, specifically facilities. Those not associated with a building were therefore also removed and the remaining non-residential crime was recoded based on the codes used in the original dataset. This resulted in a remaining subset of records that were non-residential, had a recorded address, but the type of premises was not known, so it could not be determined if it was a facility or not. Efforts were made to recode these manually, from the provided addresses, using Google, Google Maps and Google Streetview. A number of these could not be identified and were recorded as such. In addition, it is recognised that some of the manually coded premises could have changed between when the crime was committed and when the coding was carried out, introducing some degree of error (that was not possible to quantify).

Where the original location type was *not* listed as a type of outdoor space, it was assumed that the offence occurred *within* the premises recorded, or *within its curtilage* (hence the reporting officer made the connection between the premises and the offence). Whilst this assumption can be questioned, I did not want to use buffers or other forms of associating offences with premises, because of conceptual and methodological issues already discussed.

The working definition of facility used in this study was developed in part from Clarke & Eck (2007), Eck et al (2007), and Oxford Living Dictionaries (2018). For this study a facility is taken to mean:

premises provided for a particular purpose or function, to which at least some members of the public are granted access.

Land use that cannot be considered a 'premises' ("a house or building, together with its land and outbuildings, occupied by a business or considered in an official context" (Oxford Living Dictionaries, 2018: online)) was excluded. This removed places such as parks, waterways and streets. Completely private premises and residential premises were also excluded, which is different to the approach taken by Clarke & Eck (2007) and Eck et al. (2007). However, I felt that residential premises have sufficiently different place management practices, access requirements, routine activities and intervention possibilities associated with them, that they ought not to be considered alongside non-residential premises. The facilities themselves may be private or public, but at least some members of the public would be granted access to them. This access may be limited, temporary and/or revocable (as is this case with many of the premises captured by this definition, such as shops, bars, restaurants). Some public buildings (such as schools) are included, but offices and those with solely non-public facing functions are not. Finally, manufacturing and industrial premises, dedicated parking areas, and transport terminals were also excluded. Crime concentrations in these types of land use are extremely interesting and worthy of further research (such as that carried out by Newton, 2008; Newton, Partridge & Gill, 2014), but again the types of location and the routine activities associated with them were considered just too different to the other types of premises to include them in the study, and did not quite sit right as being labelled ‘facilities’.

In order to have time period categories that were equal, and not to lose the most recent data, financial rather than calendar years were used, so the data from 1/1/00 to 31/3/00 were also removed, leaving an overall dataset that covered the period 1/4/00 to 28/3/2010.

Table 1: Crime by location type

Type of location	N	%
Residential	39885	39.34
Outdoor space	24357	24.03
Nonresidential (including facility)	36565	36.07
Unknown	572	0.56
Total	101379	100.00
Nonresidential (not facility)	5830	5.75
Facility	**30735**	**30.32**

This resulted in a facility crime dataset consisting of 30,735 offences (30.32% of the original crime data) as shown in table 1, above. This covered 15 facility types across 2,905 different premises, to each of which I assigned a unique, anonymised identifier (in the created field Addnum). For the analysis presented throughout this study, all facility crime means all of these premises (addresses) and offences, regardless of facility category. Where categories are reported individually or comparatively 12 main categories were used (as shown in table 1) and a thirteenth category of *retail other* is also considered (but the findings are thought of as less useful because of the heterogeneous nature of the retail types included). The remaining two categories (*other* and *liquor store*) are only present in the all facility crime analyses, as *liquor store* only contained 13 premises (the next lowest was *leisure* with 40, but contributing a much greater amount of crime), whilst *other* was not sufficiently homogeneous to allow a comparison of ‘like’ with ‘like’, which is what the risky facility concept requires. The breakdown by location type is shown in table 4 (chapter 7).

6.4.1.2.2 Spatial and location data

Addresses and geocoding

A concern of much research involving spatial crime patterns is the issue of accuracy of the geography. Geocoding, especially that carried out by the police for recorded crime datatsets can be quite inaccurate. This is especially problematic if the geocode is being used to identify the premises or to determine what crime ought to be attributed to it. It seems that this is also more of a concern with American (*cf* UK) data (Ratcliffe, 2012). As Mazeika and Summerton (2017: 460) state: "…the degree to which…efforts and findings are tenable, rests upon, to a degree, being able to properly spatially reference…data, such as offenses reported to police…It is certainly an important endeavor to precisely measure and study micro-level crime generators…but it must be met with an equal zeal to precisely spatially reference the dependent variable." Unfortunately, they find that this is often not the case and that geocoding accuracy is variable (and that hit rate can be artificially inflated by the criteria applied). In fact, their research was able to achieve more accurate geocoding by using Google Maps and Google Earth. As my intention was always to use crime that was *already attributed to facilities* in the dataset, this inaccuracy was much less of an issue in this study.

Crime data geocoding

There were no concerns, then, with attributing a crime to a facility[4] and for much of the research carried out the geocode attached to the facility address was not required. However, analysis using mapped facilities, and the journey-to-crime calculations (offence location) *were* based on the geocodes in the originally provided dataset. Detailed meta-data were not supplied and it has not been possible to ascertain a geocoding hit rate for the police recorded data, but from the information supplied it appears that an address gazetteer was used. This same file of geocoded address point data (part of a set of base geographic data also supplied by the PD) was used to match the offender addresses supplied when calculating journey-to-crime distances (see below).

Research suggests there are likely to be geocoding errors in data recorded by police, but that using gazetteers (that is a database of address point or parcel data) might produce somewhat better results than some of the other common alternatives (such as matching to street centerlines or TIGER[5] data) (Cayo & Talbot, 2003; Hart & Zandbergen, 2012; Mazeika &

[4] From a reference perspective, there were of course concerns about interpretation, completeness and identifying categories and named premises, as discussed elsewhere.
[5] US Census Topologically Integrated Geographic Encoding and Referencing data

Summerton, 2017). Although the hit rate for the police recorded crime data is not known, it can be assumed to be similar to the hit rate produced on the first pass geocoding for offender address data as the same address point data file was used (below). The first pass hit rate, for exact match addresses within the PD, was 94%. Therefore, the anticipated hit rate for facility address geocoding is taken to be the same. This is substantially higher than the minimum rate of 85% proposed by Ratcliffe (2004), though it is recognised that for micro-level analysis, more accurate matching may be required (Mazeika & Summerton, 2017). Overall, given the purposes for which the geocoded data were used, and the assumed hit rate achieved, the accuracy of the spatial data is deemed to be acceptable.

Offender address data

In order to analyse journeys made to crime it was necessary to create a dataset containing all facility crime for which an offender had been charged. There is the possibility that some of these offenders were processed and found not guilty, but these were the only data available, therefore it had to be assumed that the crime journey recorded had indeed been made. A further issue that became apparent upon recoding the data, was that the offender information provided only included an address field for where the arrest was made. Whilst in many cases this was likely the offender's home address, this could not be assumed safely. Using these addresses would have been particularly problematic when calculating distance travelled to crime, as some arrests were made at the scene of the crime, thus the journey was 0km. This could have significantly skewed the results.

The data provider was contacted and agreed to forward further offender address data, but this was limited to a unique offender identifier and a list of addresses. This included historical data, but it was felt that it would be far too time consuming to try to match the crime data to the exact address recorded for the offender at that time. This decision was also made because the dates related to historical addresses represented when the police had contact with the offender and found out about a 'new' address; it was not the date when the offender moved to that address, therefore was still inaccurate. In addition, a quick visual inspection of the data suggested that many individuals moved around between a series of addresses, so that some addresses appeared multiple times in their records. This further questions the accuracy of being able to determine any one address as the main, permanent residence of at least some of these individuals. Ultimately, then, the decision was made that the best way to proceed was to use the current recorded home address and accept that this would introduce some inaccuracy. However, it was felt that this was less problematic than to use the arrest address and more consistent than trying to match to historic addresses.

For future research using data of this type, it is suggested that if address at time of offence (or address travelled from, which even with an accurately recorded 'home address' may not be the same) cannot be ascertained, then it would be worth investigating the impact of using multiple, weighted locations as the origin when calculating journey-to-crime distances. This was, however, outside the scope of the current study.

The original offender dataset was merged with the facility crime dataset, such that for each recorded crime with one or more associated offenders, new variables were created containing offender information (including age, sex, ethnicity and a unique ID). Where there were multiple offenders recorded against the same crime number initially, further new variables were created to capture this information. Therefore, any crime for which there was multiple offender information would have a set of variables for offender.1, offender.2, and so on. The maximum number of offenders for any single crime was seven. After cleaning and restructuring the provided data relating to current known address, this was then merged with the newly created dataset (using the unique offender ID). Ultimately this final dataset was restructured so that there was one row per offender journey. This meant that there would be multiple rows for each crime that had multiple offenders, or one row for each crime associated with only one offender. This also contained all facility crime not associated with an offender (one row per crime), but these were ultimately removed from the set for analysis. Overall, there were 8,768 crimes that could be associated with one or more offenders. This equates to 28.5% of facility crime.

Unlike the provided crime data, offender address data were not already geocoded. This therefore had to be done within ArcGIS (ESRI), after further address field variables were created so that these were in a format to allow them to be matched to (geocoded) address information that had been made available for the study area. Once X and Y co-ordinates had been assigned to each address that matched exactly, the unmatched addresses were categorised as either within the study area or outside the study area (geocoding could only be completed for addresses within the study area as there were no available data to match to outside of this area; the implications of this are considered further below). This was the first pass geocoding. Unmatched addresses that fell within the study area were then searched for and, where possible, geocoded manually, either to the exact address (very few) or to an approximate location. Even after doing this, there remained a number of addresses in the study area that were unmatched. This was usually because the information was not complete or was in a format such that it was not possible to ascertain with sufficient certainty where the premises were meant to be. Approximate locations were coded to the next nearest address. This was usually within a few 'house numbers' of the originally recorded premises (such as if there did not appear to be a 1001 High Street, so instead this would be geocoded to the nearest recorded address, which could be, say, 1010 High Street). If this was not possible it would be geocoded to the first address in the next block. If the address could not be geocoded to less than two one-hundred blocks away, it was left uncoded (an example of this might be an address recorded as Main Street, but the area only has North Main and South Main, therefore it would not be possible to know which of these was correct and the difference would be too large to be acceptable). This was classed as the second, and final, pass geocoding.

Overall, the first pass resulted in 7,179 addresses being geocoded to the exact address. A few addresses resulted in approximate geocoding at this stage and 411 addresses were considered to be within the study area but as yet uncoded. A further 2,703 addresses were categorised as outside the study area. Upon the second pass, these changed to 7,190 addresses exactly geocoded, 163 approximately coded and 196 addresses within the study area left uncoded (with 2,756 now recorded as outside the area). Exact and approximately geocoded addresses were used for the journey-to-crime analysis, which meant that of the ('known') journeys occurring in the study area, 97.4% were included in the analysis. However, this only equates to analysing 71.4% of *all* known crime journeys because those outside the study area are excluded. The implications of this are discussed in chapter 9, where these data are used.

6.4.1.2.3 Temporal data

Recoding was also necessary in order to analyse temporal patterns relating to (risky) facility crime. The resulting variables were to be used for analysing patterns over time, and of when crime occurred. Several new variables were constructed from the original full date and time variables, including the day of the week, month, year and financial year (amongst others).

Unfortunately, the data provided in the recorded crime datatset only related to the date/time *from* (that is to say, the earliest an offence could have taken place). For some offences this would be the actual time (or close to it) that the crime occurred, but for others there was likely to be a temporal range over which the offence could have occurred (because the actual time is often not known). This is particularly the case for offences where the victim was not present, such as burglary. A number of methods are available for dealing with offence time ranges, the most promising of which is a weighting approach based on aoristic analysis (Ratcliffe, 2000). However, none of these could be employed as the range was not known. There was no date/time *to* variable, so this could not be calculated.

Unlike most longitudinal research, there were few decisions to be made regarding analysing patterns over time. However, it was still necessary to select an interval size (Brandon Tuma, 2009). Monthly or quarterly differences across the ten years would have been interesting, but it was decided that the crime counts at this interval would be too small for patterns to be meaningful and that the number of intervals would be too large for the type of analysis that was to be employed. Therefore, for this study the interval was set as each financial year. With regards the accuracy of assigning each crime to a particular year, this choice of larger interval also ameliorated the impact of inaccurate time/date information. Although a small number of offences might have an occurrence possibility range this large, it is highly unlikely for most crimes. This means the likelihood of offences incorrectly attributed to a particular year was low, although greater towards both ends of the given period.

This assignation error is *much* more significant for the time of day analysis. In fact, it is substantial enough that the conclusions based on analysis of such data must be considered tentative. However, it was anticipated that different facility types and crime types would experience different patterns of offending at different times of the day, and that these would relate to the operating practices of facilities and the routine activities of those involved. Such patterns were deemed of sufficient importance that it was appropriate to forge ahead, albeit with these concerns explicitly acknowledged.

In determining the interval to use for time of day, thought was given to differences in operating hours, different types of business that might be carried out and different routine activities (and even offenders). Ultimately it was felt that more than two categories would be too much in terms of accuracy and that the evening category would be too small compared to the other two (which would have been day and night). Therefore, opportunity explanations for offending (routines and risk) were used to justify a focus on daylight versus darkness offences.

In order to produce the daylight/darkness variable, the *time from* variable was recoded using data from the US Naval Observatory website (from which yearly tables listing daily sunrise/set times can be obtained for a given city/town, see http://aa.usno.navy.mil/data/docs/RS_OneYear.php). The hours of darkness and daylight were available for every day of the ten-year period for the city being studied, therefore each offence was recoded as accurately as possible (given only *time from* could be used), rather than estimating daylight and darkness from more simplistic rules of thumb, or even only seasonally. Daylight savings time (DST) was also taken into consideration when using these data.

6.4.2 Errors and limitations

All research contains errors, missing data and inaccuracies, and this is particularly the case when relying on data compiled by someone other than the researcher (Bryman, 2015). Ideally these errors will be eliminated, if they cannot they will be reduced, and where possible they should also be quantified. There are particular issues and limitations relating to the data is this study and the ways it has been coded, that have already been considered. However, because of the type of data used and the specific method of selecting crime to include in the analysis by the facility it is recorded against (rather than using larger geographic units, or buffers) there are five errors that need to be specifically highlighted here.

6.4.2.1 Zero-crime facilities

The facilities and their crime counts used throughout this study are identified from police recorded crime data. Because of this, all the facilities present in the final dataset have experienced at least one crime within the ten-year period studied. In other words, there are no zero-crime facilities. Ideally, all premises in the area that met my definition of being a facility, would have been included. This was not possible given the data access and time constraints of the research, because for many types of facility, such information is not readily (or publicly) available. Even when it is, it is not necessarily in a format conducive to analysis or where the different datasets can be easily joined. The implications of this may be minimised in this research, as for much of the analysis, the whole ten-year period is included. It is less likely that a facility will have experienced no recorded crime over such a long period of time, but it is not impossible, and in fact research into zero-crime facilities themselves could be very revealing in much the same way that research into high crime premises is anticipated to be. This is a further endeavour that should be considered part of the risky facility research agenda.

The main limitation resulting from the likely exclusion of at least some zero-crime facilities (particularly when considering temporal comparisons, using one-year slices of data), is that the extent of apparent crime concentration will be less. In other words, the results are more conservative. I consider this to be more acceptable than over-estimating concentration. It is also recognised that this exclusion somewhat further reduces the generalisability of the findings. Further specific impacts of absent zero-crime facilities are considered where most relevant in the study chapters, following.

It is, of course, possible to address the issue of zero-crime exclusions, as the necessary data are obtainable (relatively easily) for certain types of facility. Therefore, it is strongly suggested that future research should use these types of premises as examples to test the impact of inclusion and exclusion of zero-crime premises on some of the findings reported here. In addition, it should be possible to also identify the effects over different time periods, to determine whether these are lessened when the data cover a longer time period (as we would expect to be the case). Examples of facilities for which full listings for a given area are more likely to be available include those premises holding alcohol service and/or public entertainment licenses and registered schools. Beyond this, White Pages or tax listing data ought to provide the necessary addresses by facility type, though this may not always be complete or up-to-date (directories) or may not be at sufficiently fine categorisation (tax listing). There also remain issues of contemporaneity, as historical data may be needed, when longer time periods are being analysed.

6.4.2.2 The coding error - excluded facilities

A number of premises in the dataset could not be coded as insufficient data were available to do so. This included missing addresses (with no facility category or an unspecified facility category recorded), incomplete addresses without facility information and addresses or facilities that could not be found, or reliably determined, through Internet searches. These were generally recorded or recoded as unknown nonresidential (123 records), unknown (572 records) and nonresidential not coded (1944). This equated to 2639 offences not coded, which might have involved a facility. However, as this could not be determined, these were all excluded from the analysis. This introduces some degree of error as some of the premises will be facilities. It is obviously not possible to quantify this error accurately without knowing what proportion were facilities. The maximum it could be, however (if all uncoded nonresidential offences occurred in facilities) is 8.59%.

A further error was introduced due to oversight. It became apparent after most of the analysis was complete that offences occurring in some facility types had been excluded from the primary dataset because the number of offences was so small. These had been coded as childcare (21 offences), hair and beauty (11), laundry (28) and library (10). These *should* have been included in the nonresidential facility total, but would have been excluded from the twelve analysed category types. As already noted, transport terminal (69) was also excluded. In part this was because it was difficult to determine whether all the offences falling within this category occurred at premises that were determined as 'facilities'. However, in hindsight it would have been useful to include this category and consider it further. Finally, retail strip (9) and retail mall (133) were excluded in error. Given that the nonspecific nature of these categories would have resulted in these being included in the retail other category, this has not affected the twelve facilities analysis. It does, however, introduce a further error into the total facility crime count and analyses based thereon. Assuming that all the offences discussed in this section occurred in facilities, the error can be calculated as 281 missing records, that is 0.91%. If transport terminal is excluded the error is 0.69%.

There are likely to be other errors in assigning addresses to facility types, in relation to the decisions made by reporting officers (Clarke & Eck, 2007), coding and extraction, but these cannot be quantified. Taking all of the above into consideration, the maximum quantifiable error in facility assignation is 9.50% missing.

6.4.2.3 Facility categories

Another issue that has to be considered is the process of categorisation and the impact of this on the results (thus conclusions). This is discussed more specifically where it is particularly relevant to the studies and in the final discussion chapter (10), because it is such a cross-cutting issue. In summary, the main concerns are the choice of categories (number and, relatedly, type) and the impact of officer decisions on coding. There were a number of flaws identified in the way premises had been assigned to categories and the categories themselves. It could be argued that looking at all facility crime instead reduces many of these, but it also masks possible different patterns displayed by different types of premises.

Some of these categorisation concerns could also be addressed through more accurate coding, but this could only be done manually (searching street addresses), which was not practical for the entire dataset.

6.4.2.4 The aggregation/assignation error

Having recoded the data and commenced the analysis, it became apparent that there were potential errors relating to the method of aggregation used. This arose from the fact that the same address was in some instances coded as different facility types. This was either because of inconsistencies in coding (reporting officers defined the premises differently, for example one thinking Walmart is a supermarket, the other thinking it is a retail premises) or because there were legitimately multiple premises sharing the same address (as was often the case with shopping malls/plazas, but even sometimes small strip malls or, say, a bank outlet in the same building as a supermarket). This meant that if the data were separated by facility type first, then aggregated on address, the total number of premises (and offences) within that category was different to if the data were first aggregated on address and then separated into different facility types. Using the latter approach, the aggregation error would be such that all facilities at that address would be coded as the same category (first, last, most frequent?), so that one would be over-represented and the other under-represented (for multiple premises). Regardless of the reason for the error, it could only be reduced through manual coding of all affected addresses, which was not feasible.

I settled on separating by facility type first. This was based on the assumption that the reporting officer visiting the premises ought to have a good idea of what type of facility it was, so it was better to accept some degree of error in this, than to collapse lots of different premises at the same address into just one crime count. Having completed the analysis, selected risky facilities and checked these addresses (as part of the comparison with the qualitative data) it was apparent that there were still addresses with multiple premises represented (such as a large shopping mall) because those of the same facility type had been aggregated together anyway. This overlaps with the categorisation concern raised above, and discussed more in chapter 10.

6.4.2.5 Crime counts versus rates

Crime counts have been used throughout these research, despite the fact that rates (using a suitable denominator) may better capture premises that are perceived to be more problematic (Sidebottom & Bowers, 2010). As has been noted elsewhere, this decision was taken due to the difficulty of identifying suitable denominators and obtaining the required data, across all the categories of facility included within the study.

In some ways, the decision to use counts to identify risky facilities is theoretically and empirically defensible, as size, turnover, population, usage, capacity, and so forth might all be *explanations* for the emergence of a facility as ‘risky’, rather than something that needs to be controlled for. Whether the aim is to test the concept of the risky facility distribution, or to identify those premises disproportionately experiencing and contributing to the incidence of crime within a city, then the rate of crime becomes inconsequential.

However, if the aim is to consider whether some premises contribute more crime than they ‘ought’ (given their size, turnover, the list is repeated) and/or to identify manipulable features or circumstances that contribute to this, then calculating and comparing rates, using suitable denominators, becomes far more important.

In this study, I argue that my endeavours are the former, therefore the use of counts is defensible. That said, consideration of risky facility patterns also ought to be carried out (and compared) using rate measures, and this is recommended as part of the future research agenda. This is discussed further in chapter 10.

6.4.3 Approach to quantitative analysis

The data were subject to various analyses in order to address the research questions. Sometimes, further data preparation was required and this is presented as part of each of the relevant studies. The specific forms of analysis and choice of tests is also detailed at the beginning of each study. However, there are certain tests that are introduced here as of particular importance, and to allow the study chapters to focus more on justification and application. These are the Kolmogorov-Smirnov (K-S) test (used for two purposes) and the Mann Whitney Wilcoxon (MWW) test.

It should also be noted that in some cases *exact* (as opposed to asymptotic) significances are (explicitly) displayed. This follows convention for small sample sizes (Hinton, 2014).

6.4.3.1 Choice of statistical tests

The Kolmogorov-Smirnov (K-S) test compares the probability distribution, either between one sample and a reference distribution (goodness of fit (Massey, 1951)) or between two independent samples. The K-S statistic is a measure of the distance between the cumulative distribution function (CDF) of the reference sample and the empirical distribution function (ECDF) of the study sample, or between the ECDFs of two samples (Hinton, 2014; Massey, 1951). In the current research, it is used both ways. Firstly, it is used as a test of normality in all cases where the sample size is greater than 50 (see discussion in chapter 9). Here the K-S test is effectively being used to determine goodness of fit, by carrying out a comparison between the sample data (standardised) and a standard normal distribution. When the result is statistically significant (here using an alpha of 0.05), the null hypothesis – that the sample is drawn from the reference distribution – can be rejected, which is interpreted as meaning that the sample is *not* normally distributed (Corder & Foreman, 2014). All tests of normality are subject to limitations and criticism, but the K-S test is commonly used, so in this research I adopt the 'norm' of using this except for samples of less than 50 (see chapter 9).

Tests of normality are carried out because the assumptions associated with parametric tests (such as the independent samples *t-test* and Pearson's correlation coefficient) include that the data are normally distributed. Therefore, the data are tested and, as will be seen, because in most cases the null hypothesis is rejected, nonparametric tests, which make no such assumptions (Corder & Foreman, 2014; Kiess, 1996), are used throughout.

The K-S test is used as a further goodness of fit test as part of the work assessing whether risky facilities follow a power-law distribution. This is discussed in chapter 7.

Secondly, the K-S test has been used as a 'test of difference', for comparing the distributions between two independent samples, notably in the studies of journeys-to-crime (chapter 9). In many cases, a Mann-Whitney-Wilcoxon test (also known as the Mann Whitney U test or Wilcoxon rank-sum test – as distinguished from the Wilcoxon signed-rank test) has (also) been used, as the 'standard' nonparametric equivalent to the independent samples t-test (Corder & Foreman, 2014; Fay & Proschan, 2010). In simple terms, this test determines whether a randomly selected value from one sample would be equally as likely to be higher or lower than a randomly selected value from the other sample (this is the null hypothesis). As such, when the null hypothesis is rejected, we can conclude that the distributions are not equal (Mann & Whitney, 1947; Wilcoxon, 1945). When certain assumptions are applied (basically that the populations from which the samples are drawn are continuous, and that the two samples, have similar distributions) the MWW can be seen as being a test of difference in the medians of the two samples (a shift interpretation). However, the test cannot be used to conclude this if the two samples possess very different distributions (Hart, 2001). When this is the case, the alternative hypothesis should be that one distribution is stochastically greater than the other (cf. H0 that the distributions are equal). Additionally, the K-S test is also used to compare the two sample distributions. As described above, this is a measure of the distance between the ECDFs of the two samples; thus is specifically a test of whether the two samples come from the same distribution (though not what this distribution is).

Thus in this research, predominantly the MWW test is used to draw a conclusion about whether there is a statistically significant difference either in the medians of two samples or that one of their distributions is stochastically greater than the other. The K-S test is also used to consider if there is a statistically significant difference in the distributions of the two samples, which is both useful for interpreting the results of the MWW test, and also of interest in its own right.

Throughout this research, an alpha of 0.05 has been used as the level for statistical significance, unless otherwise stated. In results tables, to aid the reader, the following formatting is used to highlight different levels of significance (<): *0.1*; **0.05**; ***0.01***; <u>***0.001***</u>.

6.4.3.2 Mapping and geographic data

Facility crime, offender residence and journey-to-crime data were mapped and visually inspected. Limited examples of hotspotting were also carried out. All the data preparation and the explanation of the tools used and analysis carried out, are presented in chapter 9.

6.4.4 Qualitative data

As stated, the qualitative data consisted of interviews and observations that were undertaken during a field trip to the study area. The data collection methods, and the analysis to which they were subjected, are now presented.

6.4.4.1 Field trip and introduction

The field work took place in May/June 2011. For the purposes of presenting relevant information in the thesis, only issues directly related to data gathering and other methodological considerations are presented here. Within the relevant sections, I include some reflexive account as I believe this to be an important element of qualitative research, however I have not reflectively considered issues relating to the broader research trip and data collection process, due to the constraints of the thesis and the fact this is less necessary with the research not being a 'true' ethnography. In the end, the data gathered through structured observation, ethnography ('living in the city'), photography, and informal conversation have only been used supplementally. The analysis used on these data (thick description, as briefly described below), allowed for additional confirmation and illustration of the main findings (in keeping with the mixed methods approach), but I was not able to incorporate this as fully as I might have wished because of the constraints of the thesis requirements. I intend to explore these data, and the benefits such methods might bring, further in the future.

The specific methods employed for both the interview and observation elements, and the strengths and limitations of these, are now considered in further detail.

6.4.4.2 Obtaining the interview data

In this section, I briefly set out the process of collecting interview data through the use of semi-structured interviews and map annotation. I consider the additional benefits this has brought to the overall research project, as well as some of the limitations of this approach generally, and more specifically with reference to the particular circumstances. Copies of the consent form, information sheet, interview schedule, and annotated maps that are referred to can be found in Appendix 1.

Interviews were carried out with six serving police officers to ascertain their perceptions of risky facilities in the area. This also took the research into a more innovative area, particularly through the use of map annotation concurrent with the interview discussion, which had some similarities with Ratcliffe & McCullagh (2001) but was constructed and analysed differently (qualitatively). The sampling procedure has already been set out above. Officers were recruited through word-of-mouth (via my contact) whilst I was on the premises and ready to commence, and they were carried out over two days. The interviews ranged in length from 11 minutes 53 seconds to 40 minutes 11 seconds, averaging around 20 to 25 minutes. However, this is the length of the audio recording and the whole process took a little longer as some of the recordings were paused whilst the respondent was annotating the map (this was not always the case and when it was, notes were taken if they commented).

The approach taken can be categorised as semi-structured, as the initial instructions regarding annotating the map and what I hoped they would tell me about, as well as three main issues/questions used as prompts at relevant times, covered the same content (though the wording was slightly different given what had been covered previously or whether I felt the respondent needed me to be a bit clearer, such as providing an example, reassuring them they were 'doing it right', and so forth). This was done to introduce some degree of consistency in the data that were gathered, the way the respondents annotated the map (such that it would be comparable with the others) and also to limit the differential impact of interview effects (Arksey & Knight, 1999; Bryman, 2015). That said, beyond these prompts and instructions, the interviews flowed quite freely. I intervened to bring respondents back-on-track if they began to talk about issues that were not at all relevant to the research project, but otherwise I let them define problematic premises, places and hotspots as they saw fit, and the focus they took then became part of the thematic analysis.

The respondents were also asked to 'draw on a map' locations where there were problematic premises, and then (after these had been discussed) areas of high offender residence. Most respondents annotated the map first and then talked through the places they had highlighted, adding some more marks when the discussion prompted them to remember somewhere they had missed. Some annotated as they talked. Most circled quite large areas, even when they talked about individual premises and they also tended to include areas of high offender concentration as part of their consideration of 'problematic places', even though they had not (at that time) been asked to do so. The extent of annotation varied quite substantially across the respondents, but many of the same places, especially with regards offender residences, were identified. These are discussed more in chapters 8 and 9, and replicated versions of the maps can be found in Appendix 5.

Overall, the interview element of the research was successful, providing a wealth of data and, as will be seen from discussion of the findings, it raised a number of interesting and important issues. The strengths and weaknesses of qualitative research generally have already been rehearsed, but with respect to interviews specifically, the strengths here were being able to compare two different forms of data (but arguably collected from those with a similar perspective) and to be able to discuss issues in greater depth so that the research findings *made more real world sense*. On the other hand, it is recognised that the sample was small and that a greater diversity of experiences and roles would have been beneficial. It is not possible to generalise the findings to the police generally, and maybe not even to other officers in the PD. Additionally, findings might have been affected by my behaviour in the interviews and my interpretation could have been biased by my knowledge of the crime patterns in the study area (although these were only based on preliminary analysis) and the different cultural perspective I brought (Arksey & Knight, 1999; Bryman, 2015; Wilson & Sapsford, 2006).

6.4.4.3 Carrying out the observations

In this section I report on the approach taken to observations, which involved the use of structured observation, unstructured observation and informal observation ("living in the city"). Observations are only considered briefly due to the constraints of the thesis, and because the findings from observation have become supplemental (rather than integral) to the main analysis.

The observations carried out included a structured observation (Foster, 2006) of the external features and environment of some types of facility, and certain addresses. These were selected from an initial, basic analysis of the facility crime data, and also from the suggestions of an accompanying police officer. I undertook more informal observations on my own, making notes on the locations and general 'feel' of the places I encountered, which included revisiting some of those that had already been observed, as well as other premises (particularly retail premises in plazas, malls and (uncovered) shopping centres though these were less systematic (Foster, 2006). I also made brief field notes relating to informal discussion with police personnel, my experience and impressions of being in the city generally (for a period of 10 days), and I took photographs of some of the premises visited, where I was able. These were supplemented by photos obtained from Google Maps/Streetview. These data were subjected to qualitative analysis (and reflection), particularly in the form of thick description (below).

When taking photographs, I had intended to use visual analysis techniques, thus treating these as another form of 'data', however given the scope and constraints on the thesis, their use has been reduced and they are now more used as a matter of record (part of my field study notes) and as illustration to help expose the reader to the context of the research (Emmison & Smith, 2000).

As with the other qualitative research, the findings of the observations helped situate the research, and the findings from the other methods, in the 'real world'. My understanding of the patterns produced through the quantitative analysis and, crucially, the ways I was able to interpret *why* these patterns might exist (and associated implications) would not have been possible without the field trip and time spent both formally and informally observing the city and its facilities. On the other hand, this is the most subjective element of the study, and observer bias is a real threat to the reliability of these findings, due to the my cultural knowledge and the influence of my theoretical standpoint on what I sought ought, noticed and how I interpreted this (Foster, 2006).

6.4.4.4 Analysis of qualitative data

6.4.4.4.1 Content analysis

Content analysis is a process used to reduce text or other qualitative data. This can be reduction into themes, but more usually it involves reducing into numbers so that qualitative data can be quantitatively analysed (Franzosi, 2009). A quantitative content analysis was carried out of the interview data as described in chapter 8.

6.4.4.4.2 Thematic analysis

Thematic analysis is another method used to assist the researcher in finding patterns in large amounts of qualitative data (Franzosi, 2009; Bryman, 2015). It is an iterative process, by which the data are read through a number of times and commonalities and differences are, eventually, turned into codes or themes. Although the use of coding varies, I took a more qualitative approach and used the themes and sub-themes as a form of analytical and reporting framework, to conceptualise the key ideas and variations in perception that were displayed (as opposed to counting them, for example). Thematic analysis was applied to the interview data and annotated maps, and is discussed more in chapters 8 and 9. The themes and sub-themes that were developed from the data are shown, below, in table 2. Not every sub-theme has been reported in this study, due to length constraints, but all are included here due to the overlapping nature of some of the issues (albeit separated because they were talked about in relation to different things).

Table 2: Thematic analysis themes and sub-themes

Level 1	Level 2	Level 3
Location types	Residential areas	Housing 'projects'/social housing 'Rich' areas
	Retail/commercial areas	
	Size of stores	Small/local community stores Big box stores
	Entertainment and leisure areas	
	Streets, parks and hang outs	
	Schools	
	Industrial areas and construction sites	
Crime types	Drug-related	
	Gang-related	
	Other violence and disorder	
	Acquisitive crime	
Facility issues	Small strip malls/multiple facilities	
	Proximate to offenders	Near offender residences Near gang areas Problematic individuals in vicinity
	Nodes, paths and edges	
	Drug dealing/prostitution in area	
	Place management	Managers complicit Poor or vulnerable management
	Targets/opportunities	
Offender types	'Criminals'	Gangs Drug dealers
	Time on their hands	Kids The 'feckless'
	Culture	Hispanics
Offender sources and routes	Locals	Nearby Corridors Housing 'projects'/social housing
	From south to north	
	Out of town	
Offending explanations	Drug-deals and systemic crime	
	Gang or group-related	General gang activity Gang or school rivalries
	Hanging out	Congregating

		Street drinking Unattended children
	Opportunities and good targets	Everyday activity Local and easily accessible Entertainment clusters 'Professional' offenders
Perceptions, experiences and processes	Reading map/locating places on map	
	Focus	Knowledge areas and expertise Reference to own experience vs hearsay Favoured explanations Offender locations vs offence locations
	Consistency and history	
	Annotation style	Encompassing vs specific

6.4.4.4.3 Thick description

Some forms of qualitative analysis can be difficult to define, but the approach I took to analysing the field data (observations, photographs and informal conversations) can be best thought of as thick description (Bryman, 2015). Data were recorded (variably) in a field diary and structured observation took a similar form throughout, from a list of self-prompts (such as detritus and graffiti in the area, security features, accessibility and permeability, and so forth, although these were not carried out as site surveys). Thick description involves taking the observations, notes, things seen and impressions, and describing these in a more detailed way so as to interpret the data for the reader (Thomas, 2017). Quality thick description is evocative, and may be thought of as a form of 'story-telling', based on what the researcher saw and felt, rather than what they factually report as an 'outsider' or 'bystander'. Clearly, this method of analysis carried with it the greatest strengths and weaknesses assigned to qualitative approaches, as already set out. It has been invaluable in supporting the most intepretivist elements of the study, which are also those that are most under-represented in this field.

6.5 Summary, conclusions and methodological limitations

In this chapter I have set out the elements of the research design applied to the current study. I have articulated my worldview and research orientation, as well as setting out and justifying the approach taken and the design frame employed. I have argued that mixed methods research is consistent with my (realist) pragmatist stance, as well as being a somewhat under-used approach in the field of environmental, specifically place-based, criminology. In doing this I set out why I decided on a mixed methods approach, and how I intended to use it to strengthen, cross-check, confirm and clarify the findings produced from quantitative statistical analysis of police recorded crime (and offender) data, as well as from content and thematic analysis of semi-structured interviews with serving police officers accompanied by a task of annotating a map (also included in the analysis). This would be supplemented by the findings from (a reflective) thick description of observations, field notes and photographs. I also introduced geographic data and how this would be used to produce crime and facility maps, which would be subject to visual analysis, as well as being used to help illustrate some of the findings in chapter 9.

In doing this, I recognised a number of methodological, data collection and analytical limitations. Most important of these were those that came from inaccuracies and discrepancies in coding by the police, alongside issues resulting from the methods used to categorise and aggregate facilities and the crime associated with them. Where directly relevant, some of these limitations will be raised again, below. There were also limitations, common in this field, regarding the use of police recorded (crime) data and what this actually represents, as well as concerns regarding how much the interview data could be said to be representative of other police personnel. In addition, I recognised that although the methods and type of data were mixed, the main sources were similar, so that only the policing perspective was actually represented. In a way this gives my observations and impressions, even if somewhat limited in presentation here, a greater importance (having had policing experience myself, this may be somewhat reduced, but being a cultural outsider looking in should increase their worth).

In this chapter I have provided a relatively detailed consideration of the philosophical and design issues related to the current research, but further information regarding data preparation and analysis is presented in each of the three study chapters, as relevant. I have presented here, though, some overarching issues, and tests used in multiple places. Given the range of methods, data sources and forms of analysis carried out in this research project, these are summarised, for reference, in table 3, overleaf.

Table 3: Research questions, data and analysis

(Continued overleaf)

Research question	Method and data	Analyses	Study/chapter
RQ1: How 'ubiquitous' is the concept of risky facilities?			
RQ1.1: Is crime in facilities concentrated in some premises more than others? Specifically do different classes of facilities in a given location show evidence on unequal distribution of crime, which may be considered consistent with the concept of 'risky facilities'?	*Quantitative* Statistical analyses and visual inspection of charts Police recorded crime data	Percentage contribution of crime (25% and 50% as per Weisburd, 2015; 80% as per 80-20 rule) Gini coefficients Ranked crime count charts (J-curves)	Study 1/chapter 7 *Crime concentrations in facilities*
RQ1.2: Is there variability in the concentration of crimes within facilities by: (a) facility type; (b) crime type; (c) time of offence?	*Quantitative* Statistical analyses and visual inspection of charts Police recorded crime data	As RQ1.1 but comparing by facility type; acquisitive vs violent offences; daylight vs darkness offences	Study 1/chapter 7 *Crime concentrations in facilities*
RQ2: Are risky facilities stable?			
RQ2.1: Is the distribution pattern persistent?	*Quantitative* Statistical analyses Police recorded crime data	As RQ1.1 but comparing over time (10 year period)	Study 1/chapter 7 *Crime concentrations in facilities*
RQ2.2: Within each class of facilities, is the ordering of premises consistent?	*Quantitative* Statistical analyses Police recorded crime data	Inspection of ranking tables (top facilities) over time Spearman's rho	Study 1/chapter 7 *Crime concentrations in facilities*
RQ3: How is the concentration of crime within facilities distributed?			
RQ3.1: For each class of facility, does the concentration of crime best fit a proposed power-law distribution?	*Quantitative* Statistical analyses Police recorded crime data	Statistical analysis of goodness of fit comparing power-law and log-normal distribution	Study 1/chapter 7 *Crime concentrations in facilities*
RQ4: Can risky facilities be defined empirically?			
RQ4.1: What quantitative methods could be used to empirically define risky facilities and what are the implications of using these?	*Quantitative* Ranking Police recorded crime data	Application of different methods Tabulation of methods with cut-off points and information about what is/is not identified as a risky facility	Study 2/chapter 8 *Identifying risky facilities*
RQ4.2: Which quantitative method is most appropriate?	*Discursive*	Comparison of different methods from an empirical, conceptual and practical perspective	Study 2/chapter 8 *Identifying risky facilities*
RQ4.3: Where do operational police officers perceive risky facilities to occur and how does this compare to quantitative identification from recorded crime data?	*Qualitative* Interviews and annotated maps Observations	Content analysis; thematic analysis; thick description Visual comparison of maps	Study 2/chapter 8 *Identifying risky facilities*

(Table 3 continued)

Research question	Method and data	Analyses	Study/chapter
RQ5: Where are risky facilities located and what types of crime journeys are made to them?			
RQ5.1: How are risky facilities distributed across the study area?	*Qualitative* Interviews and annotated maps Observations *Quantitative* Mapping and visual inspection of police recorded crime data	Thematic analysis; visual inspection; comparison; thick description of field study notes and photographs	Study 3/chapter 9 *Locations of risky facilities*
RQ5.2: What distances of journeys to crime are associated with risky facilities (compared to non-risk facilities)?	*Quantitative* Mapping and statistical analysis Police recorded crime and police recorded offender data	Calculation of distance of journey to crime Statistical analysis (tests of difference)	Study 3/chapter 9 *Locations of risky facilities*
RQ5.3: Do distances of crime journeys differ by (a) facility type; (b) crime type; (c) time of offence?	*Quantitative* Mapping and statistical analysis Police recorded crime and police recorded offender data	As RQ5.1 but comparing by facility type; acquisitive vs violent offences; daylight vs darkness offences	Study 3/chapter 9 *Locations of risky facilities*
RQ5.4: What are operational police officers' perceptions of journeys to crime at risky facilities?	*Qualitative* Interviews and annotated maps Observations	Content analysis and thematic analysis of initerviews, annotated maps and thick description of field study notes	Study 3/chapter 9 *Locations of risky facilities*

Chapter 7: Study 1: The ubiquity and manifestation of risky facilities

7.1 Introduction

This chapter specifically considers the following aims, and associated research questions:

Aim 1: to consider whether crime concentrates in facilities, and how this concentration manifests.

RQ1: How 'ubiquitous' is the concept of risky facilities?
RQ1.1: Is crime in facilities concentrated in some premises more than others? More specifically, do different classes of facilities in a given location show evidence of unequal distribution of crime, which may be considered consistent with the concept of 'risky facilities'?
RQ1.2: Is there variability in the concentration of crimes within facilities by: (a) facility type; (b) crime type; (c) time of offence?

RQ2: Are risky facilities stable?
RQ2.1: Is the distribution pattern persistent?
RQ2.2: Within each class of facilities, is the ordering of premises consistent?

RQ3: How is the concentration of crime within facilities distributed?
RQ3.1: For each class of facility, does the concentration of crime best fit a proposed power-law distribution?

This chapter proceeds with a discussion of how these research questions were addressed, through the use of various forms of quantitative analysis of the recorded crime dataset. Following this, the results of these analyses are presented. The distribution of crime and the extent to which it is concentrated within premises is considered with reference to the 80-20 rule, the law of crime concentration at places, and the Gini co-efficient. These distributions are presented for all facility crime, and for crime by facility type. Whether this distribution is one that is consistent with the concept of 'risky facilities' is then considered, with reference to the 'J-curve'; again for all facility crime and for each class of facility.

The variability of concentration of crime across the different facility types is then presented, as is a comparison of facility crime concentration for acquisitive versus violent crime and daylight versus darkness offending. Next, persistency and consistency are explored through a consideration of facility crime distribution over time. Finally, the results of the goodness of fit tests for a power-law distribution are presented.

These results are then discussed in light of the existing literature and what this suggests regarding the ubiquity of the risky facility concept, in a given location, and the way this concentration manifests. The results are also briefly considered with reference to the remaining research questions and the aims of the research overall. A more detailed and holistic discussion of all the findings is presented in chapter 10.

7.2 Data preparation and analysis

In order to answer the research questions set out above, the recorded crime data were cleaned and coded as set out in chapter 6. The analyses set out below were carried out on all recorded facility crime, as well as for the 13 categories selected for comparative consideration. The analysis methods employed are now explained and justified.

7.2.1 Concentration

In order to ascertain whether crime in facilities was concentrated in particular premises, three approaches were taken, drawing on previous research. Firstly, the percentage of premises contributing 25% and 50% of recorded crime was calculated, following Weisburd (2015). Thus as well as exploring the concentration of crime in facilities, this is a further test of the proposed law of crime concentration at places.

For completeness, the proportion of premises contributing 80% of recorded crime was also calculated. This was inspired by the 80-20 rule of thumb though it is recognised there are no claims that this specific ratio should be expected, rather that this indicates the nature of the pattern.

For both approaches involving calculating the proportion of premises that contribute a certain percentage of crime, a decision had to be made as to when to determine this contribution had been met. The same approach was applied in all cases and this was to calculate the proportion of premises from the number required to meet or surpass (to two decimal places) the target percentage of offences. This means that if 10 premises contributed 24.55% and 11 premises contributed 25.95% of crime, then 11 would be used to calculate the proportion of premises. When very few premises (one or two) accounted for a very large proportion of crime, this could mean that the cut-off was significantly over the 25% target (e.g. for *educational* the 25% category was actually 29% of offences and the 50% category was actually 56%). This was also more likely to be the case (though not inevitable) for categories with fewer premises. This suggests, therefore, that findings could be somewhat inconsistent, so that when comparing across studies the use of a bandwidth may capture different proportions of crime (in my study, up to 32% of offences for the 25% category, contributed by a single address). The Gini coefficient, below, is not affected by this problem. In addition, I test a number of approaches to identifying risky facilities, which can be used to compare across studies, in chapter 8.

Both of the approaches outlined so far allow a consideration of the degree to which a small proportion of a given unit of analysis (in this case premises) contributes a disproportionate amount of crime. The 80-20 rule would expect about 80% of crime to result from about 20% of premises, whilst Weisburd's law predicts that a universal 'narrow bandwidth' across studies will account for the percentage of premises contributing 25% or 50% of crime, and that given the existing literature on spatial crime concentrations, the proportion of premises will be small. Thus, as shown in earlier chapters, these approaches and the patterns they predict are consistent with the concept of risky facilities, when premises are used as the unit of analysis.

Finally with regards to concentration, the Gini coefficient was also calculated. The benefits of using the Gini coefficient for studying crime concentrations, especially when carrying out comparison, have already been set out in chapter four. It was also noted, however, that this approach does not identify the nature of the distribution, nor of where the concentration lies. This is the reason why using a combination of methods is important, and is the approach taken within this study.

The Gini coefficient (G) was calculated geometrically (Xu, 2004) using the formula [1]:

$$G = 1 - \sum_{k=1}^{n} (X_k - X_{k-1})(Y_k + Y_{k-1i})$$

Where X_k is the cumulative proportion of the number of facilities, for k = 1 to n and Y_k is the cumulative proportion of the recorded offences, for k = 1 to n. Y_k is indexed in increasing order of size (i.e. $Y_k > Y_k - 1$)

There is no 'cut off' point identified in the literature for concluding that a Gini is 'high' or that offences are ‘skewed’, though as the main aim in using such a measure is comparison, this is not considered problematic (Fox & Tracy, 1988; Piquero, 2000). Further, it has been noted that throughout this research, zero crime locations have been excluded. Doing this has the effect of reducing the size of the Gini coefficient (Deltas, 2003; Fox & Tracy, 1988), such that a smaller G can still be seen as evidence of unequal distribution. Gini coefficients of around 0.4 (and sometimes lower) have been presented as demonstrating this in previous literature. However, it is also acknowledged that using a ten-year period of study, there may very well be few zero crime premises, such that the effect of excluding these when considering all facilities may be much reduced.

It should also be noted that sample size needs to be considered when calculating the Gini coefficient (see the results section for sample sizes used in this study). Very small samples can have the effect of reducing this measure, so that distributions may seem less unequal than they actually are, that is to say there is a known small sample bias (Deltas, 2003). From consideration of the results (below) it became apparent that there was no obvious relationship between the number of premises in a given category of facility and the Gini coefficient of that facility relative to the others. However, this does not preclude the possibility that sample size has had an effect on the calculated Gini coefficients, thus the ordering of categories in the resultant index. The literature does not appear to establish a definition of ‘small sample’ nor to provide a recognised method for determining the validity of the measure when used on ‘small’ samples, just that they result in a downward bias. The extent to which this may be problematic for the current data is unclear, but as sample sizes can be in the tens, this bias may be present. Numerous mathematical corrections for very small samples have been proposed in the literature, therefore additional adjusted Gini coefficient estimates were calculated for all of the analyses based on this measure, by using the simple first order correction proposed by Deltas (2003) [2]:

$$G^{adj} = \frac{n}{n-1} G$$

Where n is the number of facilities in the sample, and G is the original Gini coefficient as calculated in equation [1] above.

Whilst this does not account for distribution specific bias, it is generally determined to perform well (Deltas, 2003; Van Ourti & Clarke, 2011) and is considered sufficient for my purposes here. The adjusted Gini coefficients are therefore also included in the results below. However, whilst this reduces any size effects, bias is not removed completely, thus results relating to categories with very small numbers of facilities should still be interpreted cautiously.

To address the research questions, these outputs are presented for all facility crime and then for each facility type to determine if there is an unequal distribution of crime across premises in different classes of facilities.

7.2.2 The risky facility pattern

In order to assess whether this distribution was consistent with the concept of ‘risky facilities’, the extent and pattern of the concentration was considered. There are a number of ways to present data that is believed to follow the particular pattern associated with the concept of ‘risky facilities’. The most simple is to produce is a J-curve chart. This involves plotting ranked crime counts. These are produced by aggregating the data by address, ranking this with the highest count at position 1 and then plotting these results with premises (individual addresses) on the x-axis and crime count on the y-axis. If consistent with the anticipated pattern, the resulting shape produced by the bars will look like a reclining letter ‘J’. Sometimes these charts are presented as point or line charts, however, as the x-axis is categorical it is more appropriate to use a bar chart.

Given that this pattern is possibly a discrete Pareto distribution (that is to say, it follows a power-law, as previously defined), it is worth noting here, that it is conventional to represent such distributions using the ranked cumulative function, rather than count, though the pattern produced is very similar. Additionally, as also noted, it is not uncommon to see such distributions presented using log (count) vs log (rank) as for a power-law distribution this tends to a straight line. Therefore, as part of the analysis, ranked cumulative distribution charts were also produced, but as both methods are considered suitable for comparing the *shape* of the distribution, the J-curve charts were selected for presentation in this chapter (as being more consistent with other criminological literature in this field).

Consideration of the J-curve charts was again completed for all facility crime and for each class of facility. Some of the conclusions presented here, regarding whether a J-curve is displayed, must be treated with caution. Although the J-curve is sometimes referred to as a Pareto or power-law distribution, no claims have been made that such a distribution *must* be present. Therefore, there is no indication in the literature of when we can (or cannot) conclude that something follows a ‘risky facility distribution’; only that we must have a small number of premises that experience a large proportion of crime and many that experience much less, which will likely produce a J-type curve when plotted using the ranked crime count approach.

Visual inspection of such plots, as presented here, will inevitably involve a (potentially large) degree of subjectivity in determining if the curve really looks like a reclining J. In addition, the smoothness of this type of plot - which may influence the decision over whether it is a 'J' - will be affected by the number of data points (both premises and counts); so for smaller sample sizes, types of facility where there are fewer premises, and facility types (or locations) with lower amounts of crime, there is a greater likelihood that a J-curve will not be as apparent, or may be open to different conclusions being drawn. In these cases, my conclusions are, therefore, explicitly more tenuous.

7.2.3 Comparison

In order to further consider how ubiquitous the risky facility pattern is, the variability of the concentration of crimes was considered, comparing the different outputs produced across each facility type, between violent and acquisitive crime, and between daylight and darkness offending. The bandwidths (Weisburd, 2015) and the Gini coefficients are the main source of comparison, although for facility type, the 80% cut-off point is also used both for completeness and because of the interesting findings when comparing across these different methods of analysis.

For facility type, as has been discussed in the general methodology chapter (6), thirteen categories were selected as suitable for further analysis. One of these is *retail other*, which is excluded from some of the analysis for being too broad and disparate, but is included for completeness in most of this comparison. It must be remembered that the category of *all facilities*, is in fact made up of all premises, and associated offences, determined to be a facility and is therefore larger than the sum of the named thirteen categories.

7.2.4 Persistency and consistency

The persistency, or stability, of the above described crime concentrations was also investigated. Here I am referring to a persistent inequality of distribution over a number of years (a significant proportion of the ten years included in the study) that involves the concentration of crime in a small number of premises (risky facilities). This was considered in two ways. Firstly, through calculating Gini coefficients for each facility type by financial year (April-March) and secondly by visually comparing J-curves for each facility type by financial year (that is to say, ten plots were produced and inspected for all facility crime and per class). It must be noted that as the number of offences per year was often very small, these findings should be treated with particular caution.

To test for consistency, the ordering of the premises for each class of facility was considered. This was done by calculating Spearman's rank order correlations for each pair of financial years covered by the data. A non-parametric, rank-based test was used to test for correlations (Kiess, 1996), as the actual number of offences experienced each year was not important (not least because these could be affected by all sorts of intervention, resource-allocation and socio-political circumstances, resulting in numerous confounding variables)[6]. What was important, was whether the position (rank) of each premises tended to change or not (Hardy, 2009). If there was a significant positive correlation, then it could be concluded that the more risky (highest ranked) facilities tended to be risky across both pairs of years, whilst the least risky tended to remain low ranked across both pairs of years. Such a finding would suggest there is at least some consistency of 'riskiness'. In addition, for two example facility types (*educational* and *supermarket*) the top ranked addresses across each of the ten years were also inspected for consistency.

7.2.5 Nature of distribution

Testing for power-law (Pareto) distributions is inherently difficult because of the nature of these distributions, particularly the "large fluctuations that occur in the tail of the distribution – the part of the distribution representing large but rare events – and by the difficulty of identifying the range over which power-law behavior holds." (Clauset et al., 2009: 661). In the current study, this would be the fluctuation in crime counts for risky facilities.

Therefore, many fitting methods are considered inappropriate (Clauset et al., 2009; Gillespie, 2015). For example, the relatively simple technique of identifying power-law distributions as a straight line on a log-log plot is discounted (Clauset et al, 2009; Gillespie, 2015; Goldstein, Morris & Yen, 2004; Stumpf & Porter, 2012) in favour of more rigorous techniques. To this end Goldstein et al. (2004) and Clauset et al. (2009) propose using a maximum-likelihood estimation method, with goodness of fit tested using the Kolmogorov-Smirnov statistic and likelihood ratios. The Kolmogorov-Smirnov test, which is also used in other parts of the study, has already been discussed in chapter 6.

The Clauset et al. (2009) method has been used as the basis for a package (Gillespie, 2013;

[6] If this had been of interest, the percentage of crime contributed by each premises, rather than the crime count, at least, would have had to be used in an attempt to control for this.

2016a) in the R statistical computing environment, which is available to freely download (The R Foundation, n.d.). The package is written for both continuous and discrete power-law distributions. As the data in this study are discrete (and counts are not large), this approach is the focus of the remaining discussion.

The equations and derivations for this can all be found in Clauset et al. (2009) and Gillespie (2013).

The distribution diverges at zero, therefore for this type of distribution there is a lower boundary or ‘cut off’ point after which the power-law distribution does not apply (x_{min}). In other words, it is only for values of x greater than this minimum that a power-law distribution can be calculated and is displayed; this equates to the end of the ‘tail’ of the distribution.

To assess conformity of a sample of data to the power-law distribution, Clauset et al’s (2009) proposed method is to estimate x_{min} and a scaling constant, α, for the model, to calculate the goodness of fit between the data and the power-law, and to then compare the power-law with alternative distributions (hypotheses) using a likelihood ratio test. The authors acknowledge this final stage could be determined using alternative model comparison, but their proposed approach is followed for this study.

The estimation of alpha is carried out using a maximum likelihood estimation (MLE) approach. Estimates for alpha are accurate to an acceptable 1%, for sample sizes of around 50 or more (according to the analyses of Clauset et al. (2009)). As some of the categories analysed in this study are less than this, particularly for leisure facilities and pharmacies, as well as for many facilities when considering violent crime, it is acknowledged that some of the results discussed below should be considered only tentative, at best.

To calculate the MLE for α it is necessary to first estimate x_{min} (and discard data that fall below this). This is often done through visual inspection of log-log plots, but this is also criticised in the literature, notably for its subjectivity (Clauset et al., 2009; Gillespie, 2015). Instead the method discussed here (and used in this study) involves a Kolmogorov-Smirnov approach, where x_{min} is estimated using the value that minimises D, the maximum distance between the cumulative distribution functions (CDFs) of the data and the fitted model. Taking such an approach is important because values of x_{min} that are unnecessarily high, exclude what would have been valid data, whilst estimates that are too low bias the estimate of α as they are calculated on data that does not (all) fit a power-law.

Despite the many claims in the literature of the almost universality of this type of distribution, Clauset et al. (2009) state that "Most previous empirical studies of ostensibly power-law distributed data have not attempted to test the power-law hypothesis quantitatively." (p675). This certainly appears to be the case for risky facilities. The main purpose of this part of the current study was therefore to statistically test the goodness of fit of this type of distribution for facility-crime data, using the poweRlaw package in R. In terms of data preparation, all that was necessary was to produce .csv files containing two columns, one populated with the crime count for each premises, the other populated just with the number 1, to represent each single data point (address). Such files were created for all the categories that were to be considered: *All* facilities, each facility category, and each of these split into acquisitive or violent crime, and daylight or darkness hours. The poweRlaw package was downloaded into R and the necessary code written to import, plot and analyse the data, a copy of which can be found in Appendix 3.

The first stage of the analysis was to use this package to calculate x_{min} and α for my datasets, using the above method, and to fit the data to a power-law distribution. Once this was done, the goodness of fit (or plausibility) was then estimated, again using the poweRlaw package. This was achieved by producing synthetic datasets drawn from the same model (that is to say, with the same alpha and x_{min} values) then, using the technique as already presented, individually fitting these sets to a power-law model and calculating the K-S statistic for each of them. The final p-value was obtained by counting the fraction of the time the K-S values for the synthetic datasets were larger than the K-S value for the empirical data (Clauset et al., 2009). Again, this was all done using the poweRlaw package's bootstrap function, which was set at 5000 iterations, to produce sequential mean and standard deviation estimates for the calculated alpha and x_{min} parameters.

The resulting p-value related to the hypothesis that data was *not* generated from a power-law distribution (thus the null hypothesis was that the data *were* generated from a power-law distribution).[7] Thus, when p tended to zero, it could be concluded that the power-law distribution was not a good fit. An acceptable significance threshold had to be determined, but simply, this meant that when p was greater than or equal to this threshold, the null hypothesis could not be rejected and the data *may* have been generated from a power-law distribution, but when p was less than this threshold, the null hypothesis could be rejected such that the data were *not* considered to be generated from a power-law distribution. The typical thresholds for significance may be applied, but Clauset et al. (2009) use 0.1, as they consider an alpha of 0.05 unnecessarily difficult to achieve, and further advise that this should be a decision for the informed researcher. Given the relatively small sample sizes in the current study (in the context

[7] In R this value is represented by the output bs_p$p.

of other empirical analyses of power-law distributions), I also chose to use an alpha of 0.1 (two-tailed) for this part of the study. It is also worth noting here that Clauset et al. (2009) caution that high p-values are more likely for small samples, and when this is the case results must be treated with caution. This was an issue in the current study, particularly when the data were separated into different categories.

A second method for considering the appropriateness of a power-law distribution was also used, because even if the previous method suggested the data did fit a power-law distribution (and it should be noted that selecting an appropriately large x_{min} can allow any data to be 'fitted' to such a distribution), this does not mean it was the *best* model. Clauset et al. (2009) therefore suggest comparing the power-law model to an alternative, to see which is a better fit. This could be done using the K-S test, again, however they suggest that using a likelihood ratios test is simpler, and perfectly suitable for this endeavour. Such an approach was used in this study, and was interpreted such that the sign of the resulting log likelihood ratio *R* indicated which of the two models was the better fit. A negative test statistic suggested the alternative distribution was a better fit, whilst a positive test statistic suggested the power-law distribution was a better fit. In addition, again following Clauset et al. (2009), the Vuong method was used to calculate, from the expected fluctuations, a p-value, thus showing if the observed sign of *R* was statistically significant. All of these calculations were also produced using the poweRlaw package.

In this case the power-law distribution was compared with a log-normal distribution. This was felt to be a plausible alternative distribution, as it also commonly occurs in relation to the distribution of natural phenomena. Additionally, for each category analysed, the complementary cumulative distribution function was first visually inspected against power-law, Poisson, and log-normal lines of best fit and, as will be shown in the results, the power-law and log-normal distributions generally appeared more feasible, whilst the Poisson distribution did not, thus it was not tested beyond this. It should be noted that in some instances, one or more of these fit lines could not be plotted, as the calculations carried out automatically in the poweRlaw package to produce these returned NaN (not a number). For the comparison, alpha was again set at the more generous 0.1 level (two-tailed).

From carrying out these two tests it was possible to draw conclusions regarding when the power-law could be rejected as the best fit form of distribution, or when it could not be ruled out. The results of this analysis are presented below.

7.3 Results

I now consider the results produced from the analyses described above, firstly looking at the general pattern and distribution of facility crime and the extent of concentration of crime in particular addresses (through presenting the law of crime concentration results, the 80-20 results and the Gini coefficients for all facility crime). I then assess the pattern of this distribution with reference to the risky facility concept and the J-curve plot. Next I consider the same outputs in comparative context, looking at facility type, crime type and time period, before outlining the results relating to the persistency of spatial facility crime concentration over the ten-year period analysed, and the consistency of ordering of the most 'risky' premises. Finally, I present outputs relating to the goodness of fit to a power-law distribution.

7.3.1 Concentration

As already stated, during the period studied there were 30,735 offences recorded in 2,905 distinct premises that had been designated as 'facilities'. The mean number of offences per facility was 10.58, but the range was 1 to 1,968 (SD=54.96). This demonstrates that offences were not equally distributed amongst premises.

Following Weisburd (2015), the percentages of premises contributing 25% and 50% of recorded crime were calculated as 0.45% and 2.03% respectively, figures that are very similar to those seen for other studies using different units of analysis, such as street segments (Gill, Wooditch & Weisburd, 2016; Weisburd, 2015). The proportion of premises contributing 80% of recorded crime was 12.22%, below the 'expected' 20%. Both sets of results suggest that offence distribution is highly skewed, and that a small number of premises have contributed a disproportionate amount of crime.

However, I have already proposed that the Gini coefficient is a better comparative measure and allows us to better assess the extent to which crime is unevenly distributed. Remembering that a completely equal distribution = 0 and a completely unequal distribution = 1, the Gini coefficient (G) for all facility crime was 0.8144. Though, as stated, there is no 'cut off' point for concluding that a Gini coefficient is high, a consideration of the research using this approach (within and outside the field of criminology) demonstrates that a coefficient of 0.8144 can be considered as evidence of extremely unequal distribution of crime; even more so if we consider that the sample excludes zero-crime locations. This result is also consistent with the limited existing research that uses such a measure in relation to spatial distribution of crime.

7.3.2 Risky facility pattern

The ranked crime count plot produced using the method described above, showed, as expected from the literature and the previous concentration outputs, a clear J-curve, as shown in figure 1. This pattern is consistent with the concept of risky facilities.

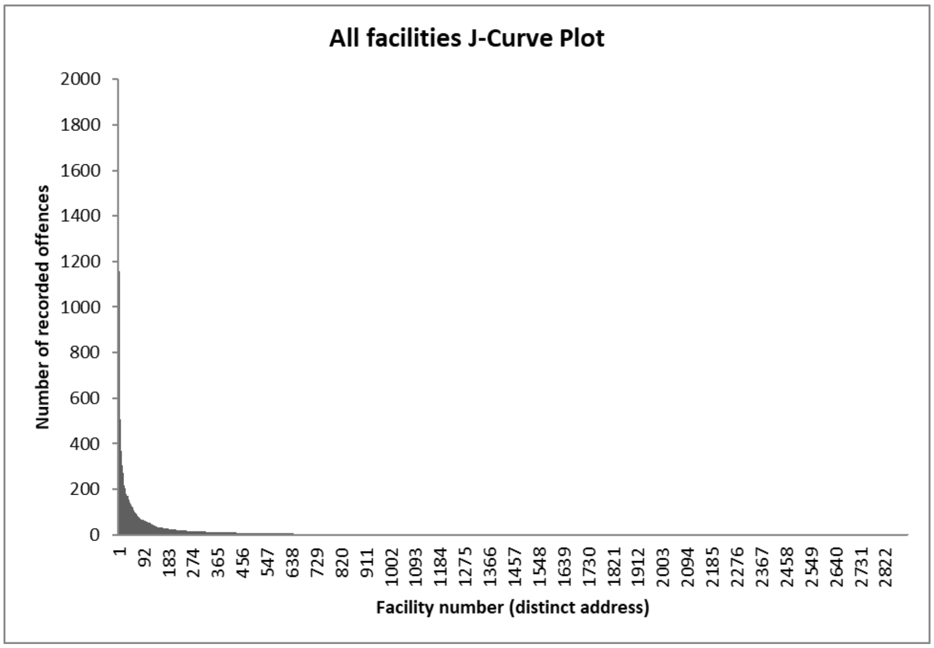


Figure 1: All facilities J-curve plot

7.3.3 Comparison

So far it has been shown that when considering all types of facility, there is strong evidence of an extremely unequal crime distribution, with a small proportion of premises contributing a large proportion of crime. In other words, that facility crime is concentrated in a relatively small number of addresses.

In this section, the results of the concentration and risky facility pattern analyses are presented comparatively. Firstly, the type of facility is considered. Next the outputs are compared by crime type (violent and acquisitive crime). Finally, different patterns for daylight and darkness offending are considered.

7.3.3.1 Comparison by facility type

Certain facility types contribute a greater proportion of overall facility crime. Table 4 also shows that when the mean number of offences per address is calculated and compared across facility types the variability remains, but the ordering of the categories changes, though those that tend to contribute higher numbers of offences overall, still tend to have higher mean levels of offences per address and vice versa. The exceptions to this are the *retail other* and *restaurant* categories, which have means in the lower half of the table, and *pharmacy* and *leisure*, which have lower overall numbers but higher means. Indeed, *pharmacy* has the highest mean rate of offending per premises with 20.75 offences per address, very closely followed by supermarket (20.74). Also higher than the *all facilities* mean of 10.58 offences per address are *convenience store*, *educational* and *gas station*.

Table 4: Facility category and number and proportion of offences

Class (basic description)	**N** (premises)	**Number of recorded offences**	**Percentage of recorded facility-crimes**	**Mean offences per facility**
Convenience store (smaller than supermarkets, with long opening hours, selling limited range of household goods and grocery items. Usually also sell tobacco products and often alcohol.[8]	255	4276	13.91	16.77
Supermarkets (supermarkets, hypermarkets, mega-stores, and specialist food stores)	164	3402	11.07	20.74
Educational (nurseries, schools, colleges and universities)	205	2735	8.90	13.34
Petrol (gas/petrol stations; may also have small shops)	205	2380	7.74	11.61
Restaurant (cafes, take-aways and restaurants; may have on-license but primary purpose is food service)	290	1688	5.49	5.82

[8] The nature of the coding means these also include most gas station addresses, but crimes recorded here were considered associated with the store. This is discussed in greater detail, elsewhere.

Class (basic description)	N (premises)	**Number of recorded offences**	**Percentage of recorded facility-crimes**	**Mean offences per facility**
Financial (banks, money-lending, and other facilities with primary purpose of money handling/transfer)	146	1162	3.78	7.96
Pharmacy (stores with primary purpose of providing medicines, including those requiring prescription)	48	996	3.24	20.75
Religious (all establishments recognised as a place of worship or spiritual contemplation; may provide additional services, e.g. limited childcare, soup kitchen, etc.)	249	719	2.34	2.89
Licensed premises (establishments with on-license; primary business is sale of alcohol for consumption on the premises; entertainment may or may not be provided)	115	627	2.04	5.45
Healthcare (hospitals, surgeries, doctors', dentists' and other healthcare providers' premises, including allied healthcare and rehabilitative services)	171	524	1.70	3.06
Hotel (all hotels and motels, regardless of length of stay)	119	434	1.41	3.65
Leisure (leisure and entertainment facilities, including fitness centres, movie theatres, bowling alleys, etc.)	40	322	1.05	8.05
Retail other (single and multiple premises involved in retail, not included in any other category; can include shopping and strip malls if share main address)	1498	10558	34.55	7.05
Other (all facilities not included in any other category)	203	852	2.77	4.20
Liquor store (shops with primary business involving sale of alcohol for consumption off-premises)	13	60	0.20	4.62
All	**2905**	**30735**	**100.00**	**10.58**

Following the method used for all facilities, the results are now presented for 25%, 50% and 80% crime contributions, the Gini coefficients and the ranked crime count plots, for each facility category. As already noted, because the Gini coefficient is a relative measure, it is particularly suited to this task.

Although the percentage of premises contributing 25% and 50% of recorded crime occurring in all facilities was small (0.45%; 2.03% respectively), this was not always the case when the data were separated into different facility types. For a 25% contribution, the bandwidth in this study (using just the twelve facility types, excluding *retail other*) was 0.58% to 6.43% (*retail other* was 0.13%), as shown in table 5. For a 50% contribution, the bandwidth was 2.44% to 19.68% (with *retail other* being lower at 0.87%) (see also table 5). It can be seen that regardless of whether using the 25% or 50% contribution measure, the ordering of facility types by percentage of premises changes very little. The only real exception to this was *healthcare*, which can be explained as by far the greatest crime contributor was one address (which unsurprisingly was the regional hospital) with 169 recorded offences, whilst the remaining addresses contributed between 1 and 17 offences each. For all facility types a disproportionately small number of premises contributed to the first 25% and 50% of crime, but this was particularly the case for *educational*, *hotel*, and then *supermarket* categories. The diverse - and large - category of *retail other*, though, required by the far the smallest percentage of facilities to account for the designated proportions of crime. Despite this final observation, there does not appear to be any pattern in the proportion of premises contributing 25% and 50% of crime, and either the number of premises within a facility category, the crime count or the mean.

Table 5: Facility type and various representations of crime concentration

Facility category	% of premises contributing 25% of recorded crime	% of premises contributing 50% of recorded crime	% of premises contributing 80% of recorded crime	Gini coefficient	Adjusted Gini coefficient
Educational	0.98	2.44	6.83	0.8643	0.8686
Supermarket	1.83	4.27	9.76	0.8451	0.8503
Petrol	2.44	6.34	20.00	0.7377	0.7413
Convenience	3.53	8.63	20.78	0.7368	0.7397
Pharmacy	4.17	10.42	22.92	0.7206	0.7359
Financial	3.42	8.90	22.60	0.7004	0.7052
Hotel	1.68	4.20	26.89	0.6708	0.6765
Licensed premises	5.22	13.04	30.43	0.6137	0.6190
Healthcare	0.58	7.02	39.77	0.5871	0.5906
Leisure	5.00	15.00	30.00	0.5829	0.5979
Restaurant	5.17	15.17	37.93	0.5586	0.5606
Religious	6.43	19.68	50.60	0.4397	0.4415
Retail other	0.13	0.87	14.49	0.7907	0.7912
All	**0.45**	**2.03**	**12.22**	**0.8144**	**0.8147**

Briefly considering the proportion of premises contributing 80% of crime, shows a somewhat different pattern. *Educational* and *supermarket* require the smallest percentage of premises with 6.83% and 9.76% respectively, suggesting these types of facilities have the greatest degree of concentration of offending within a small proportion of premises. This is consistent with what was seen for the Weisburd suggested proportions (of 25% and 50%), but there are differences. *Hotel* no longer ranks highly (sitting in the middle with 26.89%), and although *retail other* still demonstrates a highly disproportionate offence contribution, with 14.49% of premises contributing 80% of crimes, both this and the *all* category (already presented above) appear to have crime less concentrated in facilities when 80% is used as the cut-off, than do *educational* or *supermarket.* This was not the case for the 25% and 50% tests. This suggests the possibility that different value 'cut offs', which are seemingly arbitrarily selected, have the potential to identify different patterns of concentration when comparing across different categories (such as facility types, street blocks, crimes, and so forth). The differences observed could be explained by variability in the spread of crime, and should also be reflected in the shape of the J-curve (considered below).

A comparison of the inequality of distribution for each facility type using the Gini coefficient is also shown in table 5. The most unequal distribution occurs in *educational* (G=0.8643) closely followed by *supermarket*. The majority of facility types have Gini coefficients above 0.6, with just *healthcare*, *leisure*, *restaurant* and *religious* falling below this. Again, as zero crime facilities are excluded, it can be concluded that many of the Gini coefficients are particularly large (in comparison to other research in the crime and place and criminal careers fields). Overall, each facility type analysed experienced an unequal distribution of offending across the premises within that class, but for some facility types, the extent of this inequality is much greater.

As was explained, above, the Gini coefficient can be affected by a very small sample size. As already noted, there was no obvious pattern in the results presented to suggest a relationship between sample size (number of premises within a category of facility) and the size of the coefficient produced (thus *apparent* extent of inequality of distribution). However, adjusted Gini coefficients are also presented for the smaller samples (this is particularly important for the other categories of comparison considered later). These are shown in table 5, but for facility category, the results obtained are not notably different, with just two facility types switching position (*leisure* moving above *healthcare*).

Interestingly, when comparing across the different facility categories, the rank position based on the Gini coefficient (or adjusted coefficient) is much more similar to the rank position resulting from considering the proportion of premises contributing 80% of crime, than it is the proportion contributing 25% or 50%. This may be because the former captures (and is affected by) a greater amount of the sample. This finding also suggests that, if using cut-offs, 80% may be more appropriate for considering the overall inequality of crime distribution, but 25% (and less so 50%) may better capture offending that is extremely concentrated in only a few facilities. Only the Gini coefficient, however, encompasses the whole sample and this measure seems to be the clearest and most consistent way to compare across categories, units of analysis and studies. It is recognised, though, that even this measure can be less reliable when sample sizes (number of premises and/or crime count) are small.

The ranked crime count plots for all thirteen facility types showed that all conformed to a J-curve type distribution of offences, with a relatively small proportion of premises contributing a large proportion of crime. This pattern is clearest for *educational*, *healthcare*, *hotel*, *gas station* and *supermarket*. The ranked crime count plot for *educational*, as an example, is shown in figure 2 (copies of all the J-curve plots – by facility type only – can be found in Appendix 4).

Taken together, these results demonstrate that there is an inequality of distribution amongst facility crime and that this pattern holds when different facility types are considered. However, it is more apparent – more unequal – for certain types of premises. As existing research of this type is limited there is no agreed definition of what a 'narrow bandwidth' is. Nor is there a cut-off for Gini coefficients of interest. However, it is clear that facility types in the sample that experience the greatest degree of inequality of crime distribution are educational establishments, supermarkets, gas stations, convenience stores and pharmacies. Those facilities experiencing less skewed crime distributions are restaurants and, depending upon the method used, licensed premises and leisure facilities. Religious premises clearly experience the least skewed crime distribution. Healthcare facilities are unusual in that a very small number of addresses contribute a significant proportion of recorded crime, but overall the distribution of crime is more evenly spread than for many other types of facility. As proposed, this is likely to be because of the very different sizes of facility included in this category; notably the contribution of the hospital. In this data set, licensed premises also tended to be associated with less skewed recorded offences, which is inconsistent with previous research, and discussed further in later chapters.

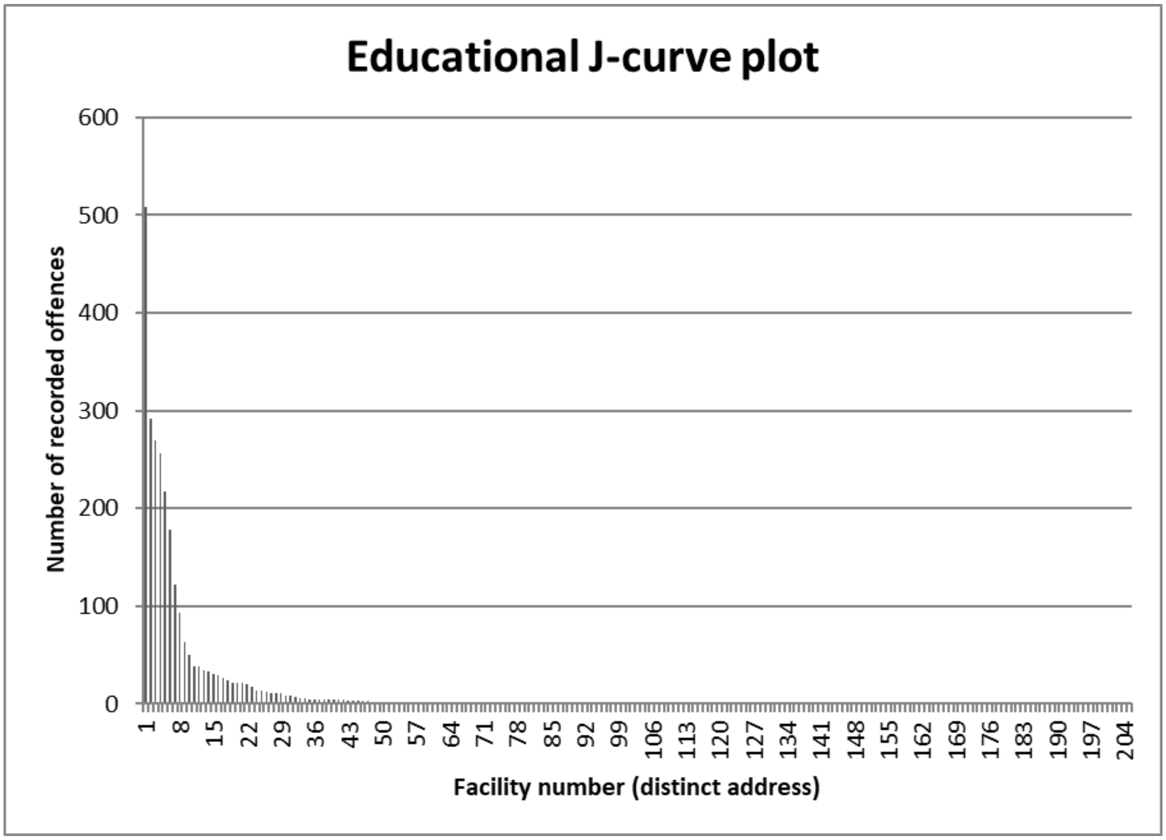


Figure 2: Educational facilities J-curve plot

These results are also consistent with the concept of risky facilities. Not only are recorded offences unequally distributed across premises within each facility type, but the J-curve plots show that it is indeed a small proportion of addresses that are responsible for a large proportion of these crimes, whilst many premises contribute very little crime each. Which of the premises should be 'labelled' as risky, and which not, is dealt with in the next chapter.

7.3.3.2 Comparison by crime type

Similar analysis was carried out for each facility type, comparing acquisitive and violent crimes, to see if one type of crime tended to be more unequally distributed and to see if different types of facility demonstrated different patterns.

Considering all facility crime, 85.74% (n=26,351; facilities=2,470) of recorded offences were categorised as acquisitive crime and 5.10% (n= 1,567; facilities =404) were categorised as violent crime. The fact that such a small proportion of offences recorded in facilities would be categorised as violent is not unexpected, as in most jurisdictions this type of offence accounts for a smaller proportion of total recorded crime than do acquisitive offences. Further, it is anticipated that violence is also more likely to occur in domestic settings and on the street. Going back to the non-facility crime recorded in the original dataset, the breakdown was 61.5% acquisitive and 10.7% violent.

Considering all facilities, a 25% contribution to acquisitive crime resulted from 0.45% of premises, whilst for violent crime this was 0.99%. A 50% contribution to acquisitive crime was made by 2.15% of facilities, with 5.20% of premises contributing this proportion of violent crimes. The acquisitive crime figures are very similar to those for all crime, which is to be expected as this crime type accounts for a large proportion of recorded offences. In addition, albeit a greater proportion of facilities is required, the figures for violent crime also show that a small proportion of premises contribute a large proportion of crime. So far, it appears that the law of crime concentration may also apply across different types of crime.

The concentration of different crime types for the twelve facility categories was also analysed. For acquisitive crime the bandwidth was 0.68% (*healthcare*) to 6.67% (*leisure*) for 25% of crime and 3.94% (*educational*) to 19.07% (*religious*) for 50% of crime. The figures for all analysed facility categories are shown in table 6. As above, it seems that bandwidth tends to be larger as categories get smaller: all facility types considered together demonstrate a much greater degree of concentration in a small number of premises, than do individual facility categories. That said, acquisitive crime can still be seen to be disproportionately concentrated in a (often small) number of premises for each facility type.

The bandwidths for violent crime were 2.38% (*educational*) to 50.00% (*pharmacy*) for a 25% crime contribution and 3.45% (*healthcare*) to 60.00% (*financial*). The full results are shown in table 6. Some of the premises and crime counts here were very small, thus little meaning can be extracted from these results. For example, 50% of pharmacies contribute both 25% and 50% of violent crime. This is because there were only two pharmacies in the dataset that recorded any violent crime (5 offences in total). The financial category was also very small (five premises and five offences). This suggests that when considering the ubiquity of concepts such as the law of crime concentration and risky facility distributions, there may need to be a minimum category or count size applied. As research in this area moves towards ever smaller and distinct units of analysis this issue will need to be further addressed.

Table 6: Facility categories and concentrations by crime type

Facility category	Acquisitive Crime: N premises (crime count)	Acquisitive Crime: % of premises contributing 25% of recorded crime	Acquisitive Crime: % of premises contributing 50% of recorded crime	Acquisitive Crime: Gini coefficient	Acquisitive Crime: Adjusted Gini coefficient	Violent Crime: N premises (crime count)	Violent Crime: % of premises contributing 25% of recorded crime	Violent Crime: % of premises contributing 50% of recorded crime	Violent Crime: Gini coefficient	Violent Crime: Adjusted Gini coefficient
Supermarket	155 (3265)	1.94	4.52	0.8436	0.8491	29 (49)	10.34	27.59	0.3125	0.3236
Educational	127 (1211)	1.57	3.94	0.8090	0.8155	84 (802)	2.38	4.76	0.7822	0.7916
Petrol	195 (2251)	2.56	6.67	0.7394	0.7432	24 (32)	12.50	33.33	0.2057	0.2147
Pharmacy	47 (977)	4.26	10.64	0.7170	0.7362	2 (5)	50.00*	50.00#	0.3000+	0.6000
Convenience	216 (3884)	4.17	9.72	0.7069	0.7102	60 (129)	8.33	21.67	0.3975	0.4043
Financial	142 (1130)	3.52	9.15	0.6996	0.7046	5 (5)	40.00**	60.00##	0.0000++	0.0000
Hotel	48 (237)	4.17	8.33	0.6838	0.6984	28 (39)	7.14	32.14	0.2555	0.2650
Leisure	30 (252)	6.67	16.67	0.5249	0.5430	15 (17)	20.00	46.67	0.1020	0.1092
Healthcare	148 (392)	0.68	10.14	0.5392	0.5429	29 (65)	3.45***	3.45###	0.5337	0.5527
Restaurant	267 (1411)	5.62	15.73	0.5389	0.5409	45 (76)	8.89	26.67	0.3123	0.3194
Licensed premises	87 (316)	5.75	16.09	0.5313	0.5375	46 (142)	8.70	21.74	0.4329	0.4426
Religious	215 (563)	6.51	19.07	0.4352	0.4372	14 (16)	14.29	42.86	0.1161	0.1250
Retail other	1353 (9839)	0.15	0.81	0.7983	0.7989	67 (123)	4.48	19.40	0.3973	0.4033
All	**2470 (26351)**	**0.45**	**2.15**	**0.8111**	**0.8114**	**404 (1567)**	**0.99**	**5.20**	**0.6561**	**0.6577**

Next, the Gini coefficient and adjusted Gini coefficient (for small samples) were calculated. When considering all facilities together, it was clear that acquisitive crime was more unevenly distributed (G=0.8111; N=2470) than violent crime (G=0.6561; N=404), as also shown in table 6. Possible explanations for this difference are considered in the discussion, below. However, it is also necessary to note that, whilst the Gini coefficient is not affected by the underlying prevalence (hence its suitability for comparison) we do know that small samples (and the exclusion of zero crimes) results in a downward bias. The smaller the number of premises and/or the crime count the less highly concentrated the crime count can be. Therefore, smaller Gini coefficients would be expected for violent crime than for acquisitive crime as the number of offences (in this study in particular) is much fewer. Additionally, the smaller Gini coefficient seen for violent crime could also be explained by there being a larger number of zero crime premises, which continue to be excluded.[9]

The Gini coefficients were then compared for each facility type, by and across crime type. These results must be interpreted with caution. As noted above, some of the sample sizes for violent crime were too small to be meaningful (though the results are included for completeness, those of greatest concern are highlighted). Additionally, the adjusted Gini (hereinafter G_{adj}) has been compared throughout to reduce some of the possible bias. All the result are again shown in table 6.

The categories that experienced the most unevenly distributed acquisitive crime were *supermarket* (G_{adj}=0.8491) and *educational* (G_{adj}=0.8155). The former may be expected as acquisitive crime is likely to be most prevalent in retail premises and within such environments, certain premises may be 'better' crime targets than others (for reasons such as layout, place management, presence of 'hot products'). Indeed, the other retail categories: *gas station*, *pharmacy* and *convenience store*, all had adjusted Gini coefficients greater than 0.7. Financial establishments, such as banks, also experienced a concentration of crime in a small proportion of premises. A relatively high prevalence of acquisitive crime may be expected here, but the distribution of crime might have been anticipated to be more even. That educational establishments had the second highest degree of concentration may be more surprising.

Educational establishments also experienced the most unevenly distributed violent crime (G_{adj}=0.7916). Indeed, this type of facility also had the greatest number of recorded violent

---

[9] Only premises with one or more violent crimes are included in this part of the analysis, even though the addresses of a number of zero-crime premises are known (because they recorded some other form of crime). However, to include these would introduce inconsistencies across the analyses by category, whilst also continuing to exclude those zero-crime premises not in the original, full dataset

offences across the greatest number of facilities. No other facility types recording an analysable number of offences were as unevenly distributed as this. The next highest adjusted Gini co-efficient (for a category with more than a minimal number of offences) was recorded for the *healthcare* category. Considering the small sample sizes and the exclusion of zero crime premises, relatively high adjusted Gini coefficients were also found for *convenience store, licensed premises* and, to a lesser extent, *supermarket* and *restaurant*. These results suggest that violent crime is unevenly distributed within specific categories of facilities, but less so than acquisitive crime. For both types of crime, offences were not evenly distributed across premises, but the extent of the inequality of distribution varies quite substantially by facility type. The results are also affected by the very small number of offences and premises experiencing these for many of the categories of facility.

With caution because of the small number of violent offences recorded for most facility types, it can also be stated that for nearly all the categories analysed acquisitive crime is more unevenly distributed than violent crime. However healthcare premises did have a slightly higher adjusted Gini coefficient for violent crime (0.5527 versus 0.5429) despite experiencing less offences across less facilities, compared to acquisitive crime. Educational establishments and, to a lesser extent, licensed premises also had adjusted Gini coefficients that were not vastly lower for violent crime (again, despite having fewer offences).

To explore further the nature of the apparent unequal distribution, J-curve plots were also produced for acquisitive and violent crime in all facilities and for each facility category. It is arguable whether *healthcare* demonstrates an acquisitive J-curve distribution, though this is in part (like many of the other results for this type of facility) because of the large contribution of the regional hospital. For all other facility types, acquisitive crime is distributed amongst facilities along what could be considered a J-curve of concentration (albeit this is not always ‘sharp’). For violent crime, the *educational* category displays a clear J-curve and licensed premises also follow this pattern (though not as neatly). For all other facility types the number of facilities and the range of offences recorded are too small to present a true J-curve when plotted. That said, for nearly all of these, visually they conform loosely to this type of distribution, with some premises disproportionately contributing crime.

7.3.3.3 Comparison by time period

Finally for this part of the analysis, daylight and darkness offences were compared. More offences occurred during daylight hours, but the difference was not as large as for crime type.

Recorded facility crimes were deemed to have occurred during daylight hours in 68.47% (n=21,044) of cases, thus 31.53% (n=9,691) were attributed to hours of darkness. However, this did vary quite considerably for different types of facility as will be shown.

As seen with the other methods of categorisation, above, when all crime is considered, a very small proportion of facilities contributed 25% and 50% of offences. In this case, for daylight crimes it was 0.43% and 2.02% of facilities (for 25% and 50% respectively). For darkness crimes, of which there were fewer, the results were 1.12% and 4.49%.

Once again, the pattern by facility type was also considered. For daylight offences, the bandwidth for a 25% offence contribution was 0.69% (*healthcare*) to 9.09% (*licensed premises*). As healthcare is known to be affected by the regional hospital, it may be more appropriate to consider the bandwidth range starting with *educational*, at 1.20%. The bandwidth for a 50% contribution was 2.40% (*educational* again) to 21.21% (*licensed premises* again). These results are once again similar in magnitude and range to those found for the other methods of categorisation.

For darkness offences, the proportion of facilities required for a 25% or 50% crime contribution was greater. This is consistent with the above findings, as there tended to be fewer darkness offences. The bandwidths were 1.61% (*healthcare*) or 2.47% (*educational*) to 11.11% (*leisure*), and 6.45% (*healthcare*) or 6.76% (*hotel*) to 26.39% (*religious*) for 25% and 50% of crime respectively. The results are shown in table 7.

Table 7: Facility categories and concentrations by time of day

Facility type	Daylight offences: N premises (crime count)	Daylight offences: % of premises contributing 25% of recorded crime	Daylight offences: % of premises contributing 50% of recorded crime	Daylight offences: Gini coefficient	Daylight offences: Adjusted Gini coefficient	Darkness offences: N premises (crime count)	Darkness offences: % of premises contributing 25% of recorded crime	Darkness offences: % of premises contributing 50% of recorded crime	Darkness offences: Gini coefficient	Darkness offences: Adjusted Gini coefficient
Supermarket	114 (2463)	2.63	6.14	0.8153	0.8226	102 (939)	3.92	7.84	0.7339	0.7412
Educational	167 (2394)	1.20	2.40	0.8622	0.8674	81 (341)	2.47	7.41	0.6389	0.6469
Petrol station	180 (14910	2.78	7.22	0.6997	0.7036	126 (889)	3.97	8.73	0.6657	0.6711
Pharmacy	43 (786)	4.65	11.63	0.7025	0.7192	28 (210)	3.57	14.29	0.5925	0.6145
Convenience store	201 (2192)	4.48	11.94	0.6570	0.6603	164 (2084)	4.27	10.98	0.6716	0.6757
Financial	122 (878)	4.10	9.84	0.6676	0.6731	72 (284)	6.94	15.28	0.5355	0.5430
Hotel	63 (215)	3.17	7.94	0.6317	0.6419	74 (219)	2.70	6.76	0.5829	0.5909
Leisure	35 (208)	2.86	17.14	0.5571	0.5735	27 (114)	11.11	18.52	0.4568	0.4744
Healthcare	144 (379)	0.69	9.72	0.5330	0.5367	62 (145)	1.61	6.45	0.5398	0.5487
Restaurant	218 (801)	5.96	18.81	0.4742	0.4764	231 (887)	5.63	16.88	0.5006	0.5028
Licensed premises	66 (135)	9.09	21.21	0.3813	0.3871	92 (492)	5.43	14.13	0.5890	0.5955
Religious	208 (499)	7.69	21.15	0.4046	0.4065	144 (220)	8.33	26.39	0.2816	0.2836
Retail other	1174 (7942)	0.17	0.85	0.7920	0.7926	732 (2616)	0.41	4.10	0.6463	0.6472
All	**2323 (21044)**	**0.43**	**2.02**	**0.8004**	**0.8007**	**1602 (9691)**	**1.12**	**4.49**	**0.7124**	**0.7128**

Here it can be seen that the percentage of facilities contributing 25% of crime tended to be quite similar for both daylight and darkness offences within any given category, particularly for *convenience store*, *hotel*, *religious* and *restaurant*. Following the patterns above, it was also anticipated that the proportion of facilities contributing this amount of crime would tend be smaller for daylight than darkness offences, simply because there were more of the former, though it is noted that certain types of establishment, by virtue of their business, would likely experience more darkness crimes, when this pattern would not hold.

The results showed this was partly as anticipated, but not entirely. As well as there being some facility types that were quite different in their contribution across daylight and darkness (notably *leisure* and *licensed premises*), there were five types for which a smaller proportion of facilities were required to achieve a 25% contribution to darkness crime than to daylight crime and of these, three (*hotel*, *licensed premises* and *restaurant*) experienced a greater overall number of darkness offences. The remaining two categories were *convenience store* and *pharmacy*. For these types of premises, even though they had fewer offences during hours of darkness, offending at this time was concentrated in a (slightly) smaller proportion of facilities. It should be noted, however, that for pharmacies, the 25% of darkness offences were recorded at just one premises (contributing 27.62% of crime).

The Gini coefficients and, as some of the samples are small, the adjusted Gini coefficients, by facility type for offences during daylight and offences during darkness were compared, and are also shown in table 7. Overall, many more offences occurred during the hours of daylight and the Gini coefficient was higher (G=0.8004), indicating daylight offences were more unequally distributed. It should be noted, however, that the concentration of darkness offences for all facility crime was also high, with a Gini coefficient of 0.7214.

Calculated adjusted Gini coefficients were of a magnitude to conclude that crime is unevenly distributed, for nearly all facility types studied, for both daylight offending and darkness offending, albeit the extent of this concentration varied. Educational establishments and supermarkets experienced particularly highly concentrated daylight offending in contrast to restaurants, religious establishments and licensed premises. The adjusted Gini coefficients for darkness offending tended, following the pattern seen throughout for lower volume offences, to be smaller than the daylight coefficients. There was again variation by facility type, with *religious* showing only minimal concentration in particular premises (G_{adj}=0.2836). The *supermarket* category experienced the highest degree of unevenness of distribution during darkness hours (G_{adj}=0.7412) but all other facility types had adjusted Gini coefficients between 0.6757 (*convenience store*) and 0.4744 (*leisure*). Therefore, it can be concluded that for nearly all facility types, offending was not evenly distributed during darkness hours, that offending was generally less concentrated in particular premises during darkness hours, but that the variability in this concentration was less than for daylight offending.

However, this general pattern did not hold for every facility type. Darkness offences were more (not less) unequally distributed than daylight offences for licensed premises and restaurants. In other words, 'risky' premises within these facility types are more likely to be found during hours of darkness than during daylight. Offences were slightly more unequally distributed for *convenience store* and *healthcare* during darkness hours as well, though the difference was small, suggesting that the riskiness of addresses for these facility types tended to be fairly consistent (as was also the case for *gas station*).

As already also seen for the other forms of comparison, there was a tendency for the period with more recorded offences to also have the higher Gini coefficient. When looking also at the facility type, in most cases the period with the greater number of offences was also the period when usage of the facility was likely to be higher. However, as can be seen in table 7, for convenience store, hotel and healthcare, the period with the largest Gini coefficient was not the same as the period with the most offences. Therefore, although the number of offences may influence the Gini coefficient statistically, or the pattern of distribution in reality, this does not always happen, thus it cannot be assumed that this is the only influence.

Finally, ranked crime count plots were produced and visually inspected. For all facility crime, both daylight and darkness plots were clearly J-curves. As with the previous analyses, when considered by facility type, all plots adhered – at least roughly – to a J-curve type distribution, though *healthcare* would be better described as having an outlier (as discussed previously).

7.3.3.4 Summary of comparison results

It can therefore be concluded that for different facility types, for acquisitive and violent crime (across facility types) and for daylight and darkness offences (across facility types) the risky facilities pattern appears to hold, though the extent to which this is the case, and the degree of inequality of distribution varies by facility type. I have also noted that the extent to which offences are distributed unequally across address tends to be in many cases related to the number of recorded offences. Therefore, it is difficult to separate offence incidence and offence concentration as for many outputs, higher incidence tends to coincide with higher concentration and/or closer adherence to the risky facility distribution. However, this is not always the case, suggesting that variability in concentration of offending and inequality of distribution are also affected by other factors, which are worthy of further exploration.

7.3.4 Persistency and consistency

The next results to be considered are those relating to the extent to which the risky facility pattern persists over time, and then how consistent this is in relation to the ordering of premises.

7.3.4.1 Persistency

Persistency was tested, as set out above, by considering the Gini coefficients and ranked crime count plots for each of the ten financial years in the study.

When all facilities were considered, the Gini coefficients per year ranged from 0.6205 (year 7) to 0.6875 (year 10) demonstrating that there was inequality of distribution for facility crime in each of the ten years studied, and that the extent of this varied very little over time. A similar pattern was seen when each facility category was analysed. This was particularly the case for *educational* and *supermarket* (as well as *retail other*). There was also little variation for *convenience store*, *hotel* or *licensed premises*. Indeed, when the Gini coefficient for one year was compared with that for the next year, there was usually variation of 0.05 or less. The results are shown in table 8. These results suggest that inequality of distribution is seen consistently over time for facility crime and for different classes of facility within this.

The nature of this concentration also appears, from visually inspecting J-curves, to be consistent with that attributed to risky facilities. This is certainly the case for all facility crime, with each year displaying a clear J-curve. As far as can be ascertained with the relatively small sample sizes, this was also true of each facility type, albeit the curves are less clearly defined and the data are less convincing for categories such as *religious*, *leisure* and *hotel* as they experience so few recorded offences. The *healthcare* facility outputs continue to be skewed by the presence of the hospital.

Therefore, it may be tentatively concluded that the risky facility distribution of offences is a persistent one, as although the number of crimes may be too small to draw firm conclusions, the pattern of offending suggests that the disproportionate contribution of some premises to the total amount of crimes within a facility type, holds across a number of years (in this case ten).

Table 8: facility category and Gini coefficient by year

		Convenience	Educational	Financial	Healthcare	Hotel	Leisure	LP	Petrol	Pharmacy	Religious	Restaurant	Supermarket	Retail other	All
FY1	N premises	103	42	49	30	21	17	41	84	16	45	115	54	353	847
	Offences	556	258	142	49	60	31	98	327	58	60	221	413	1158	3517
	Gini coefficient	**0.5731**	**0.6608**	**0.4829**	**0.3667**	**0.4952**	**0.3302**	**0.4151**	**0.5740**	**0.4461**	**0.1948**	**0.3636**	**0.6819**	**0.6506**	**0.6567**
FY2	N premises	98	43	44	28	14	12	39	79	21	58	103	58	397	873
	Offences	536	287	159	60	35	29	65	275	127	84	213	325	1289	3581
	Gini coefficient	**0.5240**	**0.6525**	**0.5530**	**0.4583**	**0.3980**	**0.3937**	**0.3457**	**0.5329**	**0.4949**	**0.2221**	**0.3725**	**0.6293**	**0.6377**	**0.6455**
FY3	N premises	96	48	40	34	10	16	31	70	21	54	94	48	353	809
	Offences	506	329	79	55	34	38	61	317	117	71	142	275	1082	3199
	Gini coefficient	**0.5340**	**0.6785**	**0.3744**	**0.3599**	**0.3765**	**0.2928**	**0.3670**	**0.5758**	**0.5120**	**0.1912**	**0.2616**	**0.5970**	**0.6193**	**0.6423**
FY4	N premises	94	55	39	33	16	12	36	72	17	58	91	42	299	738
	Offences	464	273	122	59	40	33	69	234	113	66	184	314	1064	3186
	Gini coefficient	**0.4764**	**0.6302**	**0.4313**	**0.4150**	**0.4125**	**0.3510**	**0.3643**	**0.5342**	**0.6132**	**0.1082**	**0.3356**	**0.6228**	**0.6535**	**0.6470**
FY5	N premises	90	43	40	25	22	14	28	69	15	57	85	37	320	753
	Offences	346	269	117	53	42	36	52	215	74	75	147	418	922	2851
	Gini coefficient	**0.4456**	**0.6861**	**0.4088**	**0.4951**	**0.3918**	**0.4008**	**0.3091**	**0.4942**	**0.5586**	**0.1918**	**0.3080**	**0.6007**	**0.6073**	**0.6352**
FY6	N premises	83	46	44	33	28	16	26	73	18	68	74	46	284	740
	Offences	402	312	118	51	57	34	47	240	71	116	111	307	902	2864
	Gini coefficient	**0.5356**	**0.6957**	**0.4792**	**0.3197**	**0.4167**	**0.4228**	**0.2921**	**0.5013**	**0.4734**	**0.3060**	**0.2621**	**0.6144**	**0.6275**	**0.6251**
FY7	N premises	91	49	38	32	33	16	42	70	19	47	99	35	293	765
	Offences	397	300	101	50	59	25	96	220	90	60	166	266	844	2762
	Gini coefficient	**0.4633**	**0.7162**	**0.4250**	**0.3388**	**0.3903**	**0.2825**	**0.3656**	**0.5036**	**0.5860**	**0.1652**	**0.3213**	**0.6294**	**0.6106**	**0.6205**
FY8	N premises	97	47	41	39	30	16	33	67	19	54	115	42	297	765
	Offences	389	267	133	59	51	50	75	274	123	64	189	298	1041	3084
	Gini coefficient	**0.5056**	**0.6950**	**0.4823**	**0.3173**	**0.3458**	**0.4275**	**0.3337**	**0.6204**	**0.4869**	**0.1406**	**0.2849**	**0.6203**	**0.6576**	**0.6447**
FY9	N premises	98	49	37	28	19	12	24	53	20	48	102	38	313	733
	Offences	361	227	83	50	29	23	37	117	118	62	179	373	1146	2879
	Gini coefficient	**0.5015**	**0.6617**	**0.3732**	**0.4129**	**0.2759**	**0.2790**	**0.2736**	**0.4109**	**0.5907**	**0.1841**	**0.3171**	**0.6775**	**0.6695**	**0.6478**
FY10	N premises	96	51	38	22	19	10	16	55	16	43	80	42	234	625
	Offences	319	213	108	38	27	23	27	161	105	61	136	413	1110	2812
	Gini coefficient	**0.5003**	**0.6337**	**0.4240**	**0.3828**	**0.2456**	**0.4304**	**0.3218**	**0.5563**	**0.4994**	**0.2242**	**0.3068**	**0.6968**	**0.7427**	**0.6875**
Minimum G FY1-10		**0.4456**	**0.6302**	**0.3732**	**0.3173**	**0.2456**	**0.2790**	**0.2736**	**0.4109**	**0.4461**	**0.1082**	**0.2616**	**0.5970**	**0.6073**	**0.6204**
Maximum G FY1-10		**0.5731**	**0.7162**	**0.5530**	**0.4951**	**0.4952**	**0.4304**	**0.4151**	**0.6204**	**0.6132**	**0.3060**	**0.3725**	**0.6968**	**0.7427**	**0.6875**

7.3.4.2 Consistency

To assess consistency, Spearman's rank order correlations were produced between each pair of financial years, on the data aggregated by premises (address). If there was a significant positive correlation, then it could be concluded that the more risky (highest ranked) facilities tended to be risky across both pairs of years, whilst the least risky tended to remain low ranked across both pairs of years, that is to say there was at least some consistency of 'riskiness'.

When all facility crime was considered, there was consistency of risky facilities, with all pairs of years being statistically significantly positively correlated at the 0.01 level. At the facility category level, this was also the case for *convenience store*, *educational*, *gas station* and *supermarket*. There were statistically significant positive correlations for all pairs of years at the 0.01 or 0.05 level for *financial* and *restaurant*. Also generally demonstrating consistency, with most pairs of years being statistically significantly positively correlated at the 0.01 or 0.05 level were *licensed premises* (financial year 10 was not significantly correlated with any year other than 3 and 9), *pharmacy* (years that were further apart were less well correlated) and *hotel* (financial year 1 was less often correlated with other years). Finally, there were very few statistically significantly correlated pairs of years for *healthcare* (6 out of 45 pairs), *leisure* (17 out of 45 pairs) and *religious* (8 out of 45 pairs), suggesting there was not consistency of risky facilities for these types of establishments. That is to say, the same addresses did not tend to be risky from one year to the next. As with much of the previous analysis, however, this latter set of facility types tended to have lower numbers of recorded offences, therefore it is harder to reach levels of statistical significance. It should also be noted that crime in these types of premises was more evenly distributed across addresses, as evidenced by the smaller Gini coefficients, thus the addresses that are 'risky' are less apparent, and movement in ranking is much more likely.

Interestingly, although most facility types showed statistically significant correlations across the ten years analysed, the correlation coefficients tended, in the main, to be small, suggesting that the precise ordering of the facilities within a category showed some variation. As the output tables for this part of the analysis are extensive, only the output for all facilities is included here (table 9).

Table 9: All facilities Spearman's rank order correlation by year

All Facilities	FY2	FY3	FY4	FY5	FY6	FY7	FY8	FY9	FY10	
FY1	0.419	0.37	0.348	0.349	0.327	0.303	0.276	0.28	0.285	Correlation Coefficient
	<0.001	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***	p
	2905	2905	2905	2905	2905	2905	2905	2905	2905	N
FY2		0.401	0.37	0.343	0.318	0.296	0.293	0.294	0.286	Correlation Coefficient
		<0.001	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***	p
		2905	2905	2905	2905	2905	2905	2905	2905	N
FY3			0.415	0.411	0.372	0.332	0.335	0.299	0.3	Correlation Coefficient
			<0.001	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***	p
			2905	2905	2905	2905	2905	2905	2905	N
FY4				0.416	0.395	0.353	0.36	0.371	0.319	Correlation Coefficient
				<0.001	***<0.001***	***<0.001***	***<0.001***	***<0.001***	***<0.001***	p
				2905	2905	2905	2905	2905	2905	N
FY5					0.435	0.352	0.362	0.349	0.355	Correlation Coefficient
					<0.001	***<0.001***	***<0.001***	***<0.001***	***<0.001***	p
					2905	2905	2905	2905	2905	N
FY6						0.402	0.385	0.367	0.344	Correlation Coefficient
						<0.001	***<0.001***	***<0.001***	***<0.001***	p
						2905	2905	2905	2905	N
FY7							0.375	0.377	0.355	Correlation Coefficient
							<0.001	***<0.001***	***<0.001***	p
							2905	2905	2905	N
FY8								0.404	0.369	Correlation Coefficient
								<0.001	***<0.001***	p
								2905	2905	N
FY9									0.424	Correlation Coefficient
									<0.001	p
									2905	N

The visual inspection of the top-ranked facilities for two example categories also provided support for the conclusion of consistency. Using my preferred method for identifying which premises should be considered ‘risky’ (as set out in the next chapter), the ranking positions of such premises tended to vary little over the period studied and often remained, if not in the same position, still within the top ranking premises (though there were exceptions).

The yearly ranks for all risky facilities for the example categories of *educational* and *supermarket* are shown in tables 10 and 11. A pattern of particular note is that the most highly ranked addresses overall appeared generally to be the most consistent. This was more the case for *educational.* The lower end of the top ranked premises tended to include some addresses that drop down when considered annually. There seemed to be somewhat more variability for *supermarket*, but overall the ‘problematic’ premises tended to remain the same (without any particular address dominating the rankings for the whole period). In addition, there are some premises that were not open for the full ten years, either closing down before the end of this period, or not opening until part way through. Though these accumulated enough offences during the years they were open to be identified as risky facilities overall, this obviously affected consistency of ranking.

Table 10: Consistency of rank position for educational establishments

Overall rank	Addnum	Count	FY1	FY2	FY3	FY4	FY5	FY6	FY7	FY8	FY9	FY10
			Annual rank									
1	2266	508	1	1	2	1	1	1	1	1	2	3
2	965	291	3	2	1	2	4	4	4	3	1	2
3	2660	269	5	3	5	3	2	3	3	6	3	6
4	2549	256	7	6	4	5	3	2	2	4	5	4
5	847	217	4	4	6	7	6	5	5	2	4	1
6	2758	178	12	8	3	4	5	6	8	8	6	5
7	2764	122	8	7	9	9	7	7	6	5	7	7
8	2272	93	2	5	8	14	18	14	19	9	10	11
9	397	63	6	9	7	8	17	16	13	13	15	24
10	213	50	20	14	11	6	11	8	7	NR	8	NR

[NR=not ranked]

Table 11: Consistency of rank position for supermarkets

Overall rank	Addnum	Count	FY1	FY2	FY3	FY4	FY5	FY6	FY7	FY8	FY9	FY10
			Annual rank									
1	170	328	5	7	8	3	3	1	2	3	2	3
2	244	311	3	3	2	9	8	4	1	1	4	2
3	599	297	1	1	6	4	7	5	3	6	6	9
4	1323	225	6	6	11	8	6	3	4	7	3	6
5	1655	221	4	4	5	5	5	7	5	5	9	4
6	807	184	2	5	1	2	2	9	12	20	28	NR
7	1615	179	8	11	3	1	1	2	NR	NR	NR	NR
8	1637	168	NR	NR	NR	NR	NR	NR	NR	2	1	1
9	1683	168	9	2	12	12	12	10	8	4	5	7
10	1965	137	11	8	4	6	15	8	6	10	14	11
11	1315	119	10	9	10	11	10	15	10	9	7	16
12	1614	118	7	10	9	7	4	6	13	38	NR	NR
13	2122	88	0	17	7	10	9	12	7	17	18	15
14	508	87	0	12	14	15	14	14	9	8	10	10
15	750	86	17	20	16	14	11	11	11	11	12	8
16	1697	82	47	22	NR	18	16	23	18	21	8	5

This also highlights a further important issue; that although the findings suggest there *is* a general consistency to risky premises, the nature of facilities is that they can change ownership, chain or not exist for the entirety of the period studied. That these things might (or will) affect how much crime they experience and thus whether or not they are identified as risky does not undermine the general pattern of consistency found in the categories interrogated, but it does make such patterns harder to establish and also affects the validity of using long-term data to rank problematic premises, as these 'overall' ranks may not be the same as the current (more recent) rank positions.

With this caveat in mind, for the categories selected, there was a final attempt to assess consistency through applying the risky facility definition for each of the ten financial years studied, to see if the overall risky facilities met the required criteria for each year as well.[10] There was greater variability here, but again the general pattern was that many, but certainly not all, premises were risky over a number of years. For example, for *educational*, the top ranked address overall was risky for all ten years, the second ranked was risky for nine years; the third for eight years; and so on as shown in table 12. For *supermarket*, this consistency of achieving risky facility status followed a somewhat similar pattern (with number of years identified as risky being higher for higher ranked premises) but no addresses were risky for more than seven years (with two achieving this), and most were only risky for five or less, again as shown in table 12. This shows that for supermarkets, although high ranking premises tend to remain relatively high ranking, riskiness is generally less consistent than for educational establishments. As specific addresses are explored more in later chapters, this is not discussed further here.

Table 12: Number of years that overall risky facilities are classed as risky (example categories)

Educational			**Supermarket**		
Addnum	Overall rank	Number of years a risky facility (by RF_F1 definition)	Addnum	Overall rank	Number of years a risky facility (by RF_F1 definition)
2266	1	10	244	1	7
965	2	9	1655	2	5
2660	3	8	1323	3	4
2549	4	8	599	4	5
847	5	6	2122	5	1
2758	6	5	508	6	None
2764	7	None	170	7	7
2272	8	1	1579	8	1
397	9	None	1315	9	None
213	10	None	1683	10	3
			577	11	None
			1965	12	2
			1614	13	None
			2724	14	None
			807	15	5
			1697	16	1

[10] In order to present these results as part of this chapter, I here pre-empt the material yet to be covered. The issue of defining and selecting risky facilities is covered in the next chapter, 8, where a favoured selection method is identified and applied to the remaining analyses. Unfortunately, this means the reader will here simply have to accept that a method was chosen, as yet to be described.

7.3.4.3 Summary of persistency and consistency results

It can be concluded from this analysis, that for the majority of facility types risky facilities (addresses) tend to stay risky, or at least in the top ranks, over time and non-risky facilities tend to remain non-risky. This, along with the consideration of the persistency of the risky facility distribution, suggests that the risky facility phenomenon is not a short-lived artefact of the data. Not only does the pattern of some premises contributing a disproportionate amount of crime continue over a relatively long period, but there is a tendency for this contribution to come from the same addresses. Possible reasons for the differences seen in patterns of (particularly) consistency, and the implications of these, are considered in the discussion below.

7.3.5 Nature of distribution

The approach taken to determine the goodness of fit of a power-law distribution to the sample data, and compare this with an alternative (log-normal) distribution, has been set out in the data preparation and analysis section, above. Prior to running these tests, the complementary CDFs were plotted, using the poweRlaw package, alongside fit lines based on the power-law, log-normal and Poisson distributions. The plot for all facility crime, is shown in figure 3, where it can be seen that the Poisson distribution has not been displayed, whilst the power-law and log-normal distributions are a reasonable fit, though less so for the tail.

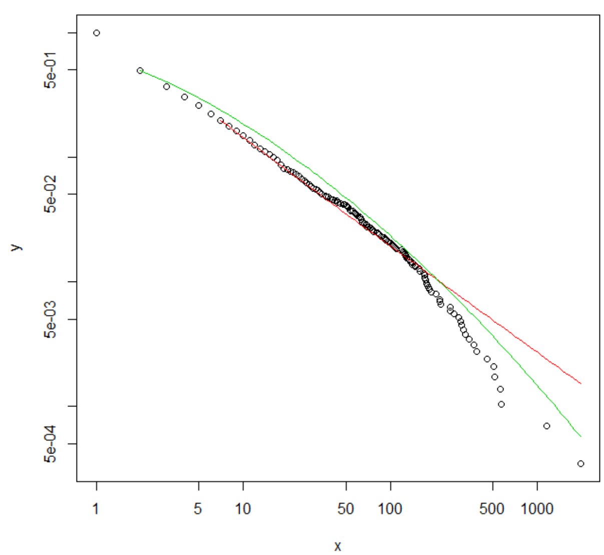


Figure 3: fit lines for all facility crime

The two parts of the statistical tests relate to two sets of hypotheses. Firstly, with regards to the goodness of fit of the power-law distribution alone these are:

H_1: the sample data are not generated from a power-law distribution
H_0: the sample data are generated from a power-law distribution

As set out, the K-S test is used to determine goodness of fit, but the p-value results from a 5000-iteration bootstrap function using synthetic data drawn from the same model. When the p-value is smaller than the chosen significance level (0.1), the null hypothesis is rejected and the conclusion drawn that the data are not taken from a power-law distribution. For all recorded facility crime, this p-value (bs_p$p using the notation from R to avoid confusion between different p-values referred to in this section) was 0.0672. Thus the null hypothesis is rejected (though it is noted that the more usual 0.05 level of significance is not reached). From this first stage, then, it would appear that when looked at together all facility crime does not conform to a power-law distribution.

The second part of the test is to compare to the alternative, log-normal, distribution. As set out in the data preparation and analysis section, above, the sign of the resulting test statistic indicates which distribution is a better fit, so that in this case, a negative statistic implies the log-normal distribution is likely to be a better fit than the power-law distribution (and vice versa). However, because of the possibility of fluctuations in outputs close to 0, the Vuong method is used to produce a p-value that allows us to determine if the observed sign is statistically significant. Therefore, if the 0.1 significance level is not achieved, even with a negative test statistic, we cannot reject the null hypothesis. So, here:

H_1: one of the two distributions is closer to the true distribution
H_0: both distributions are equally far from the true distribution

Although arguably there was no need to complete this second part of the analysis if the power-law form had already been rejected, for completeness this was done for all categories tested. For all facility crime the resulting R was -2.010 (negative, suggesting the log-normal distribution was likely to be a better fit) and the p-value (2-tailed) was 0.044. Thus it was concluded that the observed sign of R was statistically significant and the null hypothesis was rejected. That is to say, for all facility crime, a log-normal distribution was closer to the true distribution than was a power-law distribution.

The analysis was repeated for each facility category, and then again for all facility crime and each facility category for: acquisitive crime only, violent crime only, daylight offending only, and darkness offending only. With caution in interpretation due to the small sample sizes in a number of these cases, the results are shown (separated into three tables for ease of presentation) in tables 13, 14 and 15.

Table 13: Goodness of fit to power-law distribution and comparison with log normal distribution

Facility	N	x_{min}	bs_p$p	Log likelihood ratio [*R*]	p (2 sided)
Convenience	255	1	***<0.001***	-11.594	***<0.001***
Educational	205	2	*0.098*	-1.229	0.219
Financial	146	1	***0.007***	-4.406	***<0.001***
Healthcare*	171	3	0.300	-0.268	0.789
Hotel*	119	2	0.186	-0.580	0.562
Leisure*	40	6	0.211	-0.630	0.529
Licensed premises	115	1	0.167	-5.290	***<0.001***
Petrol	205	1	*0.096*	-3.816	***<0.001***
Pharmacy*	48	1	0.309	-2.096	0.036
Religious	249	3	0.294	-2.045	**0.041**
Restaurant*	290	17	0.306	-0.014	0.989
Supermarket*	164	2	0.205	-1.409	0.159
Retail other*	1498	2	0.141	-0.440	0.660
All	2905	7	*0.067*	-2.010	**0.044**

*indicates power-law cannot be rejected

Table 14: Goodness of fit to power-law distribution and comparison with log normal distribution (crime type)

Facility	Subset	N	x_{min}	bs_p$p	Log likelihood ratio [*R*]	p (2 sided)
Convenience	Acquisitive	216	1	***0.001***	-4.654	***<u><0.001</u>***
	Violent	60	1	0.276	-3.534	***<u><0.001</u>***
Educational	Acquisitive*	127	2	0.633	-0.949	0.342
	Violent*	84	7	0.680	-0.484	0.628
Financial	Acquisitive	142	1	***0.008***	-4.552	***<u><0.001</u>***
	Violent	5	n/a	n/a	n/a	n/a
Healthcare	Acquisitive*	148	3	0.206	-0.286	0.775
	Violent*	29	1	0.680	-1.315	0.189
Hotel	Acquisitive*	48	17	0.311	-0.283	0.777
	Violent*	28	1	0.680	-1.348	0.178
Leisure	Acquisitive*	30	4	0.267	-1.110	0.267
	Violent	15	n/a	n/a	n/a	n/a
Licensed premises	Acquisitive	87	1	0.104	-4.301	***<u><0.001</u>***
	Violent*	46	3	0.102	-1.464	0.143
Petrol	Acquisitive	195	2	0.249	-2.047	**0.041**
	Violent	24	1	0.102	-2.432	**0.015**
Pharmacy	Acquisitive	47	1	0.295	-2.037	**0.042**
	Violent	2	n/a	n/a	n/a	n/a
Religious	Acquisitive	215	3	0.184	-1.653	*0.098*
	Violent	14	n/a	n/a	n/a	n/a
Restaurant	Acquisitive	267	7	***0.007***	-1.610	0.107
	Violent*	45	2	0.102	-1.255	0.210
Supermarket	Acquisitive*	155	2	0.171	-1.489	0.136
	Violent*	29	2	0.289	-0.953	0.341
Retail other	Acquisitive*	1353	3	0.172	<u>0.028</u>	0.977
	Violent*	67	2	0.289	-0.325	0.745
All	Acquisitive	2470	4	0.154	-2.148	**0.032**
	Violent	404	2	*0.087*	-1.435	0.151

n/a = sample deemed too small to calculate/run test

Table 15: Goodness of fit to power-law distribution and comparison with log-normal distribution (time of day)

Facility	Subset	N	x_{min}	bs_p$p	Log likelihood ratio [*R*]	p (2 sided)
Convenience	Daylight	201	1	***0.001***	-4.579	***<0.001***
	Darkness	164	14	**0.034**	-1.480	0.139
Educational	Daylight	167	2	**0.021**	-1.427	0.153
	Darkness*	81	1	0.401	-1.604	0.109
Financial	Daylight	122	1	0.232	-4.847	***<0.001***
	Darkness*	72	10	0.421	-0.627	0.531
Healthcare	Daylight*	144	2	0.326	-0.264	0.792
	Darkness	62	1	**0.019**	-1.560	0.119
Hotel	Daylight*	63	10	0.375	-0.683	0.495
	Darkness*	74	2	0.119	-0.426	0.670
Leisure	Daylight*	35	7	0.830	-0.034	0.973
	Darkness*	27	2	0.416	-1.074	0.283
Licensed premises	Daylight*	66	5	0.674	-0.608	0.543
	Darkness*	92	14	0.724	-0.242	0.809
Petrol	Daylight	180	2	0.343	-1.947	*0.052*
	Darkness	126	1	0.242	-3.678	***<0.001***
Pharmacy	Daylight	43	1	0.234	-1.967	**0.049**
	Darkness*	28	9	0.967	-0.043	0.966
Religious	Daylight*	208	4	0.744	-0.987	0.324
	Darkness*	144	2	0.536	-0.997	0.319
Restaurant	Daylight*	218	5	0.696	-0.451	0.652
	Darkness*	231	5	0.521	-0.901	0.367
Supermarket	Daylight	114	2	0.130	-1.769	*0.077*
	Darkness*	102	1	0.403	-1.818	0.069
Retail other	Daylight*	1174	2	0.145	-0.353	0.724
	Darkness*	732	2	0.352	0.420	0.687
All	Daylight*	2323	4	0.252	-1.128	0.259
	Darkness	1602	2	***0.007***	-3.825	***<0.001***

For categories of facility, the goodness of fit test resulted in a rejection of the null hypothesis for *convenience*, *educational*, *financial* and *petrol.* For these four types of premises, it was concluded that the data were not generated from a power-law distribution. When the second test was carried out, comparing the power-law distribution to the alternative log-normal distribution, it was concluded that the log-normal was a better fit to the data than the power-law distribution for *convenience*, *financial*, *licensed premises*, *petrol* and *religious*. Taken with the results of the first test, this led to the conclusion that the data were not distributed according to the power-law form for a number of facility types (as here listed), but that the power-law distribution could not be rejected for some others (identified by * in the relevant tables): *healthcare*, *hotel*, *leisure*, *pharmacy*, *restaurant*, *supermarket* and *retail other.* The sample sizes overall are relatively small compared to the type of data often analysed in this manner. In particular, it is debatable whether there are sufficient data points for *leisure* and *pharmacy* to be able to achieve even the relatively generous level of significance used here. Therefore, the conclusion that the power-law distribution cannot be ruled out is a conservative one, and it should not be seen - particularly in the cases with very small samples - as evidence of support that facility crime follows a power-law distribution, only that it is feasible that it may do.

Adopting the same approach, but considering crime separated by offence type, the power-law distribution was rejected after the first test for all facility crime (violent crime), *convenience* (acquisitive crime), *financial* (acquisitive) and *restaurant* (acquisitive) and the log-normal distribution was deemed to be a better fit (from the second test) for all facility crime (acquisitive), *convenience* (acquisitive and violent), *financial* (acquisitive), *licensed premises* (acquisitive), *petrol* (acquisitive and violent), *pharmacy* (acquisitive) and *religious* (acquisitive). Therefore, after both tests, when the data were separated into acquisitive crime only and violent crime only, the power-law distribution could not be ruled out for acquisitive crime committed in educational establishments, healthcare facilities, hotels, leisure facilities, supermarkets, and other types of retail premises, nor could it be ruled out for violent crime occurring in educational establishments, healthcare facilities, hotels, licensed premises, restaurants, supermarkets and other retail premises. It is interesting to note that these results suggest that there may be different patterns of distribution for facility crime, when it is separated into different crime types. It is recognised, again, that in some cases sample size will have affected the ability to achieve statistical significance (thus although the power-law distribution has not been rejected, this might be not because it is a reasonable fit, but because there are not sufficient data points to be able to 'disprove' this). However, it certainly seems to be the case that facility crime may be differently distributed in different categories of facility and for different types of crime.

Turning to crime separated by whether it was recorded as occurring during hours of daylight or hours of darkness, the power-law distribution was rejected after the first test for all facility crime (darkness only), *convenience* (daylight and darkness), *educational* (daylight) and *healthcare* (darkness). The log-normal distribution was deemed to be a better fit (test two) for all facility crime (darkness only, again), *financial* (daylight), *petrol* (daylight and darkness), *pharmacy* (daylight) and *supermarket* (daylight). Therefore, after both tests the power-law distribution could not be ruled out for daylight offending committed in healthcare premises, hotels, leisure facilities, licensed premises, religious establishments, restaurants, or *other retail*. This distribution could also not be ruled out for darkness offending committed in educational establishments, financial premises, hotels, leisure facilities, licensed premises, pharmacies, religious establishments, restaurants, supermarkets, and other retail premises.

It is worth briefly noting that in many cases the x_{min} value is small, usually in the range one to seven (though this must, of course, be considered relative to the sample size in question). This suggests that not a lot of the data are discarded in the calculations, and that when the data are fitted to a power-law distribution, in most cases this encompasses a significant portion of the dataset, not just a small portion of the tail.

It is particularly interesting that regardless of the ways the data have been separated here, the power-law distribution was always rejected for convenience stores and petrol stations, whilst it was never ruled out for hotels or those premises categorised as *retail other* (and possibly leisure, but the sample is very small, which will likely have biased the estimation of the scaling parameter, as well as making it difficult to achieve statistical significance in any of the tests). Other facility types that have featured significantly throughout the research into facility-crime concentrations were educational establishments, for which the power-law distribution was nearly always ruled out (excepting acquisitive and darkness crime), and supermarkets for which the power-law could not be ruled out for all but daylight only offending.

Considering research question 3.1: For each class of facility, does the concentration of crime best fit a proposed power-law distribution? as far as can be concluded from the data available for analysis, it is not possible to completely rule out that facility crime samples might be taken from a power-law distribution. However, for certain facilities and certain types or times of crime, it can be said that a power-law distribution is not the best fit, and that in these cases, a log-normal distribution is more plausible. However, this does not mean that distributions of other forms may not yet be a better fit. The implications of these findings will be discussed, below.

7.4 Conclusion

The results presented above demonstrated that the risky facility distribution holds for facility crime in the study area. It also holds for most, arguably all, facility types considered. Overall this pattern tends to persist over time, and though there is some variation in rank, the same premises appear to be risky from one year to the next. Risky facility distributions were also apparent when considered by crime type and by daylight and darkness offending, though some of the conclusions drawn were hampered by the small sample sizes.

The research also shows that the distribution of crime at some facilities does not follow a power-law distribution, but for others, this cannot be rejected. The implications of these finding are discussed more in chapter ten, in light of further contextual findings presented across the next two chapters.

Chapter 8: Study 2: Empirically identifying risky facilities

8.1 Introduction

This chapter focuses on addressing research question 4: Can risky facilities be defined empirically?

In chapter 7, it was established that in the study sample, facility crime was concentrated in a relatively small proportion of premises. The extent of this concentration varied by facility type, but the shape of the distribution could always (when the sample size was large enough) be described, as anticipated, as a 'J-curve'. It was concluded that the results were consistent with the concept of risky facilities and the existing research findings. They also, to some extent, supported the idea of a universal law of crime concentration at place. Although the bandwidth required to cover the whole range of facility types was larger than might have been expected, it was small for many types of facility and when all facility crime was considered together. In addition, it was established that some of the distributions inspected (by facility type, crime type, or time of day) did not follow a power-law distribution, but for others it was not possible to conclude they did not. Overall, there was an inequality of distribution of offences across premises within each given facility category, which could be described as heavy-tailed, though the extent of this concentration in a small number of problematic premises varied by category.

The results so far, then, provide further evidence of the existence and ubiquity of the risky facility phenomenon, whilst showing there is variability in the extent to which offences are unequally distributed across premises and that the nature of this distribution may also vary, so that for at least some methods of categorisation, this does not follow a power-law distribution. Nevertheless, risky facilities clearly exist within the current dataset, but which premises are these? A particular issue that remains under-explored in the risky facility literature is that there is no agreed upon, empirical definition or 'cut off' point for what is 'risky' and what is not. As noted in the literature review, this is characteristic of much work in this discipline, with theorists and researchers tending to employ arbitrary cut points and *ad hoc* methods for categorising places as high crime, hot or risky. I have already explained why I believe this to be problematic. This chapter sets out to consider a number of possible ways that risky facilities themselves could be defined, focusing on quantitative methods, before proposing one of these as the favoured definition, such that when the term is used (by academics or practitioners) there would be greater consistency and comparability. Thus this chapter addresses research questions 4.1, what quantitative methods could be used to empirically define risky facilities and what are the implications of using these? and 4.2, which quantitative method is most appropriate?

In addition, the appropriateness of using police recorded crime data to empirically identify risky facilities is further considered, and critiqued, by exploring serving police officers' perceptions of risky facilities (in their PD) and comparing these to the premises categorised as risky using the preferred quantitative criteria. The types of facility identified, their locations, and some of the explanations proffered for their problematic status are also considered, to provide further insight into how risky facilities are 'recognised' and so labelled. In this way, research question 4.3 (Where do operational police officers perceive risky facilities to occur and how does this compare to quantitative identification from recorded crime data?) is addressed. Such considerations as they relate to journeys-to-crime and sources of offender residences are considered in chapter 9.

The current chapter proceeds in a slightly different fashion, given the extent of discursive and qualitative analyses involved. Firstly, the methods of data preparation and analysis are presented. Next, the findings and discussion regarding the assessment of methods for identifying risky facilities are presented. Following this, the chapter moves to the comparison of identifying problematic premises through recorded crime data versus police perceptions and sets out the findings from the thematic analysis of interviews, and map annotation, with police officers, drawing also, where appropriate, on data collected through observations in the field. Again, the discussion is presented alongside the findings. Finally, the chapter draws both approaches together with a summarising discussion during which the implications for our understanding of risky facilities, how such premises ought to be identified, and how consideration of qualitative data might add to these endeavours are also considered. Finally, a conclusion, reflecting back on the research questions, is offered.

8.2 Data preparation and analysis

The analyses relating to empirical methods for categorising premises as risky, or not, were carried out on the recorded crime dataset already described. No further coding was necessary, therefore the first two parts of this section of the chapter simply set out and justify the approaches taken and the tests used.

The analyses of the qualitative data, has already been presented in general terms in the methodology chapter (6). Specific considerations with regards to the content analysis coding, the themes identified, use of the annotated maps and the approach to comparison with the empirical data are set out below.

8.2.1 Police recorded crime data

In order to consider defining risky facilities empirically, a number of quantitative methods were identified and a selection of these applied to the study data. It was always my intention that the final selected definition be one that would be useful to academics, but could also be adopted by practitioners, such as police analysts. The rule of parsimony was also applied; it was never the aim to identify the most mathematically defensible approach (which would likely involve the point of deviation from some identified distribution), but at the same time it was important that the method should be empirically justifiable and appropriate, whilst holding to the conceptual notion that these premises contribute a substantially disproportionate amount of crime. To be included in the analysis the method had to be replicable (that is to say, it could not be a method that would only work on certain types of data, although the J-curve form really ought to be present if risky facilities are going to be identified), accessible (in terms of the data needed to calculate it, interpreted here as calculable from recorded crime data), and simple (in terms of the statistical knowledge required). My choices for inclusion were inspired by approaches used in the literature and methods that, from my own experience, tend to be used in the real world when identifying prolific or problematic places and people.

The final selection of methods for analysis was:

- Top number of premises (Method A)
- Top proportion of premises (Method B)
- Top crime count (Method C)
- Top cumulative crime contribution (Method D)
- Top fixed proportion of offences (Method E)
- Top variable proportion of offences (Method F)
- Tail, visually selected (Method G)

Most of these were also broken down into sub-methods, using different parameters, cut-offs or multipliers. Each of these methods was applied to all facilities as well as each different facility category to calculate which premises met the criteria of risky facility using that approach. This allowed a comparison across all methods on the number of premises labelled as risky facilities and the percentage of premises this equated to, as well as the number and percentage of crimes that these ‘risky facilities’ contributed. This comparison was used to help assess the practical application, as well as the ability of the approach to produce a meaningful and manageable distinction between selected (risky) and unselected (non-risky) facilities.

When the method involved selecting a proportion or contribution of, or over, a given amount, then the following approach was used to determine the exact cut off. Throughout, two decimal places were used. Method B required the cut-off to be based on a set proportion of premises. As this could only be a whole number, the percentage was always rounded *down* to select the final address to be included. For example, for convenience stores, 51 premises was exactly 20% of the population, therefore the top 51 premises were categorised as risky. However, for hotels, 23 addresses equated to 19.33% of the population, but 24 addresses took this proportion over 20%, therefore, only the top 23 were selected.

For those methods that considered the proportion of a crime contribution, such as Method D, a slightly different approach was taken. As these required, for example, risky facilities to be those that contributed 80% of crime, the last premises included as risky was that which took the cumulative crime contribution to 80.00% or *more* (so even if the previous premises took the contribution to 79.99%, the cut-off point would be at the next premises). Finally, for Methods E and F, a greater than or equal to cut off was used, so all facilities that recorded a crime count of greater than or equal to that set out in the sub-method would be classed as risky (again, calculated to two decimal places).

From a consideration of the features and performance of each method and sub-method, the most favourable approach was selected. I then applied this approach throughout the rest of the study. Overall the determination and assessment of possible quantitative criteria for identifying risky (versus non-risky) facilities, and the final selection, address research questions 4.1 and 4.2.

8.2.2 Police interview data, annotated maps and observations

The third research sub-question was dealt with using a somewhat different approach, by comparing those premises identified as risky using the favoured method determined from the above, with those identified through interviews with serving police officers (and consideration of their map annotations). By introducing these additional data (which are supplemented with analysis based on thick description of my field notes from observations and informal conversations), the match between the two sources could be measured. A deviation between the two could be interpreted as a failure of the police recorded crime data to accurately capture ‘crime’ problems, a failure of the categorisation technique to identify those places that truly are problematic, or a mis-perception (resulting in a false positive or a false negative) of the police officers. Therefore, the qualitative analysis is not a way of validating the proposed quantitative selection technique, but some weight is added if the premises selected using both methods overlap significantly.

The interview (and annotated map) data were analysed using two approaches. Firstly, a content analysis was carried out. This was a mixture of both quantitative and qualitative content analysis. Both of these approaches are critically set out in the methodology chapter, above (6). The quantitative content analysis involved identifying all the specific premises that were mentioned across the six interviews (when these could be attributed to a particular address/named facility). These premises were then categorised according to my set of facility types and the addresses were searched for in the police recorded crime dataset, allowing each to be recorded as either present or absent (in the dataset entirely) as well as whether they were identified as risky facilities using my preferred method. All but one of the premises were found in the recorded crime dataset (which means that there was at least one recorded crime at all the other addresses in the ten-year period studied). The crime count for each facility (by category and overall) was also noted, as was the ranking position relative to all other premises within that category (and the full list of overall risky facilities). These data were then used to assess the extent of the overlap between risky facilities identified from the police recorded crime data and those identified by the police interviewees.

In addition, the annotated maps produced by the respondents were compared with maps produced in the ArcGIS(R) mapping software[11] to identify any notable similarities and differences by location. Qualitatively, the types of premises identified and the locations in which those identified were situated were also considered as part of this comparison, in an effort to find explanations for any discrepancies. In addition to this, relevant issues from the thematic analysis of the interviews (which is explained in chapter 6) were used to help interpret any patterns identified. Where it provided further insight or support to proposed explanations, the analysis of my field study notes (using thick description and as also set out in chapter 6) was also referred to. The qualitatively collected data were also analysed in an effort to explore *how* police recognise and identify problematic premises. Thus research question 4.3 is addressed using a true mixed methods approach.

Having set out the data preparation and approach to analysis for this part of the study, this

[11] See chapter 9 data preparation and analysis for an explanation of how these maps were produced and used.

chapter now turns to a presentation and discussion of the findings.

8.3 Findings and discussion

The findings from the analyses set out above are now considered and discussed in turn. Firstly, the comparison and selection of quantitative methods for identifying premises that should be categorised as risky facilities are presented, followed by a consideration of the qualitatively collected data, with particular reference to comparison of identified risky facilities and an initial exploration of the types of - and reasons for - selection of risky facilities by serving police officers.

8.3.1 Identifying RF cut-off

8.3.1.1 Initial assessment of methods

The approach taken, and the (loose) parameters for inclusion, have been set out in the data preparation and analysis section above. The methods, sub-methods and a description are set out in table 16, whilst table 17 assesses how each method performs against a selection of key, initial attributes (that is to say, before considering the outputs of applying the methods to the sample dataset). Though most of these criteria are ‘advantages’, this is sometimes open to interpretation (as what may be an advantage in some circumstances could be a weakness in others). This is particularly the case for *Accounts for underlying population size* (marked, therefore, with * in table 17, below). Some methods are marked against certain features tentatively, therein indicated by [?], as whether they possess this attribute may be arguable and/or they may only partially possess it.

As can be seen, a range of approaches were selected, and sub-methods were used for most so that different cut-off points or multipliers could be tested. As stated in the data preparation and analysis section, the choice of methods reflected ideas taken from the literature, from my experience in enforcement practice, and my own ideas for how we might capture this disproportionate contribution. The methods selected obviously do not cover every possible approach, and the sub-methods are clearly not exhaustive. The choice, for example, of 3%, 5% or 10% crime contribution (Method E), could have also, or instead, been 1%, 2% or 7%. It was always going to be the case that decisions needed to be made to keep the comparison manageable, and this obviously resulted in some of the arbitrary decision-making for which I have criticised much of the extant literature. However, this is somewhat ameliorated by the fact this was done in order to perform an assessment of different sub-methods and the implications of employing different parameters. That is, the *ad hoc* decisions were made with the best of intentions to reduce how much this is done in the future.

Table 16: Models of selection

Code	Method	Sub-method	Description
A	Top number of premises	(1) Top 5 premises (2) Top 10 premises	Premises ranked by crime contribution and risky facilities are those in the top x positions
B	Top proportion of premises	(1) Top 10% of premises (2) Top 20% of premises	Premises ranked by crime contribution and risky facilities are the top x%
C	Top crime count	n/a	Premises ranked by crime contribution and risky facilities are those that contribute the largest amount of crime. Break point determined by inspection of general crime distribution (i.e. cut off when crime count no longer seen as large)
D	Top cumulative crime contribution	(1) Contributes 80% of crime (2) Contributes 70% of crime	Premises ranked by crime count and cumulative crime contribution calculated. Risky facilities are those that cumulatively contribute
E	Top fixed proportion of offences	(1) Premises that contribute >=3% of crime (2) Premises that contribute >=5% of crime (3) Premises that contribute >=10% of crime	Risky facilities are those that contribute more than x% of crime each
F	Top variable proportion of offences	(1) Premises that contribute >= 3x mean crime count (2) Premises that contribute >= 2x mean crime count (3) Premises that contribute >= 4x mean crime count	Risky facilities are those contributing a given mutiple of the mean crime count
G	Tail, visually selected	n/a	Risky facilities are all those premises to the left (in the tail) of the J-curve. Break point is where the curve is deemed to 'flatten out' upon visual (not mathematical) inspection

Table 17: Assessment of methods against criteria

	A Top number of premises	**B** Top proportion of premises	**C** Top crime count	**D** Top cumulative contribution	**E** Top fixed proportion of offences	**F** Top variable proportion of offences	**G** Tail, visually selected
Simple	✓	✓	?	✓	✓	?	
Accounts for underlying population size*		✓					
Does not matter what underlying population size is	✓		✓	✓	?	✓	
Accounts for underlying crime contribution			✓	✓	✓	✓	✓
Reduced likelihood of producing 'absurd' or impractical results		?	✓			✓	✓
Is less arbitrary/provides justification for selected cut-off		?	?	?		✓	?
Reduced likelihood of (mis-)identifying risky facilities in absence of a J-curve type distribution			✓		?	✓	✓
Considers the (whole) distribution			?			?	✓
Objective	✓	✓		✓	✓	✓	
Not affected by decision on inclusion/ exclusion of zero crime facilities	✓		✓	✓	✓		?

Before considering the features and outputs of the different approaches it is worth noting the influence of the key literature on the methods included. Methods B and D both include nods to the 80-20 rule. In addition, Method D, based on the cumulative percentage contribution of crime would also include Weisburd's approach. However, there are no sub-methods directly taken from this (which would be a 25% contribution, and a 50% contribution) because this part of the analysis was completed before the proposed law was published (it being necessary to settle upon a favoured selection option so that risky and non-risky facilities could be compared in subsequent parts of the study). The outputs produced using a 25% and 50% cumulative crime contribution as the cut-off have, however, been considered *ex post facto*, and are included below, after consideration of the original methods.

The list of possible 'advantages' and the methods for which they are present are now considered. These are summarised in table 17, above.

Simple

As can be seen in table 17, most of the methods are marked as being 'simple'. This is not unsurprising, as it was one of the factors for choice of methods in the first place. Methods C and F are both marked tentatively here, for different reasons. For Method C, this is because the approach involves looking at the premises in ranked crime count order (largest first) and making a subjective decision *in each individual case*, as to what is a sensible break point between those premises that are contributing a large amount of crime (risky) and those that are not. Thus, although this is relatively simple to do, there is no guidance as to how that decision should be made, what should be considered a large proportion, or how long to keep moving down the list if there is no natural 'break' in the crime contribution. Method F, on the other hand, is objective and prescriptive as to where the cut-off should be, but I have somewhat harshly considered it to only partially meet the 'simple' criteria because calculating the mean is marginally more difficult than some of the other approaches. Method G does not meet the 'simple' criteria, as it requires the ability to produced ranked crime count plots (J-curves) and then to (visually) determine what is the end of the 'heavy-tail' part of the distribution. This, therefore, requires more data processing than the other approaches, as well as requiring similar judgement as Method C (with which it overlaps, though the cut-off assessment in Method C is done through looking at the crime counts themselves, whilst in G it is done by looking at the whole, plotted distribution).

Underlying population

Only Method B accounts for the underlying population (number of premises). It should be noted that this can be interpreted as an advantage or a disadvantage. It is advantageous compared to methods such as A, which takes a fixed number of top ranked premises regardless of how many such facilities there are. When there are very few facilities (which is more likely if the approach is applied to relatively fine categories), taking the top ten could actually be a very large proportion (and include, therefore, a number of false positives). On the other hand, if there are very many facilities, considering only ten of them to be risky may result in many problematic places being discounted. In such cases, then, an approach that accounts to some extent for the size of the underlying population, as B does by considering the top ranked *proportion* of premises, is a good thing.

However, being based only on the size of the underlying population is usually a disadvantage, as it means the crime contribution - which is the fundamentally important feature of risky facilities - is not being used to influence the decision (other than to rank the premises in the first place to decide what is 'top'). Therefore, the criteria 'does not matter what underlying population is', is taken as *more* advantageous, and is a feature of a number of the other approaches (A, C, D and F). It is also considered partially a feature of Method E, though the underlying population does indirectly have an effect, because the smaller the number of facilities, the greater the percentage crime contribution of each is likely to be and vice versa; thus small populations are likely to result in a greater chance of false positives and large populations are more likely to produce false negatives).

Finally, G is not marked against either of the *underlying population* criteria, because although it is not affected by this, small populations make smooth, clear J-curve plots less likely (even if the risky facility distribution is present) thus making it harder to visually identify the point where the heavy tail ends.

Underlying crime contribution

Those methods that are based on a number or proportion of premises (A and B), do not account for the underlying crime contribution other than, as explained above, for the purposes of ranking them. Obviously, this ranking process means that those with the higher crime counts will be the ones most likely selected, but here I am referring to the cut-off point being related to the crime contribution, which for identifying risky facilities is particularly desirable.

All of the remaining methods (C to G) involve a cut-off produced from the crime contribution, be this facilities that contribute greater than: a particular number of offences (C); a particular percentage of offences (E); a particular cumulative proportion (D); a certain multiple of the mean (F); or that contribute a particularly unequal portion of the distribution (G).

Absurd or impractical results

Part of the purpose of assessing the outputs from each method is to determine if some approaches tend to produce results that are 'absurd' or impractical, thus to discount these as suitable. Absurd results would be those where nearly all (or none) of the facilities are identified as risky (even when the underlying distribution is appropriate), or those which include a large proportion of what are clearly false negatives or false positives. In other words, those that do not make sense and are not useful. Impractical results are similar, but here they are problematic from a practical perspective, such as too many facilities to effectively engage with in reduction work. It was possible to predict which methods were more or less likely to produce such results, hence this criterion is included here.

The methods most likely to fall foul of this problem are those that do not account for either the underlying population size or the underlying crime contribution (see above) and those that are not able to account for the underlying distribution (see below). Methods C, F and G therefore perform best against this criteria, and Method B at least will produce a set of facilities that is proportional to the population, which makes more sense than to just look at, say, the top five regardless.

Arbitrariness vs justifiability

One of the key aims of this part of the study is to respond to the criticism that the identification or selection of specific facilities that are 'risky', either to engage in practical work or to carry out further research that seeks to explain their existence or compare them to non-risky facilities, is entirely arbitrary and lacking in established convention. Resolving the latter issue is easier, though it requires sufficient influence in order to engage and encourage researchers (and practitioners) to adopt a particular approach. To some degree, Weisburd's conceptualisation seems to have started to do this, as more recent papers do tend to refer to and use the 25% and 50% parameters (though the applicability of this to selecting risky facilities has not yet been considered). The implications of doing so are discussed further, below. However, the figures of 25% and 50% still seem to be *ad hoc*, as has been seen with most other approaches. Being able to limit this arbitrariness was seen as a significant criterion to assess the selected methods against.

As can be seen from table 17, Methods A and E do not meet this criterion at all. If a suitable percentage contribution could be determined, then Method E is more defensible (and once established could be applied routinely), but that choice of percentage is entirely arbitrary. All the other methods are marked as partially meeting this, except Method F, which I have recorded in the positive. Sub-methods of B and D reference the 80-20 rule, so although the choice of these proportions is arbitrary (and noted in the literature as rarely being this 'exact' proportion in the real world), there is *some* theoretical justification. Method C could be argued to at least be driven by looking at the actual, underlying crime distribution so although the choice of cut-off is subjective and unfixed, it is to some extent based on an empirically driven decision (not just a 'randomly' selected number). Method G, in principle, should perform well here, as it is a more sophisticated form of Method C. However, in practice, because the decision is a subjective, visual one, rather than a mathematically determined one, I felt this could only be considered to partially meet the criteria.

Finally, I have considered Method F slightly generously and recorded it as possessing this feature. A somewhat arbitrary decision has to be made regarding the appropriate multiplier to use, but that can be determined through this assessment. The sub-methods are twice, three times and four times the mean. As the only choice is from two upwards, and as five-times was deemed to be too great to capture sufficient facilities, all feasible multipliers have been considered, thus it could be argued that this is not arbitrary at all. It is also justified in that it clearly references the pattern of risky facility distribution (that some facilities contribute a disproportionate amount of crime).

Reduced chance of identifying risky facilities when inappropriate

This criterion is also very important and relates to the chances of the method identifying premises as risky when the risky facility, J-curve distribution is not present. It was highlighted in the data preparation and analysis section that replicability was important, and that I use this term to mean that the approach ought to be applicable to any set of data. In other words, it could not be an approach only derived from and applicable to the current dataset. I also made the point, however, that unless the J-curve risky facility pattern of disproportionate crime concentration was present, it was not appropriate to be categorising some facilities as risky and others not.

Although it may not be possible to completely design-out the chances of identifying some high crime or top-ranked premises as risky when the distribution of offences across addresses is relatively even, certain approaches make it much less likely, as they use a selection method based on an unequal distribution of crime. These are F, G, C (although this does rely on the selector being able to identify that some crime counts are significantly higher than others) and to some extent E (but only if the percentage cut-off is set high enough and there are sufficient facilities within a given category that to achieve that percentage contribution would be unlikely without a skewed distribution). One of the particular advantages of the Method F approach, is that premises can only be identified as risky if they contribute a disproportionate amount of crime (calculated as *a* x mean). Therefore, if crime were evenly (or almost evenly) distributed, no risky facilities would be identified, which would be an appropriate outcome.

Consideration of the whole distribution

This is another important criterion. Methods that do this are less likely to produce 'odd' results and are also more capable of excluding false positives. Arguably, only Method G truly considers the whole distribution (and the nature of this), but I have also noted that Method C partially does this (as the ranked crime counts are considered, in a similar way, looking for a natural break point) as does Method F, in that it is based on those premises that deviate (by a given multiple) upwards from the mean, that is to say crime counts that are greater than they would be if the number of offences was equally distributed.

Objectivity

Arbitrary decisions can, of course, impact upon whether something is considered objective or not. Here, however, I am using this to mean that when using any given method, the decision on what is or is not a risky facility is done in an objective manner. If it is, then this also means that the method can be deemed (easily) replicable and consistent in application as the decision is not open to subjective interpretation or bias, and the facilities selected as risky would be the same regardless of who carried it out. The two methods that require inspection of the data (thus perform well when considering their ability to account for distribution, crime contribution, and so forth) obviously do not meet this criterion as each individual dataset has to be inspected and a (subjective) decision made. Indeed, the selections I have made using these criteria, may themselves be open to question, and might have been different had they been made by someone else. It would have been useful to seek a second opinion on the break points selected (a kind of inter-rater reliability test). This could have been done simply, by using peers as second raters, or more comprehensively, using a range of different individuals to assess for any differences in approach and decision (for example, students, crime analysts and police officers). However, this was not practicable at the time.

All the other methods are considered objective in their application. That said, there are other decisions that have to be made when analysing facility-crime to identify risky premises, such as whether to include zero crime locations, and what categories of facility to use, which are discussed below.

Impact of zero-crime decision

One of the key decisions that needs to be made when considering the existence of risky facilities, and the associated distribution, is whether to include zero crime premises or not. The decision taken for this study has already been discussed elsewhere, and as there was no choice but to exclude premises that did not appear in the crime data, this is the approach that has had to be applied in this section as well. Further implications of this are discussed shortly. Here, however, the proposed methods are assessed against whether such a decision impacts on the outputs, or not. When it does not, this is seen as advantageous, because researchers and practitioners are free to make the decision in the way that best suits them, but will still be comparing 'like with like' when looking at what is risky, or not.

Methods A, C, D and E score positively here, as the process used results in the same selection of facilities as risky whether there are zero-crime premises included or not. Methods B and F do not meet this criterion. Method B is calculated from the total number of premises (so including those with no crimes will mean that the selected proportion of facilities (e.g. 20%) will equate to a smaller crime contribution, and possibly increase the number of false negatives, compared to how they were calculated here). Method F is calculated from the mean, so again, if zero-crime locations are included, the calculated mean will be lower, thus the proportion of facilities contributing crime counts greater than a given multiple of this mean will be *greater* than when zero-crime locations are excluded.

Finally, Method G is noted to only partly meet this criterion. This is because although which facilities are deemed to be in the heavy-tail will not be affected by the inclusion or exclusion of zero-crime locations, the J-curve plot will present slightly differently, which could affect the steepness of the curve and/or the subjective decision-making process of the researcher.

8.3.1.2 Summary of methods

The different methods selected have a number of positive features, but these vary by approach. None of the methods is positively marked against all the criteria. Method F, followed by Method C are associated with the largest number of advantages, but each criterion is not evenly weighted, so this does not in itself demonstrate that these are the 'best' approaches. That said, Method F, whilst having some disadvantages as highlighted above, does seem to perform well and, crucially, it is the only method that is objective *and* is less likely to mis-identify risky facilities when the appropriate distribution pattern is not apparent. That said, many of the approaches assess well when considered against the range of criteria covered here.

8.3.1.3 Performance of methods

A summary of the performance of each method and sub-method (showing the number and percentage of facilities defined as risky, and the number and percentage of crimes associated with these facilities, alongside reference information relating to the number of offences, number of facilities, highest number of offences attributed to a single facility, and Gini coefficient) is presented in table 18 (a larger version of which is available from the author).

Table 18: Outputs of applying each method of selection

			Method	A		B		C	D		E			F			G
			Sub-method	(1) Top 5	(2) Top 10	(1) Top 10%	(2) Top 20%	None	(1) Contribute 80%	(2) Contribute 70%	(1) Contribute >=3%	(2) Contribute >=5%	(3) Contribute >=10%	(1) Contribute >= 3x mean	(2) Contribute >= 2x mean	(3) Contribute >= 4x mean	None
Facility	Category summary		RF selection														
Convenience	Top crime count	148	N facilities	5	10	25	51	11	53	40	4	0	0	28	45	18	17
	N offences	4276	% facilities	1.96	3.92	9.80	20.00	4.31	20.78	15.69	1.57	0.00	0.00	10.98	17.65	7.06	6.67
	N facilities	255	N offences	672	1237	2345	3371	1340	3428	3009	552	0	0	2505	3185	1913	1845
	Gini co-efficient	0.7368	% offences	15.72	31.34	54.84	78.84	31.34	80.17	70.37	12.91	0.00	0.00	58.58	74.49	44.74	43.15
Educational	Top crime count	508	N facilities	5	10	20	40	10	14	8	8	6	2	10	17	9	9
	N offences	2735	% facilities	2.44	4.88	9.76	19.51	4.88	6.83	3.90	3.90	2.93	0.98	4.88	8.29	4.39	4.39
	N facilities	205	N offences	1541	2047	2345	2538	2047	2190	1934	1934	1719	799	2047	2277	1997	1997
	Gini co-efficient	0.8643	% offences	56.34	74.84	85.74	92.80	74.84	80.07	70.71	70.71	62.85	29.21	74.84	83.25	73.02	73.02
Financial	Top crime count	94	N facilities	5	10	14	29	21	33	23	8	2	0	18	22	11	8
	N offences	1162	% facilities	3.42	6.85	9.59	19.86	14.38	22.60	15.75	5.48	1.37	0.00	12.33	15.07	7.53	5.48
	N facilities	146	N offences	313	510	625	894	790	932	822	442	160	0	726	807	542	442
	Gini co-efficient	0.7004	% offences	26.94	43.89	53.79	76.94	67.99	80.21	70.74	38.04	13.77	0.00	62.48	69.45	46.64	38.04
Healthcare	Top crime count	169	N facilities	5	10	17	34	4	68	42	2	1	1	4	11	2	1
	N offences	524	% facilities	2.92	5.85	9.94	19.88	2.34	39.77	24.56	1.17	0.58	0.58	2.34	6.43	1.17	0.58
	N facilities	171	N offences	217	252	288	344	206	420	368	186	169	169	206	259	186	169
	Gini co-efficient	0.5871	% offences	41.41	48.09	54.96	65.65	39.89	80.15	70.23	35.50	32.25	32.25	39.89	49.43	35.50	32.25
Hotel	Top crime count	70	N facilities	5	10	11	23	6	32	13	7	6	3	7	10	6	8
	N offences	434	% facilities	4.20	8.40	9.24	19.33	5.04	26.89	10.92	5.88	5.04	2.52	5.88	8.40	5.04	6.72
	N facilities	119	N offences	231	295	299	330	254	348	306	270	254	167	270	295	254	278
	Gini co-efficient	0.6708	% offences	53.23	67.97	68.89	76.04	58.53	80.18	70.51	61.75	58.53	38.48	61.75	67.97	58.53	64.06
Leisure	Top crime count	63	N facilities	5	10	4	8	5	12	11	10	5	1	2	5	1	3
	N offences	322	% facilities	12.50	25.00	10.00	20.00	12.50	30.00	27.50	25.00	12.50	2.50	5.00	12.50	2.50	7.50
	N facilities	40	N offences	151	219	131	196	151	260	228	219	151	63	90	151	63	111
	Gini co-efficient	0.5829	% offences	46.89	68.01	40.68	60.87	46.89	80.75	70.81	68.01	46.89	19.57	27.95	46.89	19.57	34.47
Licensed premises	Top crime count	42	N facilities	5	10	11	22	17	35	25	10	2	0	10	21	5	5
	N offences	627	% facilities	4.35	8.70	9.57	19.13	14.78	30.43	21.74	8.70	1.74	0.00	8.70	18.26	4.35	4.35
	N facilities	115	N offences	150	249	265	415	358	505	444	249	76	0	249	405	150	150
	Gini co-efficient	0.6137	% offences	23.92	39.71	42.26	66.19	57.10	80.54	70.81	39.71	12.12	0.00	39.71	64.59	23.92	23.92
Petrol	Top crime count	188	N facilities	5	10	20	40	10	41	27	9	3	0	19	26	15	10
	N offences	2380	% facilities	2.44	4.88	9.76	19.51	4.88	20.00	13.17	4.39	1.46	0.00	9.27	12.68	7.32	4.88
	N facilities	205	N offences	654	1028	1487	1897	1028	1998	1681	958	479	0	1454	1658	1292	1028
	Gini co-efficient	0.7377	% offences	27.48	43.19	62.48	79.71	43.19	80.17	70.63	40.25	20.13	0.00	61.09	69.66	54.29	43.19
Pharmacy	Top crime count	191	N facilities	5	10	4	9	10	11	9	10	8	3	4	10	3	3
	N offences	996	% facilities	10.42	20.83	8.33	18.75	20.83	22.92	18.75	20.83	16.67	6.25	8.33	20.83	6.25	6.25
	N facilities	48	N offences	534	780	474	738	780	804	738	780	693	409	474	780	409	409
	Gini co-efficient	0.7206	% offences	53.61	78.31	47.59	74.10	78.31	80.72	74.10	78.31	69.58	41.06	47.59	78.31	41.06	41.06
Religious	Top crime count	16	N facilities	5	10	25	50	11	126	92	0	0	0	14	32	8	18
	N offences	719	% facilities	2.01	4.02	10.04	20.08	4.42	50.60	36.95	0.00	0.00	0.00	5.62	12.85	3.21	7.23
	N facilities	249	N offences	70	127	241	367	137	577	505	0	0	0	164	283	106	195
	Gini co-efficient	0.4397	% offences	9.74	17.66	33.52	51.04	19.05	80.25	70.24	0.00	0.00	0.00	22.81	39.36	14.74	27.12
Restaurant	Top crime count	66	N facilities	5	10	29	58	15	110	82	1	0	0	18	45	9	5
	N offences	1688	% facilities	1.72	3.45	10.00	20.00	5.17	37.93	28.28	0.34	0.00	0.00	6.21	15.52	3.10	1.72
	N facilities	290	N offences	204	328	663	997	433	1353	1188	66	0	0	487	867	305	204
	Gini co-efficient	0.5586	% offences	12.09	19.43	39.28	59.06	25.65	80.15	70.38	3.91	0.00	0.00	28.85	51.36	18.07	12.09
Supermarket	Top crime count	328	N facilities	5	10	16	32	12	16	12	12	7	0	16	18	15	18
	N offences	3402	% facilities	3.05	6.10	9.76	19.51	7.32	9.76	7.32	7.32	4.27	0.00	9.76	10.98	9.15	10.98
	N facilities	164	N offences	1382	2218	2798	3131	2455	2798	2455	2455	1745	0	2798	2899	2716	2899
	Gini co-efficient	0.8451	% offences	40.62	65.20	82.25	92.03	72.16	82.25	72.16	72.16	51.29	0.00	82.25	85.21	79.84	85.21
Retail other	Top crime count	1936	N facilities	5	10	142	285	13	217	81	5	2	2	46	74	36	7
	N offences	10558	% facilities	0.33	0.67	9.48	19.03	0.87	14.49	5.41	0.33	0.13	0.13	3.07	4.94	2.40	0.47
	N facilities	1498	N offences	4126	5033	7985	8752	5357	8446	7392	4126	2992	2992	6815	7301	6561	4633
	Gini co-efficient	0.7907	% offences	39.08	47.67	75.63	82.89	50.74	80.00	70.01	39.08	28.34	28.34	64.55	69.15	62.14	43.88
All	Top crime count	1968	N facilities	5	10	294	554	59	355	180	2	1	0	157	224	130	4
	N offences	30735	% facilities	0.17	0.34	10.12	19.07	2.03	12.22	6.20	0.07	0.03	0.00	5.40	7.71	4.48	0.14
	N facilities	2905	N offences	4769	6837	23772	26595	15367	24594	21523	3124	1968	0	20863	22565	19911	4259
	Gini co-efficient	0.8144	% offences	15.52	22.24	77.35	86.53	50.00	80.02	70.03	10.16	6.40	0.00	67.88	73.42	64.78	13.86

As can be seen from table 18, within a given facility category, the number and proportion of facilities, and the number and proportion of offences selected using the different methods tends to vary quite substantially. In a way this makes it harder to know which approach is 'best' or more 'accurate' (if indeed such a thing is possible at all). On the other hand, the fact that there are differences in the outputs adds to arguments about why it is important to more consistently identify risky facilities, especially when comparing them to non-risky facilities or across different research studies and locations.

Not every output for every category can be presented and discussed here, therefore in an effort to assess the performance of the methods, they are each briefly summarised, then some approaches are selected for comparison with one another (because they are a similar or notably different approach), and finally overall patterns are drawn out. Discussion of the implications of these ‘results’, cross-cutting issues, and my conclusions on the most appropriate method for empirically defining risky facilities follow.

8.3.1.3.1 Consideration of method outputs

Method A, the top ranked number of premises, obviously consistently selected five (A1) or ten (A2) facilities as risky. The percentage of crime that the top five addresses contributed, for all facilities, was 15.52%. The percentage was generally low across all categories (with figures such as 9.74% for *religious*, 12.09% for *restaurant*, and 15.72% for *convenience*). On the other hand (staying with method A1 for comparability), some categories recorded much higher crime contributions from the top ranked premises, such as *educational* with 56.34% (the greatest proportion recorded for this sub-method), *pharmacy* (53.61%) and *hotel* (53.23%). There were also two further categories (*leisure* and *supermarket*) that were above 40%. Generally, the lower crime contributions tended to be categories where there were less crimes anyway and/or where crime was less concentrated. That said, this was clearly not *always* the case.

Method B involved taking the top 10% or 20% of premises, in an effort to account for underlying population size and, for method B2, as a nod to the 80-20 rule. The outputs from these methods varied widely in terms of the percentage crime contribution this equated to. For example, for B1 (10% of premises) this ranged from proportions as low as in the 30 and 40 per cents up to categories such as *educational* and *supermarket* (85.74% and 82.25%), with *all* and *retail other* not far behind. Those with the lowest Gini coefficients (the less unequal distributions) tended to be those with the lower crime contributions. This was expected, but the outputs help demonstrate the point.

For Method C, I made the choice of break point without looking at the recorded outputs for any of the other methods, so as not to influence my decision. This was done by considering the ranked crime counts and assessing, from the numbers only, where there seemed to be a natural break between high crime contributors and the rest. Reflecting on this process, there were some categories of facility for which this was a fairly easy decision to make, such as *educational*, *petrol* and *pharmacy* where there was a clear difference in the extent of crime contribution between quite problematic facilities and those that were less so. For others, it was more difficult (thus subjective). For example, the decisions for categories such as *religious* and *restaurant* (both of which had lower Gini coefficients than many of the other categories, though still experienced unequal distribution of offences) were fairly arbitrary, as there was no natural break point apparent in the data. This was also the case for *licensed premises*, which had a somewhat larger Gini coefficient, but did have a relatively small number of offences.

Finally, there were categories for which it was difficult to decide where, of a number of places, to set the cut-off. This included occasions when the cut-off point could have been a few premises more or less, because the decrease in crime counts from one to the next was not particularly severe. An example of this was convenience stores for which the selection made was facilities contributing 103 or more offences. The series immediately around this point was: …110, 104, 103, 99, 96, 87…, therefore it is reasonable that the break could have been between 110 and 104, 103 and 99 (as it was), or 96 and 87. In other cases, particularly for *healthcare*, it was difficult to decide because of a substantial outlier (the hospital). Arguably, this should have been selected as the only risky facility as it could clearly be distinguished in the crime counts from the next ranked premises (contributing 169 offences, whilst the 2nd address contributed only 17). However, if this address was ignored, there remained an over-contribution of offences (compared to the remaining premises) for the next few facilities as well, therefore the cut-off was set at four facilities (with crime counts of ten or more). Either decision is hard to defend. Overall, this method's reliance on subjective assessment of a list of ranked numbers made it difficult to apply, and it is highly likely that a different researcher (or a practitioner, who is influenced by somewhat different concerns and constraints, such as what seems like a 'manageable' number) would make a different risky facility selection, at least in terms of premises close to the cut point.

Method D was based on the 80-20 rule, looking for those facilities that cumulatively contributed 80% of the total crime count (sub-method (1)). It was anticipated this might capture too high a number of facilities and include premises that were not in fact risky, which would itself lead to a greater number of premises selected (as more are needed when their contribution is less). Therefore, the second sub-method was to select on premises contributing the lower 70% of the total crime count. The proportion of facilities captured in each category varied, usually in line with the extent of crime concentration, again as expected. So for 80% of crime the greatest proportions of premises needed were 50.60% (*religious*) and 39.77% (*healthcare*), whilst the smallest proportions were 6.83% (*educational*) and 9.76% (*supermarket*). Invariably the smallest percentages of premises were required when the Gini coefficients were larger. This same pattern was apparent for sub-method (2), with obviously smaller percentages of premises (because the required cumulative crime contribution was less). Interestingly, however, *hotel* did reduce more substantially than many others (from the 80% to 70% cut off). Further consideration of the data suggests this was because there were a number of high crime counts for the first few premises (up to nearly a 60% cumulative crime contribution), but by the 70% contribution, offending had dropped to only 3 or fewer crimes per premises (so it took many more addresses before the 80% cut off was reached).

These patterns demonstrate the problem of capturing some premises that are not risky when crime is less unequally distributed, but they also highlight the fact that the extent of this *worsens* the more even the distribution is. Therefore, this approach may be a good one to take when we know we have a strong risky facility-type distribution, but it will perform poorly when this is weaker (and of course when it is not present), as it will continue to select more and more (unproblematic) premises to reach the target cumulative crime contribution. In the most extreme (hypothetical) example, if crime were completely evenly distributed across premises, then the '80% of crime' cut-off would select 80% of premises to define as risky. Whilst all the research evidence points to the fact this would not be seen, it illustrates the major weakness of this approach (without first testing distributions, or having to also develop a minimum Gini coefficient or a resulting percentage of premises for which the results have to be discarded).

Rather than considering the cumulative crime contribution, Method E is based on the proportion of crime contributed by each premises. This has the advantage of not continuing to select premises just to reach some arbitrarily set cut off, but the disadvantage that the selected target percentage is itself arbitrarily set and, depending on the number of premises and overall crime count, may be too low to only select premises that are 'truly' problematic or too high to select any at all. Three different percentages were tested (1) 3%, (2) 5%, and (3) 10%. It is clear from table 18, that for the study sample, 10% is too high a cut off, as it captured no premises when all facility crime was considered, no premises for seven of the facility categories, and selected only 3 or less risky facilities for the rest (*educational*, *healthcare*, *hotel*, *leisure*, *pharmacy* and *retail other*), ranging from 19.57% of crime contributed to 41.06%. There does not appear to be any relationship between how unequal the crime distribution is and how many (if any) facilities meet the selection criteria.

For the lower cut-off percentages, more premises are, naturally, selected, though compared to many of the other approaches, the proportion of facilities identified as risky remains small. This is not necessarily a bad thing, as it may be correct that only a very few premises are indeed risky, but if the 80-20 conceptualisation holds any merit, for a risky facility-type distribution we would expect figures more around the 10-30% of premises and capturing around 70 to 90% of crime (being very generous in interpretation, both ways). Even the 3% cut off is not able to achieve these ranges in many cases (again, I reiterate that this may be specific to the current data set and choice of categories). It is acknowledged that for the highly unequal *educational* and *supermarket* categories the 3% cut off works quite well, but otherwise the categories for which a larger proportion of crime is captured tend to be the ones that have fewer total premises, as anticipated. Nothing illustrates this more than the fact that when all facility crime is considered together (2905 facilities), only two premises (accounting for 10.16% of crime) are identified as risky using this method.

Method F was seen to perform well in the criteria-based assessment and it was noted that it was one of the few (and the only objective method) that considered the underlying distribution. All three of the sub-methods (multipliers of 3x mean (1), 2x (2) and 4x (3)) performed reasonably well in terms of capturing a relatively large proportion of crime (though rarely as much as 80%, with a few notable exceptions), for a relatively small proportion of premises. Obviously using two times the mean encompassed the largest percentage of crime, and four times the mean encompassed the smallest. The categories accounting for the smallest proportions of contributed crime were always those with the smaller Gini coefficients. Sub-method (2) (2x mean) obviously involved a larger percentage of premises (though this was rarely so large as to be thought of as particularly problematic). The other two sub-methods (which often had quite similar results) tended to select a proportion of premises that fell in the middle of those selected across the whole range of methods, with no outputs standing out as being of concern (either too high or too low) with regards the likelihood of including false positive or negatives. That said, three times the mean (and certainly two times the mean) was somewhat more generous in its selection than a number of the other approaches. The sub-methods are compared further below.

The final approach to consider individually is Method G. In order to reduce the likelihood of selection-bias, I performed this first. This was done by printing the J-curve plots, and manually fitting straight lines to the distribution curve (scanned versions can be found in Appendix 4). In most cases, the curve could be separated into three straight line sections (in some it was four) and the cut-off point for risky facilities was taken to be when the line began to deviate from the first straight line. In other words, this was when the steepest part of the tail started to level out. It has been argued above that this has the potential to be the most ‘accurate’ method of selecting risky facilities, but that it is (comparatively) more difficult, and subjective. This was indeed the case. It was relatively simple to draw the lines, but as they were only visual ‘best fits’, they could have been drawn differently, and the choice of where the line started to deviate ‘enough’ to make this a break point could also have been elsewhere. That said, it was arguably easier to identify where the cut-off should be when looking at a plot of the whole distribution, than when looking only at the crime and cumulative crime counts/percentages (as was the case for Method C).

Also apparent from the results, and supported by my reflections, is the fact that this method resulted in the selection of relatively few premises as risky facilities (albeit so did some of the other approaches). Without there being a definitive way of determining what are or are not risky facilities, it is not possible to conclude whether such a small selection of premises is a good, or more accurate, thing. However, carrying out this analysis did highlight the possibility of a category of 'super risky facilities' existing, which would be that very small proportion of premises that exist as substantial outliers when the rest of the data are considered. This is briefly considered, further, below, but it is introduced here because there was some evidence when selecting cut offs from the J-curve charts (and when looking at crime counts for Method C) that this might be the case, at least for some categories. For example, it has already been mentioned that the healthcare category included one premises (i.e. 0.58% of *healthcare* facilities) that accounted for 32.25% of offences (169 offences), whilst the next highest address only contributed 3.24% (17 offences). However, if the top address was excluded, there was still an unequal distribution, accounted for by the next few addresses. This is an extreme example, and it could be argued that it is at least in part the result of 'poor' categorisation choice, or the absence of any meaningful denominators that could be used to produce a rate (e.g. footfall in premises). Categorisation is considered as a general issue further below, and again (as is the denominator issue) in chapter 10, the overall discussion. However, a similar pattern also exists when all facility crime is considered together. Two premises (first a shopping mall, second a Walmart plus the concessions and other outlets sharing that same address)[12] together contributed 10.16% of crime (3124 offences) but only equated to 0.07% of premises. The next highest facility was another mall (that included a Harris Teeter sharing the address and contributing one third of the recorded crime), but this only recorded 570 offences (1.85%). However, as Method F demonstrated, overall 130 facilities recorded crime counts of greater than or equal to four times the mean (and 157 recorded counts of greater than or equal to three times the mean), so to only select those first two 'super risky' premises (issues of whether malls should be considered as single premises aside) would clearly have meant many problematic premises would have been ignored. Yet using Method G, without having seen any of the other results or any contextual information, I selected only four premises as sitting on the steepest line of best fit (as comparison, I selected a cut off at 59 premises when looking at the ranked crime counts for Method C).

All of this demonstrates that it is difficult to select one method as 'the most appropriate', not least because it is unclear what that method should be achieving as we have no existing measure of 'true' risky facilities to test it against, as they have never been empirically defined. Without doing this (through choosing a selection method), we will continue to fall foul of the problems I have

[12] It should be noted that there were multiple facility types at the address, depending on how they were categorised. This means that this same Walmart, when considered as a supermarket, was actually the 16th ranked premises of this type (and just met the criteria of risky facility applied). This demonstrates some of the further problems faced when categorising, or inded not categorising, facilities and also of using police data to attribute crimes to specific addresses.

already set out. Therefore, I now, more briefly, consider a comparison of some of the approaches, to help draw some conclusions.

8.3.1.3.2 Comparison

Methods C and G, which have already been partly contrasted, are worthy of further comparison, because they are quite similar approaches, but carried out in a different way. Looking at the outputs shows quite different patterns. For some categories of facility, the number and proportion of facilities and crimes selected using each was very similar/the same (*educational* and *petrol*), for some it was quite similar (*healthcare* and *hotel*) and for some, one of the methods was slightly higher than the other (*convenience*, *leisure*, *religious*, *supermarket* and *retail other*). When this was the case, either method could be the one that produced the greater number of risky facilities, though it was slightly more common for Method C to identify a few less. Finally, there were some categories where the difference between the two was substantial (*financial*, *licensed premises*, *pharmacy, restaurant* and *all).* Whenever this was the case, Method C identified far more premises as risky than did Method G.

It is interesting to note that those that were the most similar, were often the ones identified as having a clear, natural break point when applying Method C. It is not surprising that this would be reflected in the choice made when looking at the associated J-curve plot for Method G. However, this was not always the case. When Method C resulted in far more premises being selected as risky than Method G, there was no obvious pattern, though this did include some of those categories highlighted as resulting in fairly arbitrary selections using Method C, and some of those where the decision was difficult because the cut off could have been set at more than one place. This is particularly the case for *all.* For Method C, I selected risky facilities as those contributing 100 or more offences (59 addresses, 2.03%), but I could just have easily set this at those contributing the much higher 457 or more offences (from seven premises, the eighth - not selected - having a crime count of 389). Upon comparison, this would have been much more consistent with the selection I made from the J-curve plot (Method G), which encompassed only four facilities (with the crime count cut off, therefore, being 565 offences). Therefore, although difficult to argue for one or the other approach, it seems that Method G might be the more consistent of the two to use.

Though the flaws of Method G have been noted conceptually (such as it being difficult and subjective) and empirically (this analysis suggests that it slightly under-identifies the premises that should be considered risky), as the only method based on the J-curve plot, it is also worth considering how the other methods compare, in terms of the proportion of premises captured. Overall, there was a tendency for the proportions (of premises and associated cumulative crime contributions) to be most similar for Methods A1, E1 or F3 (depending on the category of facility), followed by A2 and F1. However, no method(s) could be identified as consistently producing the most similar outputs to Method G.

Considering the multipliers used for Method F, prior to carrying out the analysis, it was unclear if larger multipliers would be too restrictive in terms of their selection. In comparing across the different sub-methods, arguments could be made for each of them as the strongest performer. Obviously, if the aim was to select premises that contributed the greatest proportion of crime and minimise the likelihood of false negatives, then twice the mean would best achieve this. If the aim was to minimise the chance of false positives, whilst also ensuring a relatively small number of premises were identified (to be more manageable for intervention), then four times the mean would be the more suitable method. Overall, it was felt that both three and four times the mean captured a reasonable proportion of premises and crimes, therefore two times the mean could be discounted as the most suitable option. As stated, selecting on four times the mean seemed to be more often similar to using Method G (though certainly not always) than selecting on three times the mean. However, there was very little to choose between these two multipliers. With that in mind, and considering that Method G was thought to result in slightly too few potentially problematic premises being selected, I made the decision that sub-method (1), three times the mean, was the favoured option. The implications of this decision are discussed further below.

Therefore, for the remainder of the analysis, whenever risky facilities are selected, analysed or compared, these have been identified as being those premises that contribute three times the mean crime count (for that category of facility, or overall for all facility crime). For clarity, the code on which these were selected (RF_F1) is also sometimes used to refer to these.

8.3.1.3.3 The law of crime concentration at places

The patterns of facility crime concentration in relation to Weisburd's proposed law of crime concentration at places has already been considered and discussed in chapter 7. It was noted above that this part of the analysis was completed prior to the publication of this proposal and the growing body of research seeking to test and apply it. Therefore, this method was not included when deciding which was the most appropriate approach for selecting risky facilities.

However, as this method of identifying crime concentration at places is now growing in popularity, I felt it was important to consider the outputs it would produce and compare them to the other approaches already detailed. The proportion of premises selected using the Weisburd method have already been presented in table 5, but the full set of outputs relevant to this chapter are set out in table 19, overleaf (which should be compared with table 18, above).

As this is an iteration of Method D, the same advantages (or disadvantages) apply, but the cut-off points are different. The proportion of crime selected is obviously determined by the parameters, and is set lower than many of the other approaches. In keeping with this using 25% tends to result in a selection of 'risky facilities' at the lower end (considering across all methods) and using 50% tends to select a number around the lower middle. What is particularly important to note, however, is that because this is another cumulative proportion of crime contribution method, it performs less well when there are large outliers (for *healthcare* only the one address is selected using the 25% cut-off but 12 are selected using the 50% measure). When the distribution of crime contribution is less unequal (such as for *religious* and *restaurant*) it also performs less well, having to capture a relatively large proportion of premises (compared to other categories of facility), particularly for the 25% measure. In addition, the 25% cut-off results in very few risky facilities identified when the number of facilities is large, such as for *retail other* (2 premises, 0.13%) and *all* (13 premises, 0.45%).

These criteria, then, perform as well as many of the others, but still possess the flaws discussed above for Method D. This is particularly so when the underlying distribution is less skewed, or does not follow a J-curve, though the lower percentage target (25% or 50% rather than 70% or 80%) does reduce the proportion of false positives that might be included. That said, lower crime contribution cut-offs run the risk of excluding quite problematic premises from the selection (false negatives) and this is particularly likely when this is set at a cumulative contribution of only 25%.

Given the often very small percentages of places (here, facilities) contributing the proportions of crime proposed by Weisburd (2015), it seems that this approach is particularly good for testing and - in particular - *illustrating* the highly concentrated nature of crime at places. It is not, however, the most appropriate way to *select* addresses that should be considered risky (for the purposes of research and intervention), as the cut-off point is a cumulative crime contribution that is too low, and like most other methods considered here, it does not automatically exclude selection of premises when the risky facility distribution is not present.

Table 19: selection using Weisburd (2015) suggested cut-offs (overleaf)

Facility	RF selection	Weisburd Law Contribute 25% of crime	Contribute 50% of crime
Convenience	N facilities	9	22
	% facilities	3.53	8.63
	N offences	1133	2171
	% offences	26.50	50.77
Educational	N facilities	2	5
	% facilities	0.98	2.44
	N offences	799	1541
	% offences	29.21	56.34
Financial	N facilities	5	13
	% facilities	3.42	8.90
	N offences	313	597
	% offences	26.94	51.46
Healthcare	N facilities	1	12
	% facilities	0.58	7.02
	N offences	169	265
	% offences	32.25	50.57
Hotel	N facilities	2	5
	% facilities	1.68	4.20
	N offences	119	231
	% offences	27.42	53.23
Leisure	N facilities	2	6
	% facilities	5.00	15.00
	N offences	90	167
	% offences	27.95	51.86
Licensed premises	N facilities	6	15
	% facilities	5.22	13.04
	N offences	171	328
	% offences	27.27	52.31
Petrol	N facilities	5	13
	% facilities	2.44	6.34
	N offences	654	1193
	% offences	27.48	50.13
Pharmacy	N facilities	2	5
	% facilities	4.17	10.42
	N offences	300	534
	% offences	30.12	53.61
Religious	N facilities	16	49
	% facilities	6.43	19.68
	N offences	180	363
	% offences	25.03	50.49
Restaurant	N facilities	15	44
	% facilities	5.17	15.17
	N offences	433	855
	% offences	25.65	50.65
Supermarket	N facilities	3	7
	% facilities	1.83	4.27
	N offences	936	1745
	% offences	27.51	51.29
Retail other	N facilities	2	13
	% facilities	0.13	0.87
	N offences	2992	5357
	% offences	28.34	50.74
All	N facilities	13	59
	% facilities	0.45	2.03
	N offences	7784	15367
	% offences	25.33	50.00

In seeking to answer research question 4.1, these findings have shown that there are many methods that could be used to determine what premises, within a given class of facility, ought to be considered ‘risky’. A number of these were selected, tested and compared, and it was found that their outputs were often diverse (sometimes substantially so), in terms of the proportion of facilities (and crime contribution) selected. Considering also the problems of minimising possible false positives and negatives, keeping to a relatively simple and objective approach and the need to reduce the chance of facilities being selected in the absence of an underlying unequal distribution, I conclude that the most appropriate approach is Method F. Given there was little to choose between the sub-methods using a multiplier of three times the mean or four times the mean, I ultimately decided to err on the side of inclusivity and selected method (1) - three times the mean - as the most appropriate (thus answering research question 4.2). As noted, this method (coded as RF_F1) has been applied for the remaining analyses.

8.3.1.4 Further issues: Selecting a cut-off

Having decided upon a favoured approach, it is important to consider further cross-cutting issues that relate to selecting risky facilities and limitations of the method chosen.

8.3.1.4.1 Zero-crime facilities and choice of time periods

A key issue that needs to be considered with regards the methods used in this section of the study, is the decision regarding exclusion (or not) of zero-crime facilities. The implications of doing so have already been considered for the methods tested, above, and it is noted that this particularly impacts on Method F, selected as the most appropriate. This is, in fact, its biggest weakness. I was constrained with regards the decision to exclude zero-crime facilities, as the information was not available, or would have been too onerous (and potentially incomplete) to collect. Therefore, RF_F1 was chosen as the most appropriate method for selecting risky facilities based on data that *excluded zero crime facilities*. Had these been included, then using three times the mean as the selection criteria for risky premises would have resulted, as already identified, in a larger proportion of premises being selected. Had this been the case, it may have been more suitable to settle on four-times the mean (sub-method 3) as the most appropriate method. That said, because data on zero-crime facilities has not been collected, the number of these is unknown. Therefore, it has not been possible to calculate and critique the effects their inclusion would have had.

Although it is certainly interesting from a research perspective to consider not only the differences between risky and non-risky facilities, but also premises that have no recorded crime associated with them, from a practitioner perspective this is likely to be a difficult, and unnecessary, task. For some types of establishment, such as licensed premises and schools, the data may be relatively accessible, but it will be less so for other facilities, such as shops. That is not to say that such information is not available, but that obtaining it would be time-consuming and would not really add much that could not be determined from stopping at single crime locations. It would also be much harder, maybe even impossible, to get such data for every premises in an area that could be considered a facility (if one did not wish to just look at particular categories).

This decision on inclusion or not of premises that have no recorded crime attributed to them becomes even more important when we consider the issue of time periods. In the current study, the full ten-year period is used to identify risky facilities. However, what if a shorter (or longer) time period had been used? It was established in chapter 7 that the risky facility pattern for this study area, was relatively persistent over time (comparing yearly) and that there was, overall, a good deal of consistency in the ranking of addresses, at least for most of the facility categories and facility crime overall. Nevertheless, it was also noted that there was some difference in the ordering of premises, and also that for two selected example categories (*educational* and *supermarket*) not all of those identified as risky overall (using the RF_F1 criteria settled upon in this chapter) were risky every year, when this method was applied to the annual (as opposed to overall) data.

Selecting a shorter time period will obviously result in a lower crime count. This in itself should not affect the ability of the selection method to identify risky facilities, though if the resulting samples are small, this might reduce the extent to which the crime contribution of each premises presents as a clear J-curve. What it may do, though, is increase the number (and thus proportion) of zero-crime facilities. In the current study, using ten years' worth of data means that there is a good chance that many of the facilities in the area are represented in the dataset, as they only need record one crime in the whole ten years to be present. Obviously, the shorter the time period, the less likely (very) low crime facilities will appear in the sample. Therefore, the impact of excluding these, as I have suggested ought to be done, will be greater.

Continuing the consideration of time periods, research has shown that crime tends to be more concentrated spatially over shorter time periods (in this case, that would suggest fewer premises would contribute a greater proportion of crime, as predicted above in relation to the greater chance of excluding very low crime addresses), but that such concentrations are more dynamic. That is to say, that risky facilities calculated over shorter time periods, such as one year (or even less) are more likely to contribute a greater proportion of crime, but also more likely to vary from one time period to the next. There is certainly a debate to be had regarding what is an appropriate time period, therefore, for both researchers and practitioners to use when looking to identify risky facilities. This is picked up again in chapter 10 (discussion). However, it is worth noting here that the choice of time period for analysis will have an impact on the proportion of premises selected and the proportion of crime they contribute. If more stable, consistently risky premises are desired, longer time periods should be selected. If more disproportionately problematic (and smaller numbers) of risky premises are the focus, then shorter time periods will highlight these.

It is suggested that further research looks to quantify the impacts of using different lengths of time, and including zero-crime facilities, on the outputs of the more favourable selection methods. In addition, one possibility for dealing with this would be to code facilities from the crime data over a longer time period in the first instance, so any facility recording a crime would be identified and categorised, and then to perform risky facility analysis for a shorter period (an optimum to be yet established) but including *all address* so that proxy 'zero-crime' premises (for that period) are included. This does not, however, tackle the problem of premises changing purpose, ownership or management during the period investigated. As already noted, therefore, the first step might be to carry out this analysis, and the related comparison, using only those categories of facility for which zero-crime addresses can be included (such as those requiring a license to operate, such as bars, or those that are publicly registered, such as schools). Though not covering all facilities, this analysis should at least suggest some of the implications of including or excluding zero-crime facilities and help move us forward in settling on an appropriate, universal definition (cut-off) for risky facilities.

8.3.1.4.2 Distributions and minimum crime counts

It has already been noted that any selection technique should only be applied when there is evidence that the underlying crime distribution is unequal; ideally following the 'risky facility' type distribution that presents as a J-curve output. The favoured method (F1) reduces the likelihood of facilities being selected (or at least of more than a few outlier facilities being selected) when this inequality is lacking. However, I propose that it should be considered *best* practice to test for inequality of distribution using the Gini coefficient before embarking upon any other tests, and it should be considered *good* practice to also produce ranked crime count (or ideally cumulative crime contribution) plots for visual inspection of the distribution. The results from chapter 7, however, suggest that there is little point trying to establish if the sample is drawn from a power-law distribution, as it is not only difficult to rule this out, but even if we do, many (if not most) categories of facility still demonstrate heavy-tailed distributions of crime, which should be considered the definition of the risky facility phenomenon.

Another issue that may need to be considered, but as yet has not and is beyond the scope of this chapter, is whether there is a minimum number of offences and/or premises needed before it is appropriate to test for (and select) risky facilities. It has been noted that some of the tests and associated conclusions presented as part of this study are affected by the small samples involved (particularly when further categorisation, such as crime type or time, is applied to the less populous facility types). If testing for an unequal (heavy-tailed) distribution and seeking to select premises to identify as risky facilities, it may not be valid to do so in situations where samples are very small. It is not possible, however, to suggest from the analyses carried out here, what such a minimum might be.

8.3.1.4.3 Categorisation

The final issue that needs to be mentioned as part of the consideration of selection methods is the impact of categorisation and choice of categories to apply. Decisions around categorisation, particularly regarding facility type (but also offence attributes, such as crime type and time) cut across this whole research study, therefore they are discussed in greater detail elsewhere (notably in the methodology chapter (6) and the final discussion (chapter 10)). However, there are some important points to note here, with regards risky facilities selection. The way facilities are categorised, or rather the number of categories created will impact upon what gets selected, in two different ways. Very fine, highly specific categories are likely to contain fewer premises, contributing fewer crimes. As has been seen, when this is the case it is harder to establish the extent to which the distribution of such crimes is unequal, and to determine (even visually) if this distribution is heavy-tailed/produces a J-curve plot. In addition, this approach may result in some categories of facility that do not follow a risky facility-type distribution. The crime counts across premises might be quite even, and to select risky facilities for these might be entirely inappropriate. Though the research literature generally suggests that this is unlikely, the findings reported in chapter seven suggest that it is possible. For example, the *religious* category did show evidence of an unequal distribution, which could be said to present as a J-curve, but this was far less convincing than for other categories, such as *education*, leaving open the possibility there might be other types of facility not considered here that experience an even less unequal distribution of offences. Of course, the proposed method of selection (F1) has the advantage that if this were the case, no - or very few - 'risky' facilities ought to be identified.

An alternative issue arises if large, more inclusive categories are selected, or if all facilities are considered together. With this approach, those types of facility that tend to experience higher levels of crime (such as retail establishments) are far more likely to contribute three times the mean, or indeed to be selected as risky whatever method is used, than other types of premises.

This may not be seen as problematic if the focus of practitioners is on those few addresses, regardless of type, that are contributing the most to their crime rates. Indeed, it may be seen as a favourable outcome if generally low-crime facility types (such as hotels or leisure facilities in my research) are not present on the risky facility list - because in absolute terms they may be seen as not really problematic at all.

Somewhere between these two extremes, categorisation and the use of an appropriate, consistent and objective selection technique should allow the opportunity to identify the very 'worst' premises (which will always be identified, regardless) as well as other addresses that may not be quite as significant in terms of their contribution to overall crime, but in the context of other facilities of that type, are actually very problematic. This should be important to practitioners, especially if the reasons are related to easily changeable circumstances (such as environmental opportunity reduction techniques, like altering lighting or removing foliage) or to 'poor' place management, in which case surely it is in everyone's interests that this be improved. It should also be important to researchers, as it should allow for a finer selection of type-specific risky facilities, which in turn should prove more useful for furthering our understanding of this phenomenon, for theory testing and for refining our knowledge of the causes (and therefore most suitable types of intervention). In reality, of course, categorisation accuracy is constrained by the available data, as well as practical and time considerations, as discussed in the methodology chapter (6),

Before moving on to compare risky facilities identified using the selected, preferred method with those highlighted by police officers through interviews, it is necessary to consider the actual premises identified (not just the quantities) and say a little more about the difference between category-based identification of risky facilities (which I refer to as all risky facilities) and non-category based identification (overall risky facilities).

In the methodology chapter (6), I have discussed some of the limitations of the approach taken for this study in light of issues relating to police-categorisation, study-categorisation and the existence of multiple premises with the same address. Briefly, this means that a particular premises, a good example is Walmart (North), might have been categorised as both *supermarket* and *retail other*, meaning that it appears in both categories and, crucially, its crime contribution is also split across both (depending on the category used, by the recording officer, for each offence). This means that the ranking of such premises *when looked at by category* will be affected (in other words, it will seem less problematic than it is) and, if near the cut-off point, there is the possibility that it might not be selected as risky. Therefore, although premises that are risky *within* a given category (and therefore calculated only from the crime distribution for facilities of ‘the same’ type) are of interest, the outputs produced from all facility crime considered together are actually a more accurate representation of problematic locations within the study area. Therefore, the *all* outputs form the core of the comparative considerations below.

Of course, the converse of this categorisation issue is that there are some addresses that genuinely cover multiple premises, such as the OH Shopping Mall, and sometimes these premises fall in different category types. Therefore, when considered all together, they may appear *more* risky, because the crime count is actually being contributed by more than one (sometimes several) properties. This results in an inflation of how problematic they appear, as well as pushing other, single premise-addresses down the rankings. Another issue with using all facility crime has already been mentioned above; that some premises might be significantly more problematic than others in the context of that type of facility, but do not surpass the cut-off calculated from the all facility crime mean. The result of this is less diversity in facility-type representation. For this reason, the categorical risky facilities were also compared with police perceptions, though these are referred to less.

A particular note on convenience stores and subsequent re-categorisation

In the study area, premises categorised as convenience stores were a mix of what could be considered small ‘local shops’ or neighbourhood stores[13], which in the UK we might refer to as ‘corner shops’, and convenience store/gas station combinations. The latter were mainly branches of regional or national (or even international) chains, combining a fuel sales function (provided by WilcoHess, BP, Shell, Citgo, and so on) with a shop, providing snack foods, beverages and sometimes a limited selection of toiletries and groceries (such as Pantry Shop, Kangaroo Express, and Fairway OneStop, amongst others). This did include, however, some independent stores that also sold fuel (usually with limited pumps). There was clearly, therefore, an overlap between this category and the *petrol* category. The original coding, as discussed in chapter 6, sought to separate facility categories, as much as was possible using the information provided, by premises’ function and the routine activities of those who used and worked in these types of premises. In some places, neighbourhood convenience stores may significantly overlap in function with the shops attached to gas stations. Particularly those near to residential areas, where they may be the nearest, easily accessible retail outlet for those wanting basic goods, snacks, tobacco and alcohol (as most premises sold at least beer). In other locations, the shop or kiosk may be used in a very different way and by different types of people, mainly those buying fuel, and picking up a limited quantity of other items at the same time. It was ultimately felt that the similarities in convenience store function were enough that these should be looked

[13] Distinguished from small grocery stores and supermarkets by the more limited range of products, but possibly open for extended hours; though there is a grey area between larger

at together. This meant that if a convenience/gas store combination was recorded as a convenience store by the reporting officer, it was assumed that the offence was perceived to be connected to the shop function, whereas if it was recorded as a gas station, the offence was perceived to be connected to the provision of fuel (or possibly occurred on the forecourt) in which case this should be kept separate as involving different routine activities (notably involving individuals using a vehicle and possibly stopping en route elsewhere, as opposed to 'walk ups' by local residents, primarily using the shop).

The separation of the two facility types, then, was conceptually justified. It became apparent, however, when identifying the names of premises that had been selected as risky facilities (which as stated could not be done for all premises in the dataset, so had not been done previously), that assuming officers would judiciously separate offences by whether they were associated with the convenience store or the fuel sales function, was probably flawed. In addition, by splitting the premises up in this way, each location appeared to be *less* problematic than it truly was, because the crime contribution was divided across two categories. However, it was still very important that neighbourhood stores be kept separate from gas stations as they were often very different in style, management, provision, and location. In addition, it had become apparent during my field studies (and this was echoed in the interviews) that the neighbourhood convenience stores were seen as real crime generators and attractors, variously as good targets for offenders, as places that facilitated or were otherwise involved in offending, and/or as crime hosts, because local people tended to 'hang out' in or around the premises. Whilst some of the convenience/gas station combinations also served such a function, others were very different, being on busy vehicular thoroughfares, or near highway interstate off-ramps. The two were just not comparable.

Therefore, when it came to carrying out the comparison of risky facilities from the crime data with those from the interviews, the overall risky facilities list, which covered 157 risky facilities, was manually post-categorised by type. This was done by searching for properties at that address, using the interview data to provide further information and then researcher-determined assignation. In doing this I created a 'multiple' class for when more than one type of facility shared the same address,[14] separated *convenience* into 'convenience' (neighbourhood type stores) and 'convenience/gas'[15], and introduced a 'public building' category. The counts for each

---

convenience stores and small, local groceries

[14] This was only used when the premises were of truly different types, therefore if all the premises were retail (including supermarkets or convenience stores) the retail other category was used. An example of a multiple categorisation would be a shopping area with shops, restaurants and a bowling alley.

[15] Petrol stations, using the vernacular of the interviewees, which invariably have convenience stores (albeit of different 'types' and size) attached to them

resulting category are shown in table 20, below.

When the premises identified as risky in each category were added together, this suggested there were 196, however some of these related to the same address, in different categories. So that when looking at distinct addresses the total number of risky facilities (referred to as ‘all risky facilities’) was 163. Comparing to overall risky facilities (those premises meeting the RF_F1 risky facilities definition when category of facility is ignored) this appears to be a difference of only six addresses. However, there are a number of addresses that only appear in one of the lists, meaning that the overlap between the two is actually 106 premises. This difference is predominantly made up of facilities identified as risky when looked at in their different categories, but with relatively low crime contributions when considering overall offence levels (represented in the all risky facilities list; unmatched address n=57) or those that had been recorded as different types of facility (thus splitting their crime contribution), which then moved up the ranking when these were combined together (represented in the overall risky facilities list; unmatched address n=51). Notable exclusions from overall risky facilities are all religious premises, a number of restaurants, some of the smaller, and individual shops (that found themselves in the same category as large department stores and whole shopping centres and malls). Those tending to be less represented when the data are separated into categories are pharmacies, smaller grocery stores, and convenience store/gas station combinations. This is clearly because such premises are more open to interpretation regarding how they should be categorised. Also missing were public buildings, which were not included in the original category-based analysis.

These differences again highlight the need to carefully consider the benefits, and weaknesses, of analysing all facility crime versus separating the data into categories that better encompass the range of facilities, the purposes they serve and the individuals and routine activities associated with them. All of this has been taken into account when carrying out the comparison with police perceptions. Though these differences do raise concerns regarding the ability to accurately compare, the top ranked/high crime contributing facilities are very similar across the two approaches, and, as will be seen in the following section, the patterns of which are and are not identified by police officers are not particularly affected by the distinction. It is to these findings, and discussion, that I now turn.

Table 20: re-categorisation of facility types and contribution to overall risky facilities

New category	N premises	% RF_F1 premises	N crime contribution	% RF_F1 crime contribution
Convenience/gas	42	26.75	4674	22.40
Retail other	20	12.74	4672	22.39
Neighbourhood convenience	17	10.83	999	4.79
Educational	14	8.92	2205	10.57
Financial	14	8.92	870	4.17
Supermarket	13	8.28	3200	15.34
Pharmacy	9	5.73	874	4.19
Hotel	4	2.55	214	1.03
Public building	3	1.91	349	1.67
Restaurant	3	1.91	151	0.72
Leisure	2	1.27	126	0.60
LP	2	1.27	95	0.46
Healthcare	1	0.64	173	0.83
Mutiple	13	8.28	2261	10.84
Total	157		20863	

8.3.2 Police perceptions of risky facilities

Considering police perceptions of facilities they identify as notably problematic, serves three, related purposes. It potentially acts as a form of (limited) validation of the use of recorded crime data (and the favoured selection method) as a way of appropriately identifying risky facilities. It can also be used as direct comparison, regarding the premises that experience the greatest proportion of recorded crime, and those that the police *perceive* to be the most problematic. Finally, it acts as a counter-point, allowing a different type of exploration of the way risky facilities come to attention, that can provide ideas and insight to inspire further consideration and testing, or to enhance our understanding of this phenomenon. What now follows is a presentation and discussion of the findings of both the content and thematic analyses of the police interviews (and annotated maps), supplemented with consideration of relevant findings from my field studies. Throughout, comparisons are drawn with the results obtained from analysis of the police recorded crime data, as already discussed.

The overall themes (and sub-themes) from the thematic analysis have already been presented in chapter 6. Brief information about the respondents is shown in table 21. Given that this section is being compared to results from the previous section, and because not all of the qualitative findings are relevant to the issue of *identifying* facilities that are risky, the thematic framework is not used to present this material. Rather, here I introduce a few very general issues worthy of consideration, then the types of premises and locations selected by the respondents are discussed, and this is followed by a consideration of what I refer to as 'missed' locations - which can apply either way round - broadly presented by facility type.

Table 21: Respondents and roles

Respondent number	Sex	Rank	Role	Number of years service	Previous roles/ experience
1	Male	Detective	Violent crime	14	Patrol; violent crime task force
2	Female	MPO II	Patrol (Housing Authority Unit)	11	Patrol
3	Female	Detective	Property crime (non-residential)	Not recorded	Patrol
4	Female	Detective	Property crime (pawn shops)	13	Patrol
5	Female	Officer	Patrol	9	Patrol only, but different areas (South & Central)
6	Male	Officer	Street crimes	7	Patrol

MPO = Master Police Officer

8.3.2.1 Overview

It is interesting to note from the outset that there was quite a degree of variation in the premises and locations highlighted by the different respondents. Many places were readily identified in multiple interviews, whereas others were only named in one or two. There was a much greater degree of consensus regarding larger areas, such as particular housing project areas, than facilities. That said, most respondents mentioned the big box stores (two Walmarts), the cinema area (that also included restaurants and shops), and the Oak Hollow Mall (which was in the process of closing) and shops in its vicinity. Other than a gas station/convenience store combination, all other named facilities were mentioned by only one or two respondents. This might have reflected the officers' differing current roles/beats, and/or previous experience. Alternatively, they may have focused on different issues (different types of place, different types of crime, different types of offenders, and so forth), which might itself have been the result of their different roles. It is also possible that they started out talking about different things, which naturally took them through a series of related thought processes (and premises) until they reached saturation (in terms of what they could recall, or with the interview process itself). Had they started with somewhere different, taken a break and started again, or been prompted differently, there is a possibility they may have included more, or different, locations. Further, the map that was used was the only version that was available at the time. It was an A3 map with street names and coloured beats, but some of the respondents did not find it particularly easy to orient themselves within the map, or to find some of the places they wanted to talk about.

> It's funny that when you look at a map and you think of places, it's hard to put it together. (R1)

> Seven [referring to numbering on the map]. Brentwood Grocery 1201. Oh my gosh, 1201, I go there every fricking day, it's not New Street, it's this street here, that's New Street, the 1201 Brentwood Grocery. (R2, after marking the premises at the wrong place on the map).

Different locations and facilities may have been identified if a different map size, colour or layout had been used, or if the area data was collected without reference to a map, or was collected 'on patrol' (accompanying officers and noting places referred to during the shift, for example).

8.3.2.2 Types of premises and selection

When talking about, or prompted to name, individual, specific premises (that could be defined as facilities) respondents identified convenience stores the most. Between them they also highlighted a number of supermarkets, plus other retail premises and general shopping (such as malls, shopping areas and strip malls). Some of them also talked about schools and there was some mention of licensed premises (sometimes mixed with leisure facilities, such as was the case for a billiards/snooker hall), and gas stations (or gas station/convenience store combinations).

Respondents did not really talk about any other categories of premises. Reflecting on the categories used to analyse the crime data, they did not mention financial establishments, healthcare, pharmacies or religious premises at all, and rarely mentioned hotels (n=1; 1 respondent) or leisure facilities (n=2; 1 by only one respondent, the other mentioned by four). For the latter category, neither of the premises that were identified met the criteria for being a risky facility in the crime data (though they were present). Interestingly, as will be discussed more later, there *were* two risky leisure facilities identified in the crime dataset, but neither of these were listed by interview respondents.

Already, this focus on particular types of facility requires consideration. The majority of premises were discussed with reference to theft-related (acquisitive) crimes, thus it was not surprising that these were mainly retail establishments. However, some of these were (also) noted to experience violent offences (convenience stores, for example, as discussed below) and of course educational establishments might be the locations of a variety of different types of crime. It is noted that pharmacies are retail premises, but these were not named by respondents, despite some being identified as risky from the recorded crime data (4 in the *pharmacy* category, 9 pharmacies being present in the list of overall risky facilities). Respondent 1 did mention (and annotate the map - marker 8) an area that had a number of pharmacies as experiencing some trouble, but this was discussed in light of another premises in the vicinity (see discussion of public buildings, below) and so was not counted as identification of risky facilities. Other than that, there were two premises situated in general shopping areas mentioned by the interviewees.

There is no apparent reason why pharmacies would not be identified, other than that they just were not perceived as problematic compared to the types of retail outlets that *were* identified. It is possible that the offences recorded at pharmacies are less likely themselves to be construed as problematic. It could be, for example, that offences involving frauds or other misuse of prescriptions are not detected until after the fact or are not seen as being as problematic as shoplifting or fights. Revisiting the original data provides some support for this suggestion. About half of recorded offences at pharmacies were shoplifting and there were also some other larcenies and burglaries, plus a small number of robberies. The majority of the remaining crimes were related to offences of dishonesty (frauds, forgeries, embezzlement and counterfeiting) predominantly relating to prescriptions, and it is likely that these offences did not demand an (immediate) response.

Another type of facility that was rarely highlighted was licensed premises. Although there was some mention of drinking establishments or leisure/entertainment venues with bars, these did not appear to be at the forefront of respondents' thinking, and on some occasions were only mentioned at prompting. Analysis of the interview and crime data suggests that crime related to on-licensed premises is not particularly a significant problem in the study area. This is somewhat unusual when compared to other locations in which research has been carried out, and is certainly different to the general UK experience. Thus on-licensed premises have often been the subjects of research into crime generators/attractors and risky facilities, as noted in the review of the literature. However, for the current study, crime occurring at licensed premises was neither a particularly problematic crime contributor (2.04% of facility crime) or an issue of concern for serving police officers. It is acknowledged that the crime excluded because it did not occur in a facility could well have been associated with attendance at licensed premises. However, the thematic analysis also showed that this was something that was not really a current issue, which Respondent 1 acknowledged was unusual.

When licensed premises were asked about, the three respondents who spoke about them explained there had been problems previously, including "…wall to wall people, fights, shots fired, people being shot people being hurt…" (R1) and that there had been particular issues with violence at "Hispanic Bars" (R5). However, this was now seen as something from the past, as such problems at bars had been the focus of initiatives, violation proceedings and nuisance abatement notices, resulting in a number of closures. Now licensees/owners were perceived to make efforts so that things were safer, knowing police would respond and take action if they did not. Respondent 2 adds to this, explaining "…if we do have fights round licensed premises it's either rival gang fighting or drug deal gone bad or something like that…it's like a ghost town now." Although my observations were predominantly during daylight hours, I particularly noted the lack of obvious drinking establishments in the city. The main entertainment areas were situated around the cinema, or other shopping areas, and also involved restaurants with bars/sports bars. In the city there are premises that operate 'just' as bars and clubs but, partly because it does not have a downtown area, there is not a large cluster of late-night venues. This might also have affected the nature of licensed premises crime, and how concentrated (thus problematic) it was perceived to be.

As stated, no financial premises were mentioned by respondents, though 18 were identified as risky facilities from the categorised crime data (14 from all facility crime), with this type of premises contributing 3.78% of recorded facility crime (and ranking sixth of the twelve categories that did not include retail other, seventh overall). It is possible that, in a similar way to pharmacies, the type of crime committed at this sort of location was not something the officers that were interviewed tended to respond to, or that it was simply not thought of as that problematic (in terms of public disorder or extent). Returning to the original crime data suggests the vast majority of offences occurring at financial institutions were, as anticipated, dishonesty offences (counterfeiting, fraud, forgery, and so forth), so like pharmacies, financial institutions may not be perceived as much of a drain on patrol resources. However, as this was not explored through the interviews, there is no evidence of this either way and such premises are not considered any further in this part of the study.

Somewhat unexpectedly none of the respondents mentioned healthcare facilities either. Analysis of the *healthcare* category identified four risky facilities (from a total of 171), whilst there was just one premises included in the overall risky facilities. It has already been noted that the crime contribution of this type of establishment was generally low excepting one large outlier: the regional hospital. It is reasonable that the interviewees did not particularly consider this category, then, but as they were asked about problematic or high crime locations (rather than to discuss specific categories of premises), it is a surprise that the hospital was not mentioned. Overall this address was ranked 30th by crime contribution to all facility crime. Again, this was not asked about specifically in the interviews, but it is possible that respondents just did not think of it as a ‘facility’. They tended to discuss areas where there were crime problems, including residential ones, until they were prompted to focus on specific “facilities” or “problematic premises” [interviewer] and then they discussed acquisitive crime, invariably in retail premises, or disorder and violence in restaurants or at leisure or entertainment facilities. The only exception to this was schools. Therefore, it seems that for some reason the hospital, despite being a substantial crime contributor, was not thought of in the same context as ‘other’ facilities or residential crime. In fact, on the annotated maps, the area directly around the hospital was not even encompassed by the places highlighted by any of the officers (see Appendix 5). One respondent identified the library (see below) and nearby pharmacies, which were situated close to the hospital (Respondent 1, map-marker 8), but still did not mention the hospital at all.

Additionally, no religious establishments were mentioned by respondents, despite there being 249 different addresses listed in the dataset for this type of premises. It is quite possible that religious buildings and organisations are not perceived (at least by the officers interviewed) as facilities, although, as stated, they did freely mention other types of locations that would not be either, including residential areas, parks and street corners. Of course, religious premises did have the lowest mean crime count and the least unequal distribution, so it could be that the respondents were correct to ignore these establishments. Although this type of facility showed evidence of a concentration of crimes, and 14 ‘risky facilities’ were identified within this category, it is possible that religious establishments should not have been included as one for which it was appropriate to select locations as being problematic premises. The maximum crime contribution from any one address was 16 recorded crimes (which is notably less than an average of two a year). Further, none of these premises contributed sufficiently high levels of crime to be recorded as risky when all facility crime was considered together. This supports the suggestion that there may need to be a minimum number of offences (in total or for the maximum contributor), based on overall crime incidence for that area, when identifying categories to include in risky facility testing.

As has already been mentioned, one category of premises that was not included in the category-based analysis was ‘government or public buildings’. This, therefore, is included in the *Other* category shown in table 4. There were 722 offences recorded in this category, but it was felt that the premises this encompassed would be too diverse for the category to be meaningful. However, when all facility crime was analysed, these premises were included, and some of them then met the overall risky facility criteria. However, no public or government buildings were identified in the interviews, other than respondent one considering the library, which sometimes experienced disorder problems, as a result of people who were homeless using the facilities (air conditioning, sometimes the Internet). The library was a good place for them to be able to access some comfort and resources that they would otherwise not be able to, and there was also a shelter a few blocks away, which meant such individuals were more likely to be in the vicinity. This was not seen as a big problem, but it *was* specifically mentioned; "It's not often it happens but there are some things that do occur up there." (R1). Re-visiting the crime data showed the library recorded just 25 offences during the period studied (3.46% of offences in that category).

The greatest contributors in this category were City Hall (N=126 (rising to 137 when facility-type codes are ignored); 17.45%), the Police Department (N=106 (rising to 125); 14.68%) and the County Courthouse (N=87; 12.05%), the library being the next highest. These three premises are not at all surprising, but as stated, none were mentioned in the interviews. Perhaps the officers did not think of them as 'facilities', or perhaps they simply did not perceive them as addresses that were either 'targeted' by offenders or 'caused trouble', even if potentially a large number of crimes were *recorded* as occurring there. In hindsight, it would have been interesting to include these public buildings, as even a cursory inspection suggests they follow the risky facility distribution. That said, I do not think this adds much to our understanding of risky facilities, given the types of premises that were the greatest contributors are also those where offences are 'created' by virtue of the business they carry out.

Another type of facility that was hardly mentioned by the respondents was hotels. Seven of these were selected as risky facilities from the recorded crime data, the top ranked contributing 70 offences over the ten-year period. Compared to all facilities, just four were classed as risky. Risky hotels did not contribute a particularly large amount of crime, but nor was it inconsequential (and the worst address alone contributed 16% of all recorded hotel crime). I would have expected hotels to experience much greater amounts of crime, as they appear very suitable targets for acquisitive crime (from minor thefts to burglaries and theft by employees), violence committed on the premises (including familial assaults), and drug and vice offences. However, this was not reflected in the recorded crime, nor does it appear that most respondents thought of this either, with just one flagging a hotel because it had "…cabanas and rent by the hour motel rooms." (R4), where it was implied sex work and drug deals might be arranged to take place. This would suggest that if hotels in the study area do suffer reasonably high levels of crime, that these do not regularly result in police attention.

Returning to the types of facilities that the respondents did highlight, it may be the case that the focus on retail premises could be reflective of the resource demands placed on the local police, with theft offences likely to be the most numerous, or on these respondents specifically (although two dealt with violent or street crimes, and two carried out general patrol). Thus, whilst non-retail categories might also include risky or problematic facilities, these may be less demanding of police time, so considered less of an issue. Additionally, across the interviews there was discussion of gangs and gang-related activity and drug sales (and systemic crime, Goldstein, 1985). These activities, and the violence and disorder associated with them were clearly seen as a key focus for police attention.

> …this whole area [indicating a predominantly residential area on the map] has become a hotspot too for gang activity lately. (R5)

> Again it's a drug area and gang territory. A lot of gang members I won't say members, I'd say affiliates live there. (R2)

> By violent crimes I mean, you know, shootings and stabbings and most of that stuff is usually over drugs or gang-related. (R6)

These behaviours would not necessarily be associated with facilities, *per se*, but both gang and drug crime were mentioned in an explanatory manner, as will be discussed in chapter 9.

There is also the possibility that the results are more reflective of the way the data have been collected and analysed than they are of any true differences in risky facilities between police perceptions (and knowledge) and recorded crime data. It may be that the way I described the study, or things the respondents had heard about it prior to being interviewed or even the questions or prompts I used might have (inadvertently) biased their discussion towards certain types of premises. Indeed, I had been on two ride-alongs/observations with civilian and uniformed staff prior to the interviews, and there had been much attention given during this time to convenience stores (both at mine and their prompting). This might be because they were seen as particularly problematic, or because they were simply places it was thought would be good for me to observe.

Further, and this is discussed more below as it is particularly relevant to patterns where respondents identified premises or locations that had not been selected as risky, the analysis of recorded crime data used only those offences occurring *at* facilities, whereas the interviewees were free to consider this more broadly. Indeed, there were some comments relating to crimes ‘starting’ at certain places (though that might not be where they finished) and crimes occurring around residences or along pathways taken by (potential) offenders. This was particularly mentioned in relation to housing projects and to crimes occurring in the vicinity of schools. Respondents were able to consider which facilities were risky from a wider evidence base, including crime recorded as occurring in the street, parks and residential areas.

In addition, despite being prompted to focus on facilities, they were also free to discuss offences happening in other types of location, which they did. Sometimes this was with little reference to any types of premises (such as when they talked about problematic housing projects, routes, or street-corner hang outs) and sometimes it was to refer to problems in specific locations, but that did not meet my definition of facilities, including outdoor spaces (such as parks) but also industrial premises and construction sites. This meant that as well as noting some specific facilities, the annotated maps generally ended up being more akin to hotspot and offender residence maps. This may reflect the way officers think about crime maps. Alternatively, it may be because (as discussed more in chapter 9) identified facilities were quite often (though certainly not exclusively) situated close to other facilities as well, for example in strip malls or shopping complexes, so the whole area was encircled. It is also quite possibly because crimes are not so clearly associated with specific premises as the methods I employed suggest, so that officers tended to annotate using a roughly drawn approximation of a buffer, accounting for what they perceived to be the ‘sphere of influence’ of a facility or group of facilities. This adds to the debate about using only crimes recorded against particular premises versus employing buffers or street segment counts, and also raises the idea that if buffers are to be employed, there may be some merit in research with police officers regarding suitable spheres of influence. Indeed, it would be interesting to compare the locations (addresses and general areas) that the respondents identified, with all recorded crime, not just that which was recorded as at a particular facility, to see if the findings were more (or less) similar, or if it made little difference.

A final issue to consider with regards the types of location and premises that the respondents selected, and how this differed to what was included in my analysis, was that as far as possible, I specifically only selected what in the UK would be described as recorded crime data. This has been explained in the methodology chapter. Of course, the interviewees were drawing on all their experiences and recollections when determining what areas and premises to tell me about, so this would have included crimes, but also calls for service and possibly intelligence information, as well as additional personal and institutional knowledge. This could produce substantially different findings (which is, of course, one of the reasons for including qualitative data and analysis in the study).

8.3.2.3 'Missed' locations

Here I define 'missed' locations as those specific addresses that were identified in the crime data but not in any of the interviews *and* those mentioned by one or more respondents, but not flagged as 'risky' in the crime data. Overall, there was only one premises that did not appear in the crime data at all (no recorded crime in the ten-year period studied). However, there was some confusion over this address, and it is possible that one recorded crime occurred there (listed as retail other, though it is actually a convenience store). Either way, it was certainly not categorised as 'risky' from the crime data. All remaining premises identified in the interviews were able to be associated with an address in the dataset analysed.

It is important to recognise that (as stated previously) this is not a 'test' of police officers' knowledge, nor a test of how accurately the empirical method identifies risky facilities, though discrepancies either way direct further investigation into both of these issues. It was not expected that the respondents would mention every premises that was labelled as risky through my analysis, as this would be a very large number of premises to be able to think of 'on-the-spot'. Rather, discrepancies relating to the most highly ranked premises, peculiar patterns in what was thought of as a high crime premises compared to what was not, and seemingly problematic premises that did not appear as such in the recorded crime data, are what I am focused on exploring.

Given the discussion above about the types of facility that were identified, the main establishments that I focus on are educational establishments, supermarkets, and convenience stores. Educational establishments and supermarkets are particularly interesting because all of those identified in interview were classed as risky facilities from the crime data, but there were some highly ranked risky facilities identified from the crime data that were not mentioned by the respondents. Convenience stores are also discussed. Across the interviews, they were the most frequently identified type of facility. Analysis of the qualitative data also highlighted an important distinction between neighbourhood-type convenience stores and those attached to gas stations, which means the *petrol* category is also discussed here. Finally, there is a brief consideration of entertainment areas and establishments (encompassing *leisure*, *licensed premises* and *restaurants).* Although some other retail premises were also mentioned, all the significant findings relating to patterns in retail facility crime are covered through discussion of supermarkets and convenience stores. First to be discussed, then, are educational establishments.

8.3.2.3.1 Schools/educational establishments

There were ten identified risky facilities in the recorded crime data for this category, whilst respondents identified just three. One respondent highlighted all three (R4), whilst another highlighted two (R5). The other four respondents did not mention schools as risky facilities, although respondents 1 and 6 referred to schools when talking about routine activities (such as events being put on at School Park, or young people skipping school to get into trouble), but not the school premises themselves. The overall risky facility list includes a further four schools (all with lower crime counts, and all elementary schools). Therefore, the number of 'missed' establishments varies somewhat, but the general pattern remains the same, so the *educational* category has been used as the basis for the following discussion.

There is no obvious reason why some respondents picked up on educational establishments experiencing large amounts of crime, and others did not, and there was nothing else in terms of the issues the different officers tended to focus on that would obviate consideration of schools, other than perhaps respondent 3 who was definitely more focused on discussion of premises and areas that experienced theft-related offences, mainly shoplifting, burglary and 'car B&Es'. This was not surprising given her current role as a Detective focused on non-residential property crime. A larger number of respondents, as well as officers with a broader range of specialisms, might have resulted in schools being mentioned more often. This would have been particularly likely if officers with responsibilities for school liaison or involved with youth work had been represented.

Returning to the educational establishments that were identified (or not), of particular interest is the fact that the top ranked school (in terms of crime count over the study period, with 508 recorded crimes) was *not* identified in interview, whilst the 2nd, 3rd and 4th ranked facilities (with crime counts of 291, 269 and 256 respectively) were. Also of note is the fact that where there were middle and high schools on the same or nearby sites, one pair *was* identified by respondents, one pair was not identified (this included the top ranked address) and, for the third, only the high school was mentioned. Finally, the university (which was included in this category for my analysis) was ranked as the 7th most risky educational establishment, but was not mentioned at all by the respondents. Most (but not all) of the possible methods for identifying risky facilities that were considered earlier in the chapter, would have selected eight or more premises as risky. Even if these criteria were too generous, the fifth ranked facility was certainly not a substantially lower crime contributor than the fourth (which was identified in interviews), and the sixth, seventh (and possibly eighth) ranked facilities were also not that far behind, as shown in table 22.

Table 22: 'Risky' educational establishments

Rank	N offences	% of offences in *educational*	Street	Premises type	Identified in interview
1	508	18.57	Barrow	High School	N
2	291	10.64	McGuinn	High School	Y
3	269	9.84	Ferndale	High School	Y
4	256	9.36	Ferndale	Middle School	Y
5	217	7.93	McGuinn	Middle School	N
6	178	6.51	E Washington	School for Arts (Middle and High School)	N
7	122	4.46	Montlieu	University	N
8	93	3.40	Barrow	Middle School	N
9	63	2.30	Fairfield	Middle School	N
10	50	1.83	N Centennial	Elementary School	N

Considering the university, it at first seems unusual that the respondents did not mention this 'facility'. Indeed, it is highly likely that the number of offences in the recorded crime data associated with this establishment would actually be far higher if all university property had been included, but given the method of coding and identifying 'premises', instead this will only encompass those offences that were attributed to, or occurred at, the main university building(s) (that is to say those recorded against the official university address). Although this means that the actual experience of crime on campus will be (potentially much) higher, given that I am only looking at facility crime, it is more appropriate that other locations (such as accommodation) are excluded. Either way, one might expect the university to be seen by officers as a source of crime and disorder, including under-age drinking and public drunkenness, drug offences, sexual assaults and harassment, and property crime, but this did not appear to be the case.

During the interviews, the university was mentioned only once in the context of crime, and this was in reference to a small shopping centre nearby that had previously experienced a lot of crime problems (until off-duty officers had been hired as security). It is unclear, therefore, why it has not been identified as a high crime location. It is possible that the security arrangements for the university mean that the police are less likely to deal with offences on site. Although there isn't a campus police force - as apparent from the data, offences come under the purview of the study area PD - much enforcement activity is likely to be carried out by the large security force (supplemented by off-duty police officers) who will likely deal with a lot of complaints and incidents, thus reducing the calls for service to the local PD. This would be even more the case if there is a culture of not wishing to involve the police in campus incidents, to protect the reputation of the university, though this is purely conjecture.

Whatever the reasons, the university was either not seen as problematic by the officers interviewed, or at least it was not at the forefront of their minds. It might also be the case that the sort of individuals most likely involved in offending at the university (students) do not fit the stereotype of typical offenders, and so are not called to mind so easily. Another possibility is that they did not think of it as a 'facility', although they did mention other locations that were not, so this is unlikely to be the only reason. Finally, I visited the university as part of my observations (and during an accompanied ride along). It was very neatly presented, and is known to be an expensive, modern and comfortable campus (referred to by some as a 'country club' campus). It is also surrounded by fencing and access gates. The general impression created (which may not at all be the case in reality) was of a well-looked after, safe area, and also one that felt very separate to the rest of the city. Both of these features make it more likely that it would not be identified as a 'problematic facility in the city', giving the impression it is neither problematic, nor perhaps truly part of the city.

In order to seek to explain the other differences identified, I turn to a consideration of the locations and local populations. The top ranked school, which was not considered a risky facility ("There's another school…which is up north but we don't have nearly as much problems up there as we do here" (R5)) is at the far north of the study area, therefore it might simply be that because it is close to the edge of the area covered by the PD they tend to spend less time there *unless* they have been called to a specific incident that requires police attention and are generally less well versed with the general surroundings and (criminal) activity. As none of the respondents mentioned it, though, this cannot be confirmed. A map of the study area, showing the position of facilities from the *educational* category is shown below, and the school in question is highlighted (figure 4).

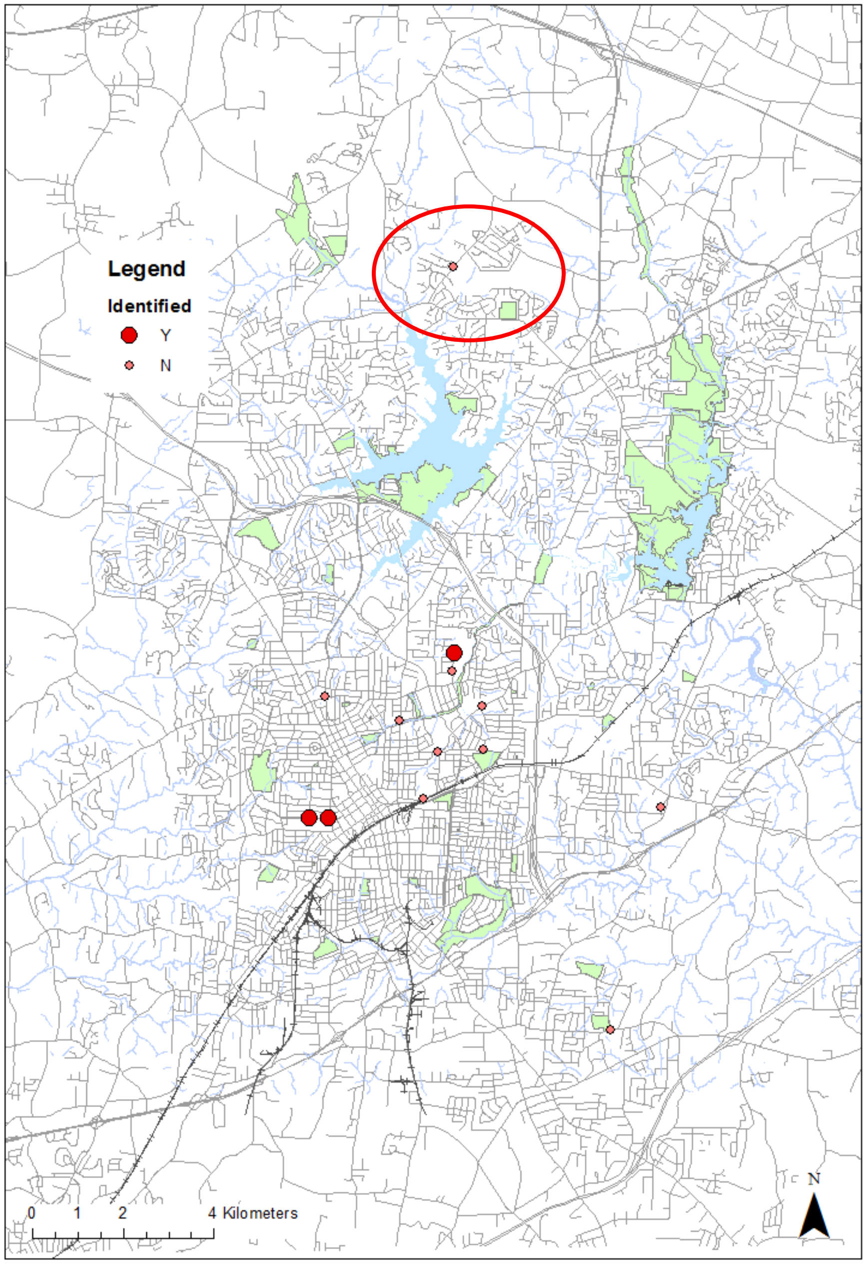


Figure 4: Map of educational point data, showing whether identified in interview or not (top ranked school (unidentified) is circled

I did not specifically set out to observe schools, however, driving around the city, it was apparent that this area to the north felt a long way out from the rest of the city and also required the crossing of a future Interstate, that was under construction. As such there were both physical and possibly psychological barriers separating the northern areas from the southern (as discussed later, and in chapter 9), and this was even more so for the area where the highest contributing school was situated. It is possible, therefore, that although officers would attend the northern fringes of their district, they would likely be in this area much less frequently than central and southern areas that were more associated with criminal activity and drug problems (and where the PD building was located), thus not only would they be less likely to think of problem premises in that area of the city, but they would likely be less familiar with it and it would be excluded from, or at least less well defined in, their 'mental maps'. Indeed, although not talking about this premises, one respondent commented: "I don't know too much about this area because I don't go up there, or this area I think, I don't know too much about that, that's the rich people." (R6, street crimes).

An alternative explanation is that the schools that *were* identified might be those that are more closely associated with particular local gangs and gang activity. As will be seen later, discussion of gangs, gang problems, 'gangbangers' and gang or area feuds was a common theme appearing in the interviews. This was both in general and in relation to specific types of premises, including schools. Some of the offending discussed at educational establishments was done so in the context of 'who' went there and what gang or housing project they were associated with. It is possible that the schools identified by respondents as risky were those where the main crime problems could be associated with gang activity, whilst those not identified may be schools where this is less of an issue. This might also explain why only one middle school was mentioned (particularly when these tended to be situated next to their associated high school). Middle schools may be perceived as having less gang problems than high schools. Of course, it could also be that they are perceived as having less *crime* problems in general, because of the younger age of pupils. This is to some degree supported by the data, with the three top ranked premises being high schools, before middle schools make an appearance, with the first elementary school listed in position ten.

It must be remembered, as with all the facilities under discussion, that there could also be a denominator effect. The size, number of users, turnover, and so forth are all likely to impact on the amount of recorded crime. This chapter does not seek to explore the nature of risky facilities or explanations for their existence, however. Rather it is interested in how we decide which premises they are. That said, it is also probable that size, popularity, centrality and the frequency with which officers attend any particular facility (either on-duty or for their own purposes) will affect its familiarity to them, thus their perceptions of it. This explanation for discrepancies between the interview and recorded crime data will potentially apply throughout.

In terms of the types of behaviour and criminality taking place that caused respondents 4 and 5 to identify the particular schools they did as high crime locations, one mainly referred to "…all sorts of activity going on with the students here." (R4). The other respondent mentioned kids from one of the schools breaking into it, and that the other school problems, which were a big concern, were violent crimes associated with either gang affiliation, or rivalries between schools, especially when pupils were sent to schools further away.

> We have a lot of violence over there, lot of assaults and we've had a couple of stabbings…in the past (R5)

> There's a big gang problem over there too [referring to a particular area] and a lot of that has to do with the High School [3rd ranked premises]. (R5)

Some schools appear to have been identified by officers as they were perceived to 'create' further crime problems in residential areas. On the other hand, some schools themselves (including the one that actually experienced the greatest amount of recorded crime) were seen as less of a problem, but the allocation policy that saw young people from the same, or bordering neighbourhoods being sent to different schools around the county meant that problems *were* created back in the areas where they lived, and it was this non-facility crime that was of most concern:

> There's a lot of these places over here…[names some residential streets]…there's all your little gang bangers that live in this area and they're the ones that stir their own little pot, you know? You know if they have beef with another gang member or from another school [pause] most of the time they live in this area, because some of the kids that go to [school in an adjacent town], I can't really remember how they separate it, it doesn't make sense really. Some of these kids go to…[the top ranked 'risky' school]…so then when they come home there's just tension in the neighbourhood. (R5)

Therefore, it would appear that few officers thought to mention schools when identifying high crime areas or facilities, but it is unclear why this was the case. The top ranked school in terms of crime contribution was either not mentioned, or was thought to be less of a problem than other schools, but it experienced two thirds more crime than the second ranked school and was the fifth biggest contributor to overall facility crime in the study area. Some explanations have been offered as to why this might have been overlooked, with the idea that it is because it is situated towards the edge of the city limits, in the more affluent part of town, being a convincing but currently unverifiable one.

It is not possible to ascertain the extent to which the perceptions of the six officers that were interviewed reflect those of the rest of the PD. As discussed in the methodology, qualitative data are generally not expected to be ‘representative’ in the way that quantitative can be, and that was certainly the case here, with the interviews expected to be illustrative (and hopefully enlightening), rather than generalisable or validatory. However, if these perceptions are broadly similar to the wider policing body, there is the the possibility that the necessary pro-active and preventative work that ought to be carried out with this school would not take place. Of course, good problem-orientated analysis should still highlight this location, but there is the risk this might not be acted upon if it does not match officer perceptions.

An alternative (or additional) possibility is that the schools ranked 2nd, 3rd and 4th, that *were* mentioned by the respondents, actually have worse problems than the top ranked school, but do not report as many of them, thus the recorded crime contributions are an under-representation of the crime problems there. If this is the case, but officers do not realise this, then although they may target their enforcement resources appropriately, they may not take action to improve reporting rates. Accurate reporting of offences is important in helping enforcement agencies identify and understand the nature of problems requiring their attention. High reporting rates also tend to be indicative of good police-organisation (and possibly community) relationships, which in turn are associated with more effective interventions, and lower crime rates. Of course, there is also the possibility that the crimes occurring at these schools are more serious or resource intensive, again suggesting more attention needs to be paid here than would be apparent from a consideration of the overall recorded crime statistics *only.*

8.3.2.3.2 Supermarkets

Supermarkets are one of the facility types to be noticeably affected by categorisation issues, which impact on the number of offences attributed to each address, and consequently to what is selected as risky and their rank position. The breakdown of ranks and crime count for *supermarket* are not shown, as although this proved helpful in furthering the comparison, supermarkets identified in the overall risky facility list have been used for the remainder of this section, unless otherwise stated, as this is considered more accurate.

There were 13 premises identified as overall risky facilities that could be considered to be supermarkets (including big box stores that might contain concessions). This excluded premises where the address was shared with a number of other shops, in a dedicated shopping (and sometimes entertainment area). Such an arrangement covered all the Food Lion supermarkets, therefore it was not appropriate to count them here, though they are discussed further below. Having separated the data in this way, only four of the 13 risky supermarkets were mentioned in interview. These were the two Walmarts, a Harris Teeter, and a smaller grocery, Superior

Foods.

The most obvious issue to consider is why so few 'risky' supermarkets were mentioned in interview. There are two Walmarts in the study area, both are risky facilities by my definition and both were identified.

> I marked down here at the bottom (5), that's South Main…that's where the Walmart is, that's where we have a lot of shoplifting, lot of problems down that way with property crimes more so than our violent crimes just because it is a more commercial district. For the shoplifting purposes that happen down there and the Kmart [categorised as *retail other*] is also down there, so it also is probably one of our more problematic areas for that.
> I'll mark this up here (6), which is our North Main area…that's where the Walmart is for the main street up there and every day it seems like there's either one of those two places, more so North Main, there are tons of shoplifting, they come out of there. (R1)

Some respondents thought, like respondent 1, that the Walmart in the north of the city was the most problematic, whilst others claimed it to be the one in the South. Based on overall facility crime (rather than facility-type categories) the North Walmart is the greater crime contributor (n=1156, 3.76% of all facility crime). In fact, it is the second most risky of all facilities in the study area (after the large, but now closed down OH Mall), whilst the South Walmart was the seventh largest crime contributor over the period studied. Differences in opinions of which is the more problematic may be related to officers' personal experiences of dealing with crime at these locations (amount or frequency) or it may result from perceptions regarding the type of offending that takes place here (and by whom). Interestingly, further exploration of these premises revealed that the South store did not open until 2007, meaning it has recorded around 152 offences per year over the three years it was open that coincided with the study period. In contrast, the North Main store was operating for the whole period studied (in fact it was built in 1991), which equates to an annual offence average of around 116. This shows offending at the two premises is actually quite similar, with possibly more offences at the newer, South store.

The city also has a number of Dollar General stores, three of which are risky facilities, but none were mentioned by respondents. It is unclear if this is because they simply did not think of them, or because they thought them not to be risky. Further exploration revealed that one of these appears to have closed down around 2006 (which tallies with the crime data, as no offences are recorded after financial year 2005/06) and the other around 2007 (again this fits with the recorded crime data). As the interviews were carried out in 2011, these stores were likely not mentioned because they had not been in operation for three to four years. There is no apparent reason why the other store was not mentioned, though as these are generally thought of as 'discount' stores, there is the possibility that respondents did not think they would be targeted. Although this appears to be erroneous in these two cases, there are currently five Dollar Generals in the study area, and the others were not identified as risky, in the crime data, so officers may, generally, be correct.

Reflecting on my observations, although this location is relatively close to the central part of the city, it could be considered to be towards the edge of the more densely populated and commercial areas. Travelling westerly from the centre, the streets begin to be less lined with retail outlets and fast food restaurants, feeling more open. There is more landscaping and foliage and properties along this thoroughfare are more spaced out. Also along this road are regional religious headquarters, law firms and nursing homes built like manor houses. It is certainly the case that this area feels very different from the south and central locations, and even the more commercialised north/central areas. As with many North American towns, the city is mainly built for the car driver, and this is certainly one of those locations that feels inaccessible without access to a vehicle.

Clearly offences still occur at this address, as it is flagged as risky from the recorded crime data, but it does not seem to be an area that the respondents were either particularly familiar with ("I don't normally work that area so I don't know" (R5)), or one that they perceived to be problematic in terms of crime. Further south along this road, the area becomes more commercialised again and more industrialised (where copper thefts were talked about, but not in relation to premises that would be defined as facilities). This is when it is then mentioned again, in relation to retail crime.

Given the nature of the area, the higher recorded crime might actually be a reflection of the willingness to notify the police when offences are carried out, possibly even because they are seen as less a normal part of trading than might be the case in stores where shoplifting is more rife. Of course, there are many other possible explanations for why this particular store might be associated with more offending than others, and these are the sorts of issues that are explored in much of the risky facility research to date, however that is not the purpose of this study. Rather, here I restrict myself to considerations of why this location might be inaccurately identified as risky in the crime data (higher reporting rates) or as not risky in interviews (not thought of as the type of place where there is a lot of crime).

Of the remaining unidentified risky supermarkets, a further three had closed down prior to the interviews (one of which had contributed a substantial proportion of crime in this category). These were Cloverleaf Grocery (which closed in 2010) and Bi-Lo Inc (closed in 2006). It would make sense that these were not highlighted as currently problematic, although given the amount of crime Cloverleaf had experienced, and the fact it had fairly recently closed, it might have been expected to be mentioned (as other historical problem locations were). That said, this premises, which I visited during my observations, but obviously could only view from outside as it was already closed down by then, was situated in the vicinity of two large pharmacies, that were only mentioned in passing by one respondent (and this just to help explain the location of library), though they too were flagged as risky in the crime data. It is possible this area was not seen as problematic, though it is unclear why. Other than mentioning the library and the homeless shelter, very little was mentioned about crime in this area which has the hospital (not identified in interviews as risky) to the west and the university (also unidentified) to the east. Again, this could possibly be because these premises were mainly targeted at car drivers (fronting directly onto a main north-south thoroughfare, with no or little sidewalk, so were not subject to local, walk-up offending (as was the case further south) but they were perhaps not seen by the police as containing desirable enough targets for commuting offenders (who would more likely 'hit' the Walmart further north on this route).

Although information regarding the closed down store was available through analysing the crime distribution for that premises over time, and Internet searching, unless this was done for every (risky) facility, this closure would not have been known about had I not been trying to identify reasons for it not being mentioned (particularly, searching on Google Maps to verify the location). Local practitioners will, of course, know of major closures or moves such as this, especially when they involve high crime locations. However, this is not the case for researchers, who are much less likely to have 'on-the-ground' knowledge. This highlights the importance of looking beyond just the basic crime date when exploring risky facilities. It also illustrates the impact that different time-frames can have on outcomes, with longer periods identifying more stable patterns, but potentially lacking currency. This is even more important when studies use datasets taken from different time periods, for example census, environmental or tax parcel data that is not contemporaneous with the crime data.

This leaves four unidentified premises. Only one of these, Lowes Foods (as distinct from different Lowes Home Improvement stores that were situated elsewhere), was flagged as risky within the categorised data. The exclusion of this premises by respondents is very likely because it is situated in the far north of the city; in fact it is listed as being in a neighbouring town. Thus this is a further example of 'edge effects' in relation to perceptions and mental maps.

The remaining supermarkets were small or non-chain retailers and experienced only 37 and 36 offences each. It is likely that these were not mentioned in the interviews because they were not actually that problematic, suggesting that the selection method applied to the recorded crime data may have been over encompassing.

Finally, there were a number of Food Lion stores, often listed as supermarkets, but also as *retail other*. There were six of these flagged as risky facilities, in both the *supermarket* and the *all* analyses, though the crime counts were expectedly higher when aggregated on address only (rather than on location type). Another reason crime counts were higher in this case was because all of these stores shared addresses with other premises, being situated on strips or small open shopping centres/plazas (to which the address was assigned, rather than to individual units). For this reason, the total crime counts were clearly not all attributable to the Food Lions. When specific named premises were extracted from the interview data for the content analysis, four of these stores were identified as risky. However, the general shopping areas within which the other two stores were situated were also identified in the interviews, therefore it is safe to conclude that the respondents were aware that all of these premises and/or the areas where they were located, were associated with high incidences of crimes. Providing some degree of validation for their selection from the recorded crime data as well.

Taken all together, it may have seemed in the first instance that officers' perceptions of high crime supermarkets did not match what was quantitatively identified through the recorded crime data, with few such premises being highlighted in interviews. However, further consideration of the data and the facilities themselves has shown that the risky facilities identified in both ways were very similar when considering locations of current concern.

8.3.2.3.3 Convenience stores

When risky facilities were selected from the categorised datasets there were 28 convenience stores identified as risky, and these were a mixture of both the neighbourhood convenience store and convenience/gas station combination types of premises. For *petrol* there were 19 risky facilities. Many addresses appeared (under-counted) on both lists. When premises were instead aggregated first by address and overall risky facilities were selected, there were 42 that I categorised as convenience/gas store combinations and a further (different) 17 neighbourhood convenience stores. For the remainder of this section, these two new categories are used.

Convenience store/gas station combinations

This type of premises was rarely discussed in the interviews, with just six of the 42 risky facilities of this type identified from the facility crime dataset. With so many not identified as problematic by the respondents, there is no obvious pattern in either the location of such premises, or the businesses represented. It is possible that there are so many gas stations in the city, that it is difficult to distinguish (particularly from memory) those ‘few’ that are more problematic than the rest. That said, some of the respondents were able to do this for some premises and in doing so they highlighted two different problem types. Firstly, those gas stations on or near the city limits, which tended to suffer beer thefts

> Gas stations. The ones that border the city limits. It’s not a constant problem, but every now and again I’ll get one, like in the Beat three area there’s a Wilco…and it is literally just right on the city limits and it’s that one that really gets hit the most and it’s basically out-of-city offenders coming in…[U]sually what it is, maybe a soda or two but usually it’s alcohol-related shoplifting and that is pretty much all hours of the day. It’s a 24 hour store… (R3)

The journeys taken to offend, and explanations regarding the locations of problematic premises are considered more in chapter 9. However, it is interesting to note here that respondent three identified another gas station, problematic for the same type of offence, that was also on the city limits (a BP garage), but this one did not meet the criteria to be selected as a risky facility from the recorded crime data (being associated with just 29 offences over the ten years). What also stands out, is that for supermarkets and educational establishments, I proposed that some of the facilities not selected were those towards the borders of the city and that this might be because officers were less familiar with locations on the edges of their mental maps. However, that is exactly the locations that were being talked about by this respondent. This raises the possibility that officers may have more accurate perceptions of crime occurring towards the edges of their mental maps (and/or the boundaries of their jurisdiction) when being in such a location (or on a specific route/pathway, here, for example, the highway interstates to the south and the northwest of the city) is what creates the opportunity for crime or makes something a more attractive target.

The second type of problematic convenience store/gas station combination was those near residential areas with high rates of offending, truanting or gang activity. In this case, it was because the shops were used more like neighbourhood convenience stores (as discussed below).

> …this is another little store over here…Like it’s a gas station and convenience store, a lot of the kids [pause] from the apartment complex right over here…So it kind of goes hand-in-hand because a lot of those kids are like selling drugs and assaulting people

> and stuff over there and they just walk straight across the street from their housing complex or their apartments to the little store. It always seems to be a little hotspot. (R5)

The distinction made between the two types was also apparent not just in my own observations of the city and the distribution of certain types of facility in it, but also through my ride-along and accompanied observations, when I visited many neighbourhood stores, and gas stations, and was told of all the problems that occur at them. It became apparent very quickly that the offenders involved (and the crimes they committed) were very different depending upon the location and type of gas station, but that those near concentrations of offender residence were viewed more like convenience stores (that happened to have fuel pumps) and, in fact, in a very similar light to neighbourhood convenience stores. It is clear, therefore, that although the original method of categorising gas stations (like the first example) and convenience stores (like the second example) separately may not have been entirely successful in practice, the conceptual distinction was appropriate and seems to be reflected in police perceptions as well. Therefore, I now turn to neighbourhood convenience stores (whilst noting that some of the patterns discussed in relation to such premises, both here and in chapter 9, also related to the second type of gas station discussed above).

Neighbourhood convenience stores

As stated, convenience stores featured heavily throughout this study. Though I had some communication with the PD regarding my initial analysis of the data provided, which may have had some influence on their subsequent behaviour and discussions, I was readily taken around to see premises of this type and it was clear that they were perceived to be extremely problematic. Indeed, one of the first places I was taken to visit was a very rundown, but typical store, in one of the most deprived, high crime areas of the city. As I was with non-uniformed staff at the time, we waited for an armed, uniformed officer to accompany us inside, so that I could have a look around. It was no surprise, then, that nine such premises were identified through the interviews, including eight of the seventeen that were deemed risky from the all facility crime dataset. Leaving one (identified by respondent 2) that was not risky according to the crime data (being associated with only 23 recorded offences).

There were no obvious patterns in the risky facilities (according to the crime data) that were not identified in the interviews. ‘High crime’ neighbourhood convenience stores were all of a fairly similar style and set-up, small, independent, often somewhat unkempt, and they were all located in the south and south/central areas, which given the populations they serve and the nature of these premises is again not surprising. The premises that were identified by the respondents were more likely to be the higher crime contributors, but this was not always the case, therefore it is possible that those mentioned were not necessarily substantially more problematic than those they did not pick up on, but may have stuck in the mind more, because of particular incidents officers had dealt with, because they knew the owners, which my accompanied observations would suggest was true, or because the address had a particular reputation because of its location and/or things that had happened there (for example one of the respondents talked about the owner of a risky convenience store having been killed (and I was also told of a homicide at one of the neighbourhood-type convenience/gas station combinations whilst on an accompanied observation)).

> Another one I just thought of…is a local convenience store…it's right here at this corner…right outside one of the housing projects that was torn down and has been rebuilt. But that is a place where numerous robberies, I think it's called Tommy's or One Stop Tommy's. They had a homicide there, the guy who used to own it he was killed there. Numerous robberies, shopliftings and that one was a real hotspot for us because it was right beside that housing project… (R1)

The premises in question was typical of those in this category standing alone with a relatively large parking area (though many of their customers will be 'walk ups') around it, often near scrubland, empty or dilapidated buildings, or on street corners, as shown in figure 5.

Speaking of the same premises, Tommy's, R4 identified it as: "…in the middle of 'Shitville'". Many of the convenience stores talked about in the interviews were also discussed in relation to drinking in the parking lot and open air drug markets, involving locals and those travelling into the area for this purpose. These are discussed more in the next chapter.

It might also be the case that the premises identified by the interviewees were the most problematic, but that some of the properties were under-represented in the recorded crime statistics because of a lack of reporting or of co-operation (as will be seen in chapter 9, some of the place managers of these types of facility are thought to be themselves involved in crime, or at least turn a blind eye to things such as drug dealing taking place on their premises). Also as proposed above, they could be more problematic than is apparent in the study data, because of offences associated with those addresses but not recorded as occurring *at* them (for example, they may have been coded as in a parking lot, or the street).

Figure 5: Tommy's, a typical walk-up neighbourhood convenience store

Though several of the risky convenience stores were not perceived (or at least remembered) by the interviewees to be problematic, those facilities identified using quantitative methods applied to recorded crime data and those highlighted by police officers during interviews were very similar types of premises and nearly half of those selected from the crime data were also identified by the respondents, more so those with the higher crime counts. In addition, some of the stores not identified in interview were pointed out to me by other officers during accompanied observation. Therefore, it seems clear that this type of premises is problematic in the study area, and that either method of identifying which addresses are the most risky is likely to produce similar, overlapping lists. If using only recorded crime data, the list of risky facilities is likely to be more comprehensive (and longer), but the insight that can be gained through qualitative methods is also invaluable. More findings drawn from this will be discussed in chapter 9.

8.3.2.3.4 Entertainment locations

Before moving on to summarise the main issues arising from the analyses reported in this chapter, brief mention needs to be made of this final group of facilities. This broader group consists of *licensed premises* (as already discussed), *restaurants* (from fine dining, to sports grills, to fast food), and *leisure* (which covers such premises as sports facilities and cinemas). There is not the scope to consider all of these types of premises separately and, in any case, the number of identified risky facilities, and their crime contributions, are relatively small (in light of the other categories discussed above). However, there are a couple of interesting findings that are worth consideration.

In interview, the only fast food premises that was specifically identified was a McDonald's, situated at a confluence of several roads that had become a known 'gang' area and hang out. As has been highlighted throughout, gang activity was a particular concern of respondents, so that locations and facilities seen to be associated with, or affected by, gangs were more readily and thoroughly discussed. Overall risky facilities included a further two fast food restaurants, with no obvious reason why these were not identified, though the crime counts for both were low (52 and 33). The categorised selection of risky facilities identified many more restaurants as being risky (18), a further two of which were mentioned in interview seemingly because they were part of wider entertainment complexes that were known to be targeted by offenders (particularly for thefts from vehicles in car parks).

What is also interesting, is that in identifying such locations, four respondents talked about problems in and around the Palladium Movie Theater. This was part of one of these 'entertainment' and retail complexes, in the north of the city, that was seen as a significant crime problem, as a result of car theft, as mentioned, but also general disorder, fights and drunkenness, often involving teenagers. Explanations that were proffered regarding this are considered in chapter 9, below, but it certainly seemed to be the case from the way the respondents talked about this area, that it was very resource intensive, which may well explain why so many of them thought to mention it. However, this highlights an issue with using only 'within' facility crime to identify risky premises. The Palladium was not listed as a risky facility in any of the categories, or overall, as there were only 16 offences recorded at this address. Therefore, it seems that this cinema did not experience a great deal of crime itself, or at least if it did this did not translate into many *recorded* offences, but the area in which it was located *at least in part because of its existence* was indeed a crime and disorder 'hotspot'.

This finding supports approaches that employ buffers around facilities (or use street segments or hotspotting techniques) to identify problem area, then looking at the premises falling within the designated area to see what might ‘cause’, generate, attract or enable crime. On the other hand, it supports critique related to determining which crimes in an area only occurred because of the premises in question and, further, when there are multiple premises what influence did each of these have, individually and together. In light of this, I would propose that to further explore the nature and explanations for crime concentrations in microgeographic units, a combination of within facility, buffered facility/hotspot *and* qualitative research is required. Alone, none of these seem capable of revealing all the patterns and issues that need to be explored. On this note, a final summarising discussion of the findings from this chapter is now required.

8.4 Summarising discussion

Given the format of this chapter, the implications of most of the findings have been discussed as they were presented. However, there remain some issues that cut-across the sections, which must now be discussed.

It has been shown that there was a substantial amount of overlap between the premises that were identified as problematic by police officers and those that were selected, using my preferred approach, from the data. As would be expected, a greater number of premises were deemed risky from recorded crime data, as officers could not be expected to think of or recall all possible candidates during a single interview. However, there were some high crime addresses (notably the top ranked school, but also the regional hospital and the university) that one would have expected officers to recognise as problematic, and there were some types of facility that were either not, or only rarely, mentioned despite experiencing not inconsequential amounts of crime (pharmacies and financial establishments in particular). Perhaps even more importantly, there were a small number of premises that police perceived to be risky, but had not been identified through analysis of the recorded crime data. There are a number of implications of officers not identifying the same problematic premises as an analysis of the crime data.

For the purposes of this discussion, if we take problematic premises as either truly of concern, or not (true problem and false problem, respectively), there are four possible iterations of discrepancy between risky facilities identified in the crime data (using my selected method) and those identified in interview. The implications of each of these, assuming the perceptions gathered from the interviews in this study to be representative of wider perceptions in the PD, are considered in turn.

When there are truly problematic premises identified by the crime data but not by police officers there is the possibility that some locations that warrant attention might not receive this. When deployment decisions are made, or if calls for service are prioritised due to high levels of demand, such misperceptions could divert officers away from places that are already problematic (but not seen as such) or are on the brink of becoming so. This could similarly result in such facilities being excluded from reduction initiatives, or receiving a smaller proportion of patrol officer time. This could have the impact of allowing crime to continue to increase unchecked.

For example, if officers do not really think of the university as part of the city, perceive it as a relatively unproblematic and safe location, or generally leave it for campus security to manage, then there is the potential for all sorts of undesirable behaviour to be left unchecked. Given ongoing concerns regarding the handling of sexual violence on campuses (The Whitehouse, 2011 *cf*. Guardian online, 2017;) and the possibility of under-recording, thus reporting of campus crime as required by the Clery Act,[16] (Clery Center, 2018) it seems important that local police departments have an accurate understanding of crime and disorder taking place on campuses that they have overall responsibility for, or even those situated within their districts that they do not directly police.

Officers also seemed to feel the High School with the highest amount of recorded crime was not that problematic, especially in relation to other schools in the area. It was proposed this could be because the school was towards the northerly border of the city, and/or that the *nature* of crime committed was not thought to be as serious, or related to other more serious types of offending - such as gang activity - resulting in the establishment being discounted as a problem. Though the latter explanation could well be true, this school still experiences high levels of crime, but if not seen as a problem, then it may receive less police attention: less time building up relationships between the school community and the PD, less time spent involved in youth intervention work, less investment in crime reduction work, and so on.

This problem also applies to whole facility groups, and the misperceptions regarding what *type* of premises are substantial crime contributors (as opposed to which specific addresses). Though the facility categories that experienced the most recorded crime were represented in the premises discussed in interview, other facilities in this case pharmacies and financial establishments were not discussed. It has been proposed that this may, in part, be because of the type of offences committed at such locations. However, although some convenience store/gas station combinations were mentioned, these were the most numerous of the overall

[16] Jeanne Clery Disclosure of Campus Security Policy and Campus Crime Statistics. 20 U.S.C section 1092 (f) 1990

risky facilities, and most of them were not identified as such in the interviews, whilst a greater proportion of neighbourhood convenience stores were. It is worth noting, then, that there may be a tendency for police officers to focus on types of location they think of as stereotypically problematic, when there may be others that do not receive such attention. This also suggests that even if a particular address is problematic (a good example here would be the hospital) it may not be at the forefront of officers minds because it does not fit into one of the more typical problematic categories.

This point further reinforces the argument that research into facility crime concentrations needs to focus much more on the concentration of crime across premises into the few risky addresses (arguably regardless of facility category), than to consider crime contribution by whole facility category. In other words, for intervention and deployment purposes as least, the risky facility concept is more useful than that of crime generators (or attractors) - as this term was originally used (I discuss this much more in chapter 10). Of course, thorough analysis of the crime data should help ameliorate this, but police officers' individual perceptions could still result in them making decisions whilst on (directed) patrol that take them away from the areas that most need it.

When there appears to be a problem in crime data but not to police officers, and it is a false problem, this suggests weaknesses in the methods of selection, or that crime is not sufficiently high (or possibly concentrated) enough, to warrant including this type of facility in any analysis. It is also possible for an address to appear more risky than others because, for some reason, a greater proportion of the crime occurring there is reported or recorded. This could again result in unnecessary or inappropriate interventions, but in this situation quality analysis is less able to identify that this is the case. This could have negative effects on the relationship between the police and place managers, who may then start to report less crime. However, if analysis *did* reveal such a pattern to exist, and police were adamant that in reality this was not a problematic location, it could suggest that there was substantial under-reporting of offending at other establishments (in the same category). This brings us to the third possible reason for discrepancies between police perceived, and recorded crime identified, risky facilities.

When premises are truly problematic and this is recognised by the police, but not reflected in the crime data, this is likely, then, to be due to under-reporting and/or under-recording. If this is the case, efforts need to be made to gather information that paints a more accurate picture, and if possible to encourage more reporting. Doing this should result in more appropriate responses and the actions taken may even help reduce crime (as a result of the better relationship between the police and place managers). Of course, problem premises that are not well represented in the crime data, might also be perceived as problems because of non-crime issues. These may still be of concern to police officers. They could be addresses associated with large numbers of calls for service, and so still be resource intensive. They could also have a negative impact on fear of crime or quality of life for local residents, or those who work nearby. Finally, they could be seen as a problem not because crime occurs there, but because offenders frequent the establishment, or because offences committed elsewhere are thought to have 'started' or emanated from these locations, as may often be the case with licensed premises (bar) related crime. This type of discrepancy seems (though it cannot be confirmed without further exploration) to apply to the cinema, discussed by four respondents in interview, but only experiencing a relatively low amount of recorded crime. This premises and other entertainment establishments around it (such as restaurants) seemed to generally experience high levels of disorder, that appears not to have been assigned to any particular address, to be spread across several premises or to end up as crime committed as well. The same was said regarding some other leisure facilities as well.

This type of discrepancy could also be the result of inappropriate (or lack of) categorisation methods, resulting in some premises being obscured by higher crime facilities that have been analysed alongside it. For example, there are very few restaurants in the overall risky facility list, because the recorded offences associated with each are relatively low. However, when only this type of premises are analysed, more are selected as risky (in the context of the distribution of crime across restaurants only), including two that were identified in interview.

Whichever the reason for this difference, and more than one of those proposed may be true, it is important to think about how to respond to such issues and, crucially, how to capture police knowledge so that this type of risky facility is not 'missed' in analysis and intervention-planning. It is also important to further consider how researchers could use qualitative approaches more, in order to include this type of insight in their studies, or to at least consider sources other than recorded police data, such as from other agencies (health, fire, local government) or collected from police in different ways (e.g. questionnaires, annotated maps, and so forth). This is considered in more detail in chapter 10.

Of course, deciding between whether such discrepancies are the result of 'true' problems (the police have correctly identified that this address is problematic, but it has not been picked up through analysis of the crime data) and which are actually 'false problems' (the police think this particular place is a problem, but the recorded crime data is 'correct' and they are in fact mistaken) would not be an easy task, and again multiple data sources are likely to be required.

Finally, facilities that are not risky in the crime data, but are perceived to be such (falsely) by police are likely to come about because of reliance only on limited experiences, or hearsay, being influenced by historical problems (that are no longer current), or stereotyping. This last reason might include stereotypes regarding categories of facility or specific addresses. More concerningly, these could be related to particular (types of) location or people. In the current study there was a definite focus on ‘gangbangers’ and regular mention of housing projects and ‘crap’. Although most of the places identified in relation to this type of explanation actually *were* risky in the crime data as well, this certainly seems to at least partially explain the extent to which neighbourhood convenience stores were discussed.

This type of discrepancy may lead to a focus on, or inclusion of, places that is not warranted, thus wasting time and resources. In turn, this could be harmful to police-citizen relationships (with all the attendant implications of this) and, in more extreme cases, it could result in discrimination of stereotyped groups, such as ethnic minorities. Again, using carefully produced analytical products and ensuring officers are well-trained, and managed, will reduce the chances of inappropriate attention being focused on individuals, groups or facilities, but it is important that in doing this, the benefits that on-the-ground (police) knowledge can bring, as discussed above, are not completely lost.

There are a number of further, general issues that need to be briefly considered, before bringing this part of the study to a conclusion. Categorisation, methods of categorising and size and homogeneity of categories have all been recognised as having an impact on what premises are identified from the recorded crime data as risky, thus the degree of similarity with police-identified address. This was also seen in practice when the categorised versus overall risky facilities were compared.

Another research decision that might have an impact is the time period used to identify risky facilities. Using longer time periods, as already discussed, is likely to result in a selection of risky facilities that are more stable, than if a shorter period is employed. However, what is identified as problematic over the last ten years is not necessarily the same as what is currently problematic, or what police officers are able to recall (or, if they have less years of service, what they even know about). The fact that several premises included on the recorded crime risky facilities list had closed down (which is a plausible reason for why they were not highlighted in the interviews) is evidence of this.

Finally, it is important to note that even when police officers did not identify specific addresses, on at least some occasions, particularly for retail establishments, they *had* discussed the general area in which they were located. Overall, the data obtained from both approaches was actually quite similar. Although the respondents might have thought some addresses were worse than others when the recorded crime data ordered them differently, the qualitatively obtained data could be viewed as a smaller subset of the crime data, other than for a few premises. This adds weight to the appropriateness of my proposed risky facility selection method, whilst also suggesting that it might have been somewhat over-generous in the number of premises identified. However, the benefits of the qualitatively obtained data are not just limited to this. They also helped add to my understanding of the types of premises in the city and the type(s) that are generally most problematic. They prompted further consideration of why places were viewed the way they were, or why they were not mentioned. They were also useful for beginning to explore reasons why some premises experienced much more crime than others, what types of offence were most likely to be committed and what offenders, or types of offender were likely to be involved. This study does not seek to explain why risky facilities exist, or why some premises in an area experience much higher crime than others. Indeed, this has been the focus of much of the literature and I have deliberately taken a step back from this. However, having settled on a working definition of risky facilities, the next chapter starts to explore some of the features associated with the selected premises: the locations in which they are situated and journeys-to-crime.

8.5 Conclusion

This chapter sought to answer research question four: Can risky facilities be defined empirically? Through an exploration of a number of possible quantitative methods and a comparison with qualitatively collected data, the answer to this question is yes, but further research is required to assess the suitability of the proposed method.

I have argued that selecting an empirical method for identifying risky (versus non-risky) facilities is an important endeavour, and then gone on to test several possible ways this could be done. In comparing these, it was difficult to determine which performed ‘best’, given that there is no reference for when the risky facility label should, or should not be applied, either in terms of the extent of inequality required in the crime distribution across premises (as was discussed in chapter 7) or in terms of where the cut-point should be set. Both of these decisions are discussed further in chapter 10, and constitute a significant, original contribution of this piece of work.

The method selected for defining risky facilities, based on the findings presented in this chapter, was that they be considered those premises that contribute greater than or equal to three times the mean crime count for all premises (experiencing at least one offence, over a given period of time - this will be discussed more in chapter 10). This method was conceptually strong, was relatively simple to calculate, and tended to select a number of premises that fell between the extreme highs and lows of some of the other methods. Limitations of the method were identified, notably that it is sensitive to decisions about whether to include or exclude zero-crime facilities (and in turn, therefore, the time period selected for analysis), and it was also recognised that there was equal merit in the alternative iteration of four times the mean. Further research in relation to this, and the use of more sophisticated statistical methods for selecting risky facilities (as comparison, rather than alternative) is suggested in chapter 10.

Finally, a different approach to identifying risky facilities was also considered, through content and thematic analysis of interview data and annotated crime maps, collected from a sample of serving police officers. It was found that there was substantial overlap between the addresses identified as risky using the two different methods, that recorded crime data tended to select a greater number of risky facilities and some of these were clearly high crime locations that the respondents did not identify. A number of reasons for this difference were hypothesised, but the most convincing in the absence of further evidence were that officers were less likely to mention problematic premises towards the edges of the city (and possibly their awareness spaces), in the richer, better class locations in the city (though there were exceptions), and focused heavily on retail establishments, with some mention of schools and entertainment/leisure facilities. They also tended to spend a lot of time talking about more general areas, including those seen to have a high proportion of offender residences and other signs of disorder.

It was also identified that the qualitatively obtained findings highlighted as risky, a small number of further locations that might otherwise have been missed, as well as providing a greater degree of insight into types of premises than could have been gained by analysis of the recorded crime data alone. Thus research question 4.3 is answered, and as a result I propose a greater use of qualitative and mixed methodologies in crime and place research.

Chapter 9: Study 3: Risky facility crime journeys

9.1 Introduction

This chapter presents the study that sought to address research question five: *Where are risky facilities located and what types of crime journeys are made to them?* This was broken down into four sub-questions that considered: the distributions of risky facilities across the study area, from both recorded crime and police perceptions (RQ5.1); the distances of journeys-to-crime associated with risky facilities, compared to non-risky facilities (RQ5.2); whether distances of crime journeys differed by facility type, crime type (all facilities only), and time of offence (all facilities only) (RQ 5.3); and operational police officers' perceptions of journeys-to-crime at risky facilities (RQ 5.4). Research questions 5.1 and 5.4 are supplemented with findings from my observations.

In the next section, the methods used to answer the research questions are set out and critiqued. The results are then presented. In order to incorporate discussion and draw further meaning from the qualitative findings, research questions 5.2 and 5.3 (based on statistical analysis of journey-to-crime data calculated from police records) are presented first. Next the results relating to the locations of risky facilities across the study area are considered, by combining the findings of visual analysis of crime maps, with thematic analysis of the interviews with police officers (and their annotated maps), supplemented with thick description from my observations. In doing this, issues that address research question 5.4 are also introduced as officers' discussions about the locations of risky facilities, those of offender residences, and the journeys made to commit crime (and the reasons for this) were inextricably linked. In this section I also reflect back on the distance findings, in the context of the other material presented. It should be noted that although I use the term 'distance', and only calculate this feature of the journey-to-crime, unlike some recent authors (Andresen et al., 2014; Townsley & Sidebottom, 2010) I generally refer to the 'journey-to-crime', in keeping with the majority of the literature.

Finally, all the issues are brought together in a short, summarising discussion critically considering the types of journeys made to risky and non-risky facilities, the types of areas in which risky facilities are located and the areas that 'supply' offenders, as well as the possible relationships between these different factors. This is followed by brief, concluding remarks.

9.2 Data preparation and analysis

Specific issues relating to the data preparation and analysis that have not already been considered are presented here.

9.2.1 Journeys: data preparation and analysis methods

9.2.1.1 Quantitative analyses

The preparation of the offender address data, joining this to the recorded crime data, and geocoding the current address have already been set out in chapter 6. Having created and cleaned the dataset as described, it was necessary to calculate the distance from the offender's (current) home address to the crime location (facility premises). This was simply done using Pythagoras' theorem to calculate the Euclidean distance from the recorded Cartesian co-ordinates (X and Y) of the two addresses. Alternative journey distances could have been calculated, notably Manhattan or road network distances (which, given the relatively standard North American layout of the study area, would likely have been similar). However, the Euclidean (crow flies) distance was the easiest and quickest to calculate, and it was deemed sufficient for the purposes of this study (as well as being probably the most common method employed in the journey-to-crime literature). In particular, the journey-to-crime was to be used to compare the distances travelled to different types of facility and to compare crime trips to risky facilities versus non-risky facilities. Therefore, as long as the distance was calculated consistently, the method was less important. However, it should be noted that this does affect comparison with distances in other studies, if an alternative method has been used.

As offender addresses outside the study area are excluded this has the effect of skewing the results. Those who might travel long distances or travel into the study area from neighbouring towns and cities are not represented, therefore it is not possible to determine if such journeys are more likely for risky facilities or not. The exclusion also means that the maximum and mean distances travelled will be under-estimates, as they cannot fall outside the maximum distance from an address to the edge of the study area. This will be considered later.

Ultimately, the dataset produced for analysing journeys-to-crime, contained key information about the crime and offender, with one row for each crime journey. It also included information as to whether the crime occurred in a premises that was categorised as a risky facility or not. In addition, a further dataset was created by aggregating the crime trips to each address, so that each row was a different premises and the distance measure was the mean of all the journeys to that address. This was good for comparison and considering differences by address, but it did mean that some of the variation in trips was lost because not all the different lengths of journey were apparent. Additionally, this resulted in much smaller sample sizes, so although these data were analysed, virtually none of the inferential tests achieved significance.

A test of normality of the distribution of distances travelled was carried out for all journeys, all journeys to risky facilities and all journeys to non-risky facilities, for all facility crime and for each category of facility. In nearly every case, the distribution was shown to be statistically significantly different to normal (for an alpha of 0.05, using either the Kolmogorov-Smirnov test for sample sizes of 50 or greater, or the Shapiro Wilk test for samples of less than 50). In a few cases it was not possible to reject the null hypothesis that the sample was normally distributed, however these tended to be when samples were particularly small. This was the case for *leisure* overall (n=39) and for risky (n= 9) and non-risky facilities (n=30), *healthcare* for non-risky facilities (n=18) and *petrol* for risky facilities (n=96; though there was a statistically significant difference using the Shapiro-Wilk test). Therefore, non-parametric tests were used throughout, and the median is reported as well as the mean (Kiess, 1996). The Kolmogorov-Smirnov (K-S), and the below mentioned Mann Whitney Wilcoxon (MWW) tests have been introduced and justified in the methodology chapter (6). Following convention, and the standard outputs of IBM SPSS, the Shapiro-Wilk test is used as a test of normality for small sample cases. Whilst the power of all normality tests is low for small samples (especially those under 30), Shapiro-Wilk (which calculates departure from normality using the measures of skewness and/or kurtosis) is widely accepted to outperform the K-S test in these conditions. Indeed, attitudes towards the use of normality tests have changed over recent years, given further comparative testing, such that it appears the Shapiro-Wilk test is beginning to be considered the favoured, more powerful test of normality for both symmetric and asymmetric distributions for any sample size (Mohd Razali & Yap, 2011 *cf.* Hinton, 2014), though the more traditional approach is employed here.

As stated, one of the key aims of this part of the analysis was to compare whether there was a difference in the distance of crime journeys to offend at risky facilities and those to offend at non-risky facilities. This comparison was carried out on two different measures. Firstly, it was done for each journey-to-crime, so that each recorded journey is either to a premises identified as a risky facility or to a non-risky facility, and the analysis shows whether there is a statistically significant difference between either the median distance (using the Mann Whitney Wilcoxon test) or the distribution (using the Kolmogorov-Smirnov test). Secondly, the same comparison was carried out, but using the premises address as the unit of analysis. As many addresses were associated with multiple journeys-to-crime, the distance in this case was the mean (the subsequent analysis involving the MWW test and the K-S test was calculated from this mean). This introduces a further degree of error, in that the data for each address have all been reduced to a descriptive measure that 'evens out' variability in the data.

Overall, some of the results obtained were limited by the sample sizes. This was particularly the case when the premises address was the unit of analysis, for example there were only 13 leisure facilities with journey data, only 19 healthcare premises and only 24 pharmacies. In fact, for leisure facilities, when separated into risky and non-risky premises, there were only two risky facilities (similarly for healthcare and pharmacy facilities, there were only four risky premises). Although included in the testing for completeness, no conclusions can be drawn from the results of analysing such small samples. For some other categories, sample size was large enough to warrant consideration, but it is likely that the statistical tests would not have sufficient power to achieve the required alpha level, even more generous ones. Such results must be treated with caution. Because of the limitations of using the premises as the unit of analysis, conclusions are drawn only from the analysis based on the dataset of individual journeys and others are mentioned here only when they provide support or highlight a particular departure from the main findings.

Throughout this study, efforts have been made to consider the ubiquity of patterns of micro-spatial crime concentrations. As well as comparing across different facility categories, therefore, comparison was also carried out for crime type (acquisitive versus violent crime) and for time of day (daylight versus darkness). Because of the issues relating to small sample sizes, as already identified, this analysis (using the MWW and K-S methods already set out) was only carried out on 'all facility crime'. The identifier for whether a facility was risky or not was the 'all risky facilities' measure (RF_F1), not the 'overall risky facilities' measure. As the results for these were very similar in the previously described analysis, it was felt unnecessary to run both again here.

In addition to the analysis described above, following Bowers and Johnson (2011), the journey-to-crime distance distribution was also plotted for journeys to (all) risky and non-risky facilities so that these could be visually compared with one another, as well as to demonstrate the extent to which some journeys were substantially longer than the mean. This chart was produced by plotting the number of crime journeys of distance greater than or equal to *d* on the y-axis versus the distance *d* (m) on the x-axis, binned into 500m intervals.

Finally, journeys-to-crime were compared across facility categories to determine if the distances travelled differed by type of facility. This was done for all journeys and just for journeys to risky facilities. The analysis was carried out on a dataset constructed from appending the separate facility category datasets, excluding *retail other*. Thus the differences being tested were between each of the twelve main facility categories and the eleven others (that is to say there were 66 pairs of comparison). In order to determine if there was a statistically significant difference, the Kruskal-Wallis H (KWH) test was used. This is generally considered to be a suitable non-parametric equivalent to the ANOVA, using ranks to analyse one-way variance to establish if several samples are drawn from the same distribution, or if they differ (Kruskal & Wallis, 1952). As with the MWW and K-S tests described above, where the distributions of the samples are similar, the results of the KWH test can be said to demonstrate a difference in medians. Following the calculation of the output for this test, a post hoc pairwise comparison was used to identify where any statistically significant differences lay. The default Dunn-Bonferroni post hoc test produced by IBM SPSS was used (thus pairs identified as significant are only those showing as significant, for an alpha of 0.05, for the Bonferroni adjusted p-values). This ensures that the ranks used and the pooled variance implied in the KWH test are preserved, as well as adjusting for multiple comparisons (that is to say, it is not the equivalent of running several, independent MWW tests) (Seaman, Levin & Serlin, 1991).

9.2.2 Locations and police perceptions

The methods for collecting and analysing the qualitative data have been set out in chapters six and eight already. Here, I report on the findings from these analyses that specifically relate to areas of offender residences, journeys by offenders, likely offender populations and certain offender/offence behaviour related to journeys and target choice.

In addition, the interview data and annotated maps were also used to assess where police perceived their local offender populations to live. No quantitative or geographical analyses were carried out on these data as the areas in question could really only be thought of as estimates and the sample size was too small for the results to be meaningful. However, it did allow for further comparison. For reference, these maps are reproduced in Appendix 5.

9.2.2.1 Producing maps

The facilities for each of the original categories were mapped, to show those that were risky (using the RF_F1 method as settled upon in the previous chapter) and those that were not. In light of the findings of the previous chapter regarding the appropriateness, and accuracy, of these categories, maps were also produced for *retail other* and *supermarket* combined, as well as for *convenience* and *petrol* together. In addition, the overall risky facility list (N=157), which had been manually categorised from the recorded addresses was also mapped. Although it was not feasible to apply this alternative categorisation to the whole facility dataset, it did allow a consideration of the locations of risky premises as identified from the crime data that was somewhat more comparable to the interview data.

In addition, the offender current address data were also mapped. It should be noted that this only related to the addresses that had been linked to one or more crimes in the facility crime data set. Due to the time consuming nature of cleaning and joining the data, those records not associated with facility crime had been removed prior to this process. Whilst this reduced the data available for mapping, including limiting it only to offenders who had been processed for involvement in crime recorded at a facility, it did mean that the patterns were not affected by the addresses of those offenders who committed crime elsewhere. However, as offenders are generally not specialists, exclusively committing just one type of crime, this may have made little difference.

As well as producing point maps with the data described above, journeys-to-crime were demonstrated using polylines that linked individual offender addresses to individual facilities where they offended. Due to the quantity of journeys involved, for the larger categories of facility such maps are not particularly useful in terms of visual analysis or illustration. The most obvious pattern is that the lines tend to be inward pointing, but this is in part an edge effect caused by only coding those offender addresses that fell within the study area. However, for particular facilities or patterns of travel identified through the analysis of interview data, selected journey maps are presented for illustration.

Finally, limited hotspot analysis was carried out. This was applied to the offender address data (with a reminder of the caveat that by virtue of including only points that had been associated with facility crime, the results may not be representative of broader patterns of concentrations of offender residences). It was also applied to all facilities, as a limited proxy of premises concentration (again noting that zero-crime locations were not represented) and to only risky facilities, in order to see if these were concentrated in particular locations.

All mapping was carried out in ArcGIS® (desktop and online). To produce the hotspots, the default hotspot tool in ArcGIS Online® was used. This identifies statistically significant clustering, in this case of point data. This was based on point counts, not on the amount of crime at each premises (field values) because the aim was not to see where crime clustered (this was already determined by looking at the counts for the individual addresses to identify risky facilities), but rather to see if there were areas where facilities or offender addresses clustered, as this might help explain patterns in the location of risky premises and journeys to them. The Find Hotspot feature includes a number of optional parameters. These were left at their default settings, so a fishnet grid for point counts was applied, all points were included, and the data were not normalised (that is to say a denominator, such as population, was not used). The cell size was also left to be calculated as appropriate to the data (information on this feature is taken from https://doc.arcgis.com/en/arcgis-online/analyze/find-hot-spots.htm).

As already noted, there are a number of hotspotting tools and approaches, but the one utilised by ArcGIS Online is the Getis-Ord Gi*. This statistic is calculated for each cell. Hotspots are those collections of cells where high values, or low values are clustered, and is identified as being statistically significant when the observed summed values of neighbouring cells are sufficiently different to the expected value, based on all features. In terms of display, this creates a hotspot map made up of square cells, using colour-coding to identify those collected cells that are statistically significantly high (or low) at three different confidence levels (99%, 95% and 90%). When these are high (there were no statistically significant low clusters of cells in the data), this can be interpreted as demonstrating there is clustering of facilities or of (facility) offender residences in that area.

Having set out the methods of preparing and analysing the data, the findings are now presented and discussed, starting with the results of the statistical analysis of journey-to-crime distances.

9.3 Results

The findings of the analyses discussed above are now considered. The descriptive and inferential statistics relating to distance of journey-to-crime are presented first, followed by the findings from the mapping and qualitative analysis, which are presented alongside discussion of all the results for this chapter.

9.3.1 Journeys

There were 7,350 individual crime journeys (from offender residence, as described in the methods section above, to a facility premises) associated with the recorded offences at facilities upon which this study is based.[17] The mean distance travelled in these crime journeys was

[17] The actual number of records created was 7,353 as stated in the relevant methodology

3788m (SD=2405m), with a minimum of 0.3m and a maximum of 16941m. Using the rule of thumb that the data are considered skewed when the skewness statistic is greater than 2x the standard error of skewness, it is concluded that the distribution of these journeys is also highly skewed (skewness=1.036, SE=0.029), which is consistent with the results from the test of normality (D(7350)=0.061, p<0.001). This means that the data are asymmetrically distributed (Hardy, 2009). The median distance travelled to crime was slightly shorter than the mean at 3444m. Figure 6 more clearly demonstrates the spread of journeys taken, showing the number of journeys of *d* or more, such that the left-hand most point represents the total number of crime journeys (and the shortest distance travelled (as explained by Bowers & Johnson, 2011)). This shows that journeys are much more likely to be (relatively) short, but they are not substantially concentrated in the shortest distances. There are also quite a lot of trips that are greater than the mean, and a long tail showing a small number of journeys that are much longer.

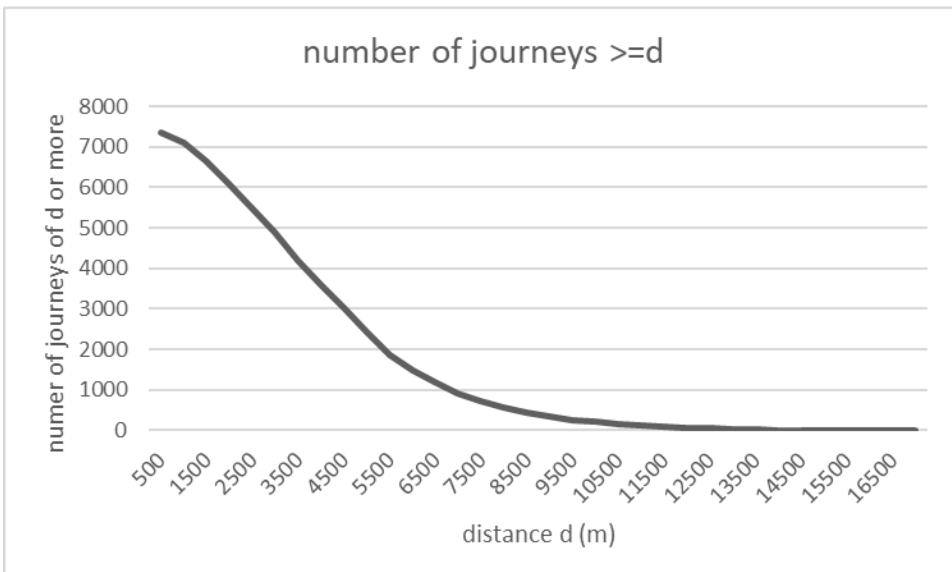


Figure 6: distribution of crime trips to all facilities

When the data were aggregated to individual premises, to produce a mean distance travelled for each, the results were similar. There were 805 premises with journey information (one or more recorded crime journeys to those premises in the ten-year period studied). The minimum mean distance travelled was 0.3m, the maximum was 15921m and the mean distance (noting this is in fact a mean of a set of means) was 3638m (Mdn=3161m). Again, the data were shown not to be normally distributed (D(805)=0.091, p<0.001). The average distances calculated using both individual journeys and premises aggregation are very similar, suggesting journeys to facility

section, above. However, upon analysis three of these were identified as having missing data, thus were excluded.

crime committed by those living within the study area, tend to be around 3 to 4km in length. Some journeys are clearly very short, whilst others are much longer, the maxima likely being curtailed by the constraints of only including those within the study area (which was roughly estimated in the mapping software as being just over 20km at its longest diagonal).

Using the crime journey dataset, the distances travelled when the journey was to a risky facility and the distances travelled for journeys to non-risky facilities were compared. This was done for all facility categories, in the first instance, so the two different forms of risky facility identification were used (all addresses categorised as risky facilities for their own type of premises (all risky facilities) and those addresses that met the risky facility criteria when premises type was ignored (overall risky facilities)). The mean distance travelled to all risky facilities was 3893m (N=5503, SD=2380, Mdn=3548m), whilst the mean distance to non-risky facilities was 3477m (N=1847, SD=2454, Mdn=3010m). For overall risky facilities the mean distance travelled was 3832m (N=6102, SD=2390, Mdn=3496m) compared to a mean distance of 3575m to non-risky facilities (N=1248, SD=2467, Mdn=3141m). In all cases, the data were determined not to be normally distributed.

The number of journeys of *d* or more were once again plotted, to consider if there was any apparent difference for ('all') risky and non-risky facilities. Figure 7, shows both types of crime trip on the same y-axis. This shows there were more journeys to risky facilities than non-risky facilities, but that this difference is greatest for shorter trips. Figure 8, alternatively displays the data on two different y-axes (one for risky and one for non-risky facilities). This removes the ability to contrast the extent of journeys, but it allows us to more clearly compare any difference in the shape of the distribution. Here it can be seen, in contrast to what might be interpreted from figure 7, that journeys to non-risky facilities appear to be slightly more concentrated in shorter to medium distances than those to risky facilities. The same charts were produced using the 'overall risky facility' categorisation and the patterns were very similar, although less pronounced.

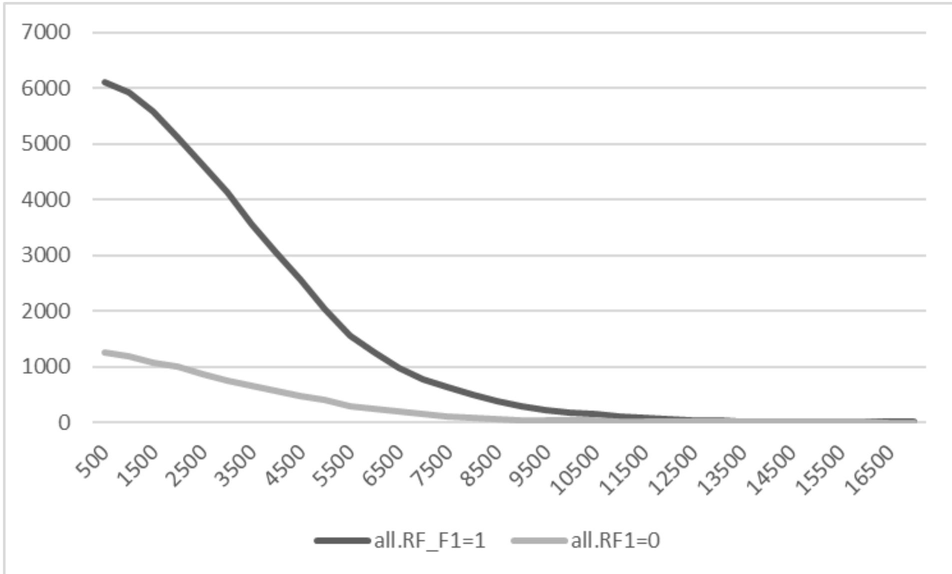


Figure 7: distribution of distance of crime journeys comparing risky and non-risky facilities

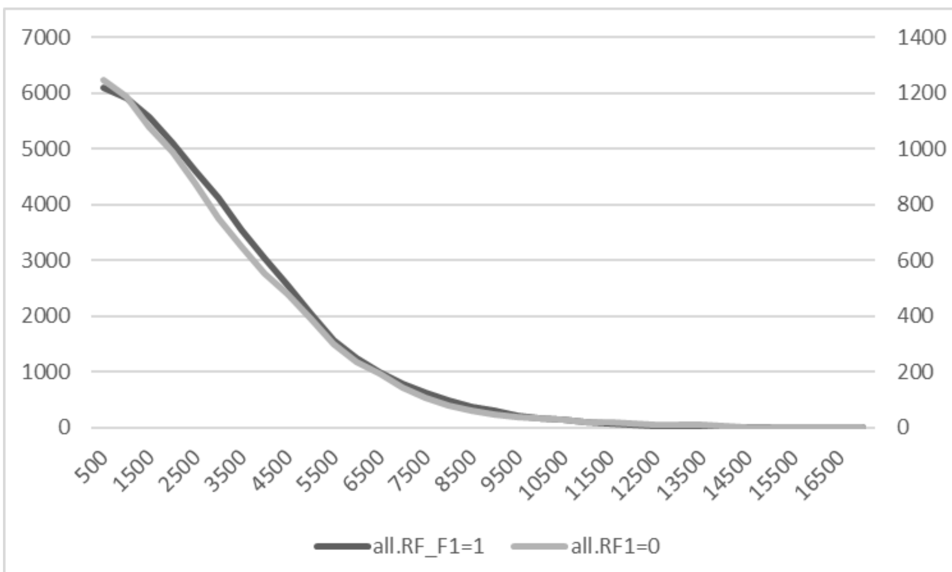


Figure 8: distribution of distance of crime journeys comparing risky and non-risky facilities (two-axes)

For both types of categorisation, there was a statistically significant difference in the distance of journeys to risky facilities and non-risky facilities. More specifically, the K-S tests showed that the two samples (journeys to risky facilities and those to non-risky facilities) were drawn from different distributions, and the MWW tests showed that there was a statistically significant difference between the ranked means of journeys to risky facilities and journeys to non-risky facilities. This was the case for both types of risky facility categorisation, and in all cases significance was at the 0.01 level. Also in both cases, this difference was that the journeys to risky facilities tended to be longer than those to non-risky facilities. The outputs relating to these findings are shown in table 23, below. The finding that generally journeys to risky facilities tend to be longer is consistent with the interpretation of figure 8, above.

When the same tests were used for journeys aggregated to addresses, there was a statistically significant difference in the distribution of (mean) journey distances to risky facilities compared with non-risky facilities, again when comparing for all risky facilities ($D(805)=1.783$, $p=0.003$) and for overall risky facilities ($D(805)=1.796$, $p=0.003$). There was not, however, a statistically significant difference in the mean rank of these aggregated journeys. As with the individual journeys, the mean distance travelled to risky facilities tended to be slightly longer than to non-risky facilities.

Table 23: Journeys-to-crime tests of difference outputs

Method of identifying risky facilities	MWW test			K-S test		
	Z	N	p	D	N	p
All risky facilities	7.646	7350	***<0.001***	3.769	7350	***<0.001***
Overall risky facilities	4.153	7350	***<0.001***	2.684	7350	***<0.001***

Thus, the answer to research question 5.2 is that crime journeys to facilities are relatively short, but that journeys to risky facilities, whilst still short, tend to be longer than those to non-risky facilities. However, each facility type needs to also be considered, as this overall pattern might disguise differences by category. This may be particularly the case if the *retail other* pattern is different, as it makes up quite a large proportion of facility crimes (and journeys).

Before moving on, it is important to recognise the critique that has already been presented regarding the possible effects of aggregation and nesting on the patterns obtained in typical journey-to-crime (distance) analysis. Notably, it has been proposed that 'prolific' offenders (those with many trips represented in the data) may skew the overall means, and that aggregating individual journeys, analysing them, and then making inferences back to individual behaviours (typical distances travelled to offend) is a form of ecological fallacy. As the purpose of the current research is to consider the general patterns relating to distance of facilities (both risky and not) from offender residences (rather than to draw conclusions about the journey-to-crime behaviour of individuals in this area), I would argue that these critiques are of somewhat less concern for this study. However, it is recognised and accepted that the patterns found may be affected by the relative contributions of individual offenders.

Though it was not possible, within the constraints of the study, to fully address this, some analysis was carried out to check if these biases were likely to be present. Firstly, for each individual the number of crime trips within the dataset and the mean distance travelled was calculated and compared. Of 4223 different offenders, 1266 (29.98%) had more than one recorded crime trip to a facility and these accounted for 59.79% (n=4396) of all journeys. The most 'prolific' of these offenders was associated with 59 crime trips (mean distance = 2983.20m), followed by two offenders that recorded 30 trips each (mean distances = 4812.12 and 4313.28m).

As a reminder, the mean distance obtained in this study was 3788m, with a range of 0.3m to 16941m.[18] Considering three groupings of offenders, those with 15 or more crime trips had a mean journey-to-crime distance of 3437.13m (range = 1725.66 to 5800.13m, n=16), those associated with five to 14 trips travelled a mean distance of 3766.75m (range = 226.76 to 13873.07m, n= 187) and those with less than five trips covered a mean distance of 3846.29m (range = 0.30 to 15921.39m, n=4018). Therefore, although the mean distance travelled was slightly higher for those with fewer crime trips, there was little difference between these groups. To further consider this, a Spearman's rank order correlation was performed on the number of trips recorded per individual and the mean distance travelled by each. As anticipated, the correlation coefficient was negative, but extremely small, and this was not a statistically significant relationship (r_s=-0.005, N=4221, p=0.764), thus suggesting the number of trips recorded for an offender was not related to the mean distance they travelled. Though not definitive, this suggests the aggregated findings are less likely to be skewed by the effects of prolific offenders.

As my analysis was focused on trips to facilities (rather than trips *by* individual offenders), I also looked at the number of 'repeat' offenders associated with each address, and whether there was a relationship between this and the mean distance travelled to that premises. There were only 106 premises associated with an offender who committed more than one offence at that address, with just 47 of these involving more than one 'repeat' offender. The maximum number of repeat offenders was 86, but only 11 addresses had ten or more repeat offenders present in the database. The numbers of journeys carried out by these repeat offenders were also calculated. The minimum number of journeys that could be recorded for a repeat was (obviously) two. With this in mind, 64 premises were associated with 3 or more journeys

[18] It must be noted that due to the methods of aggregation, and a small number of missing facilities (3), the totals and ranges obtained when considering individual offenders and their trips are slightly different to those used for the rest of the journey to crime analysis, that is aggregated by facility address.

committed by (one or more) repeat offenders. Only 18 premises recorded 11 or more journeys committed by repeat offenders, with the top numbers of repeat offender journeys being 126 (from 86 repeat offenders), followed by 51 (from 39 repeat offenders), 37 (from 27), and 35 (from 28). The number of repeat journeys generally increased with the number of repeat offenders as one would expect, but this was not always the case.

In order to try to assess whether the mean journey distance calculated for each address was affected by the extent of repeat journeys (thus possibly being skewed by a number of prolific offenders with different crime journey patterns), further Spearman's rank order correlations were carried out, considering mean distance to crime for each address with, firstly, the number of repeat *offenders* recorded at the address and, secondly, the number of *journeys* associated with repeat offenders at that address. In both cases, the correlation coefficient was positive and very small, and neither was statistically significant (r_s=0.027, N=805, p=0.452 and r_s=0.027, N=805, p=0.449, respectively).

Though only indicative, these findings suggest that the number of repeat offenders and repeat journeys recorded does not affect the mean distance travelled to offend at facilities. Though this does not preclude that particular offender journey patterns have skewed the findings on average distance travelled to crime at facilities, it does suggest that when considering aggregated patterns of individual journeys to facilities, in the current dataset, these are unlikely to be noticeably affected by the journeys of repeat (as opposed to single) offenders. That said, further research looking at journey-to-crime at facilities (as discussed below) is required and this should include applying methods to test for the effects of, and if necessary, subsequently control for, nested data. A first step could be to replicate the work of Townsley and Sidebottom (2010), using the current data.

9.3.1.1 Crime journeys to different facility categories

Following the approach taken for all facility crime, the journey-to-crime data for each facility category were explored, and the differences in distances travelled to risky and non-risky facilities were tested. The results of these analyses are shown in table 24.

When considered by facility category, it is apparent that there are indeed different patterns in distances travelled to commit crime either at risky or non-risky facilities. Although a number of these differences were not statistically significant, overall there were four categories of facility where journeys to risky facilities tended to be longer than those to non-risky facilities (*educational, supermarket* and *retail other* - all significantly different, and *religious* - not significant). For the other categories, journeys to risky facilities tended to be shorter than to non-risky facilities. Of these, *convenience, healthcare, pharmacy* and *restaurant* were all statistically significantly different (for *pharmacy* and *restaurant* this was only for the MWW test, suggesting a statistically significant shift difference, i.e. a difference in the median but not in the distribution). Consideration of these results suggests that sample size may have played a part in whether a statistically significant difference was observed. Some of the samples are very small and the results cannot be considered reliable (for example, *licensed premises* and *religious*). Those categories with few recorded crime journeys also tended to exhibit shorter mean journeys to risky facilities compared to non-risky facilities, although this was not always the case. The results for *convenience* and for *pharmacy* (and to a lesser extent *healthcare*) are considered to be sufficiently reliable to draw the conclusion that different categories of facility display different patterns with regards to whether crime journeys to risky facilities tend to be longer or shorter than those to non-risky facilities. Possible explanations for these differences are considered alongside the qualitative findings, below.

Table 24: Crime journeys tests of difference by facility category

							Tests of difference					
	Risky facilities			Non-risky facilities				Of mean ranks (MWW)		Of distribution (K-S)		
Facility category	N	Mean	Median	N	Mean	Median	N	Z	p	D	p	Difference
Convenience	505	2953.84	2470.24	382	3370.20	2858.50	887	-2.540	**0.011**	1.689	***0.007***	RF journeys are shorter
Educational	927	4522.45	3381.41	129	3653.14	2729.48	1056	2.965	***0.003***	1.687	***0.007***	RF journeys are longer
Financial	139	3541.15	3035.60	70	4058.15	3560.73	209	-1.747	0.081	1.046	0.224	RF journeys are shorter (ns)
Healthcare	39	2437.38	2028.86	18	3835.38	4220.89	57	-2.954	***0.003***	1.530	**0.019**	RF journeys are shorter
Hotel	75	4253.88	3743.78	77	4209.93	3920.04	152	-0.096	0.924	0.581	0.889	RF journeys are shorter (ns)
Leisure	9	4141.27	3202.11	30	3292.20	3493.07	39	0.700	0.501*	0.731	0.659	RF journeys are shorter (ns)
Licensed premises	26	3379.67	2287.12	70	3641.15	3799.58	96	-0.709	0.478	1.129	0.156	RF journeys are shorter (ns)
Petrol	96	3478.60	3311.61	106	4132.33	4608.15	202	-1.379	0.168	1.324	0.060	RF journeys are shorter (ns)
Pharmacy	191	2596.39	2067.25	99	3257.72	2670.48	290	-2.217	**0.027**	1.346	0.053	RF journeys are shorter
Religious	6	2223.75	2607.35	33	2828.33	1974.88	39	0.350	0.747*	0.751	0.625	RF journeys are longer (ns)*
Restaurant	79	3158.02	2497.27	116	4154.42	3532.45	195	-2.492	**0.013**	1.356	0.051	RF journeys are shorter
Supermarket	1053	3733.71	3483.31	149	3499.67	2883.56	1202	1.981	**0.048**	1.473	**0.026**	RF journeys are longer
Retail other	2358	4106.18	4113.98	445	3250.83	2861.86	2803	8.886	***<0.001***	4.510	***<0.001***	RF journeys are longer

Finally, with regards to these results, it is noted that for three categories of facility, the type of crime journey that tends to be longer (either to risky or non-risky) facilities switched depending on whether the mean distance or the median distance was used. Given the nature of the distance distributions, the median has been taken as the more appropriate measure of central tendency. Part of the reason for the different patterns could be the small sample sizes for two of the categories (*leisure* and *religious*) and it has already been noted that the results for these cannot be considered reliable. The third category, *hotel*, has an N of 152 overall (risky facility journeys n=75; non-risky facility journeys n=77), however the distances to risky and non-risky facilities (whether compared on the mean or the median) are very similar, and there is certainly no statistically significant difference, therefore overall it is safe to conclude from the data available that crime journeys to risky and non-risky hotels tend to be similar in length. Of course, this category of facility is also particularly likely to be affected by the exclusion of out-of-study-area offenders, as hotels by their nature are likely to be (temporarily) populated by many people - some of whom will be offenders - who are not local to that area. On the other hand, there may also be hotels - possibly more likely to be risky - that by their business and target clientele are more likely to be visited by those who live more locally. This suggests that even within an homogeneous set of establishments (hotels) there may well be quite different environments, place management and routine activities that either produce risky (or non-risky) facilities or maybe even produce different *types* of risky facility. Unfortunately exploring this further was outside the scope of this project.

The analysis of difference in distances travelled to commit crime at facilities was also carried out for journeys aggregated to address. Other than for all facilities together (as already noted above) and *retail other*, none of the facility categories demonstrated any statistically significant differences in either distribution or mean rank. It is likely that this was in part due to the often small sample sizes, however, not all of the categories were small. It is possible that the results were also affected by having already 'averaged' the crime journeys when aggregating to addresses, thus reducing the possible variation (and likelihood of finding statistically significant results). Despite this, the direction of the (non-significant) difference in the median distance travelled to risky facilities compared to non-risky facilities when aggregated to premises was the same for each category type as the direction identified in the crime journeys analysis (for example trips to risky convenience stores tended to be shorter than those to non-risky ones, whilst trips to risky supermarkets tended to be longer than trips to non-risky ones, for both measures). The only categories where this was not the case were *hotel* and *leisure*, for which the analysis on aggregated journeys identified longer trips to risky facilities, whilst the individual journeys analysis found these tended to be shorter. However, as already noted, the results for these categories are not considered reliable as this pattern 'flipped' depending on whether the mean or median was compared. Again the sample sizes here were also so small as to effectively exclude their consideration.

9.3.1.2 Differences across facility categories

In addition to considering the patterns for each category of facility, differences in journeys-to-crime across facility categories were also considered, as set out in the methods section, above. Differences in patterns for crime at all premises (regardless of whether risky or not) were tested, but of most interest were the differences in distances travelled to risky facilities (across each of the 12 main categories, so excluding *retail other*).

The longest median distance travelled to commit facility crime was to petrol stations, followed by hotels and supermarkets. The shortest median crime trip was to religious premises, followed by healthcare premises and pharmacies. The median (and mean) crime journey distance for each facility category are also shown in table 24, above.

When considering all journeys by facility category (regardless of whether to a risky or a non-risky facility), the Kruskal-Wallis H test (as described above) demonstrated there was a statistically significant difference in the distance travelled to offend (H(4424)=142.511, df=11, p<0.001). Using the Dunn-Bonferroni post hoc test as described in the methodology, the significant differences are shown in table 25.

Table 25: Dun-Bonferroni post hoc test: pairwise comparison of facility categories (significant results only)

Pairing (shorter median first)	standardised test statistic (Z)	Adjusted p-value
Convenience-educational	-8.564	***<0.001***
Pharmacy-supermarket	-6.208	***<0.001***
Convenience-hotel	-5.968	***<0.001***
Pharmacy-educational	7.461	***<0.001***
Pharmacy-financial	4.521	***<0.001***
Convenience-supermarket	-6.813	***<0.001***
Pharmacy-hotel	6.276	***<0.001***
Pharmacy-petrol	4.801	***<0.001***
Religious-hotel	4.27	***0.001***
Convenience-petrol	-4.302	***0.001***
Pharmacy-restaurant	-3.961	***0.005***
Convenience-financial	-3.975	***0.005***
Religious-educational	3.88	***0.007***
Healthcare-hotel	-3.664	**0.016**

Generally speaking, crime journeys to convenience stores were often statistically significantly shorter than other types of facility (*educational, hotel, supermarket, petrol* and *financial*). Crime journeys to pharmacies also tended to be shorter (statistically significant differences were seen versus *supermarket, educational, financial, hotel, petrol* and *restaurant*). Supermarkets, educational establishments, petrol stations and hotels stand out as having statistically significantly longer journeys (than *pharmacy* and *convenience* for all four, and also compared to *religious,* for educational establishments and hotels, and *healthcare*, for hotels only).

Considering only crime journeys to risky facilities, the longest median distance travelled was for hotels, followed by supermarkets, educational establishments and petrol stations. It is not unexpected that journeys to these types of facility might involve people who have travelled further, as will be discussed below. However, the actual locations of all available premises and of risky addresses must also be considered, as this may impact how far people might travel to use these (for either offending or legitimate purposes). The shortest crime journeys were made to healthcare establishments (though the numbers are small), pharmacies, licensed premises (which again has a small n), and convenience stores might also be considered here. Again, this pattern makes sense for facilities that serve fairly local communities.

The Kruskal-Wallis H test also demonstrated that there was a statistically significant difference in the distance travelled to different facility types when only journeys to *risky* facilities were considered ($H(3145)=152.623$, $df=11$, $p<0.001$). Again, the significant differences were identified through a post hoc test, and are shown in table 26.

Table 26: Dunn-Bonferroni post hoc test: pairwise comparison of facility categories (journeys to risky facilities only)

Pairing (shorter median first)	**standardised test statistic (Z)**	**Adjusted p-value**
Convenience-educational	-9.189	***<0.001***
Pharmacy-supermarket	-6.723	***<0.001***
Convenience-hotel	-4.828	***<0.001***
Pharmacy-educational	7.932	***<0.001***
Convenience-supermarket	-7.513	***<0.001***
Pharmacy-hotel	5.281	***<0.001***
Healthcare-educational	4.359	***0.001***
Healthcare-hotel	-4.061	***0.003***
Pharmacy-financial	3.84	***0.008***
Healthcare-supermarket	-3.747	**0.012**
Pharmacy-petrol	3.473	**0.034**

Many of the patterns identified through the post hoc test were the same as for all crime journeys to that type of facility (regardless of whether to a risky facility or not). This means it is not possible to determine if the patterns seen here are because of different crime journey lengths to different types of facility, or if they are differences unique to risky facilities (that have then influenced the patterns seen for all journeys). Two pairs of statistically significant crime journey distances were identified for trips to risky facilities, that were not seen for all journeys, with both *educational* and *supermarket* being significantly longer than *healthcare.* However, as the sample size for healthcare is quite small, this finding must be treated with caution.

Although crime journeys to non-risky facilities are not the focus of this study, by also considering these it was possible to ascertain whether the differences found for risky facilities simply reflected overall patterns in crime journeys across various facility categories, or whether there seemed to be something different about journeys to risky facilities. Overall, the Kruskal-Wallis test found a statistically significant difference in the distributions of crime journey distance by facility category ($H(1279)=36.132$, $df=11$, $p<0.001$). However, when the pairwise comparison was carried out (as above), there was only one pair with a statistically significant adjusted p-value, and this was *religious-hotel* ($Z=3.555$, $p=0.025$, *religious* Mdn=1975m, *hotel* Mdn=3920m). This adds some weight to the suggestion that crime trips to risky facilities differ when comparing across category types, beyond differences in facility distribution.

Taken together, the above findings suggest that, as asked in research question 5.3, crime trips do differ by facility type. Overall, the key differences appear to be that trips to risky convenience stores and pharmacies tend to be shorter, whilst trips to risky educational establishments, supermarkets, petrol stations and hotels tend to be longer. It was also the case that for these same types of establishment, there were differences in crime journeys to risky versus non-risky facilities, with convenience stores, pharmacies (and healthcare) tending to have shorter journeys to risky facilities than non-risky ones, whilst educational premises and supermarkets tended to have longer journeys to risky, over non-risky, addresses. On the other hand, for petrol stations and hotels, trips to risky facilities tended to be shorter, but the difference was not statistically significant.

9.3.1.3 Crime type

For comparing journeys by crime type, the following were tested: (1) whether there was any difference in the distance travelled to commit acquisitive crime at risky facilities versus non-risky facilities; (2) whether there was any difference in the distance travelled to commit violent crime at risky facilities versus non-risky facilities; and (3) whether there was any difference in the distance travelled to risky facilities to commit acquisitive crime versus violent crime.

For acquisitive crime the mean distance travelled to commit crime at risky facilities was 3827m (Mdn=3642m) and to offend at non-risky facilities the mean was 3447m (Mdn=2965m). The MWW test found there was a statistically significant difference in the mean rank distance travelled, with journeys to risky facilities tending to be longer (Z=7.257, N=5878, p<0.001), and that the distribution of these journeys was also different (D=3.792, p<0.001). This was also the case for violent crime, with a statistically significant difference in the mean rank distance travelled (Z=2.848, N=739, p=0.004) and in the distribution (D=1.912, p=0.001), again with crime journeys to risky facilities tending to be longer (M=4373m, Mdn=3295m) than those to non-risky facilities (M=3493m, Mdn=3094m).

The results of comparing journeys to commit acquisitive crime or violent crime, at risky facilities only, demonstrate that the mean distance travelled to commit acquisitive crime (N=4523) was 3827m (Mdn=3642m), which was less than the 4373m mean distance travelled to commit violent crime (N=556; Mdn=3295). The distribution of these crime journeys was statistically significantly different (D(5079)=3.431, p<0.001), but there was no statistically significant difference in the mean ranks (Z=0.397, p=0.691).

Therefore, it seems that journeys to risky facilities tend to be longer than those to non-risky facilities for both acquisitive and violent crime, and that trips to risky facilities tend to be differently distributed when committing violent crime than when committing acquisitive crime, with violent crime trips possibly longer. Although it is not possible to confirm this from the current analysis, it is possible that the longer trips to violence are found because of the use of recorded crime data. If the city has a (typical) concentration of licensed premises in one particular area, far from residential locations, then this could have explained longer journeys to violence. However, as it does not (and the extent of offending at the northern entertainment complex is unlikely to be enough to account for this pattern) it could instead be because violent crime committed more ‘locally’ is less likely to be reported, either because of the nature of the relationship between the police and those communities where violence is more likely to take place or because reporting may be lower for violence amongst people that know one another.

9.3.1.4 Time

Crime journeys were finally compared by time of offence, considering the same iterations as for crime type. That is to say: (1) whether there was any difference in the distance travelled to commit crime during daylight hours at risky facilities versus non-risky facilities; (2) whether there was any difference in the distance travelled to commit crime during darkness hours at risky facilities versus non-risky facilities; and (3) whether there was any difference in the distance travelled to risky facilities to commit offences during daylight versus darkness hours.

The distributions of crime journeys during the hours of darkness were statistically significantly different for trips to risky facilities and those to non-risky facilities ($D(1902)=1.927$, $p=0.001$), and the mean ranks were also statistically significantly different ($Z=2.814$, $p=0.005$), with trips to risky facilities tending to be longer (N=1223, M=3647m, Mdn=3415m) than those to non-risky facilities (N=679, M=3454m, Mdn=2862m). Crime journeys during hours of daylight displayed the same pattern, with trips to risky facilities being longer (N=4280, M=3963m, Mdn=3598m) than those to non-risky facilities (N=1168, M=3491m, Mdn=3071m). Again there was a statistically significant difference in mean ranks ($Z=6.711$, $N=5448$, $p<0.001$) and the distributions of journey distance were statistically significantly different ($D=2.983$, $p<0.001$).

When looking only at crime trips to risky facilities, the mean journey length during hours of darkness was 3647m (N=1223, Mdn=3415m), whilst for journeys during daylight it was 3963m (N=4280, Mdn=3598m). Although the medians are quite close, there was still a statistically significant difference in both the distributions of these crime trips ($D(5503)=1.772$, $p=0.004$) and the mean ranks ($Z=-3.593$, $p<0.001$).

When considering comparison across both crime type and time of offence, the over-riding pattern is that crime trips tend to be longer to risky facilities than to non-risky facilities, that there is some difference in the length of journeys to different types of crime, with the possibility that violent crime is committed further from the offender's home than acquisitive crime, and that offenders tend to travel further to risky facilities during daylight than darkness. The possible explanations for, and implications of, these findings will be discussed, below. However, it is important to note here that the same main pattern regarding trips to risky versus non-risky facilities was seen for the 'all facility crime' analysis already reported above. When the results for different types of facility were considered it became apparent that some categories followed this pattern, but that others tended towards shorter, not longer, journeys to risky facilities. Therefore, it is possible that the analyses presented here are masking variation at the facility category level and that considering journeys at this level of disaggregation might reveal different patterns. However, this has not been possible here because the sample sizes are, generally, too small.

9.3.2 Locations of risky facilities and police perceptions

In order to explore locational characteristics of facility crime further, the distribution of facility crime throughout the city is considered, with an obvious focus on risky premises (sometimes compared to non-risky ones). This section includes consideration of mapped, police recorded crime data, mapped offender data (based only on their facility crime trips) and the perceptions of the interviewed police officers, including consideration of their annotated maps and explanations for offending. In addition, where it helps reinforce or clarify a point, I refer also to my observations. The approach taken and methods of analysis of the different data sources have already been set out above. The themes identified from the analysis of the interview and annotated map data have been introduced in chapter 6. These were taken from the full thematic analysis, but it is not possible to cover all of them in detail in this study. The focus in this chapter, then, is on the themes of *Location Types*, and *Offender Sources and Routes*. However, this inevitably results in me considering some of the sub-themes identified under *Crime Types*, *Facility Issues*, *Offender Types*, and *Offending Explanations*. Also because of the relationships between these issues, they tend to be discussed together, so the chapter proceeds only loosely thematically.

Overall offending was mainly discussed in relation to gang-related activity, crime occurring as a result of drug deals or in locations where drug dealing took place, and offenders exploiting the ready opportunities to commit acquisitive crime in retail premises or at strip malls. In addition, there was some mention of more general violence, including domestic abuse (mainly in residential areas, therefore - although very important - not considered in this study), violence and disorder associated with people congregating or drinking on the street, similar types of behaviour caused by young people mainly around leisure and entertainment complexes and restaurants, and issues related to schooling (and school allocation policy in particular).

All the officers referred regularly to the ‘projects’, social housing and low income/low rent areas. It was clear that the majority of respondents believed these types of area were the source of their main offending populations. Most of the offender residence areas they identified on their maps, were locations with this type of housing, and they consistently referred back to these areas, sometimes naming ‘gangs’ (in both a loose and a more formal sense) that were associated with them and explaining how this fed into local offending (or offending along routes between different gang areas). This is an interesting dynamic that no doubt affects the patterns seen in the crime and journey data. If this is the case, it also reduces the generalisability of the findings to locations that do not have such entrenched gang and territory affiliations. Further consideration of the role of gangs, however, is beyond the scope of this study.

Another general finding that was apparent across all the interviews, and from my own observations, was the north-south split.

9.3.2.1 A city of two parts

As has already been raised in chapter 8, there is a clear distinction made by officers between the northern part of the city and the southern and central part. The PD is split into two districts (North and South) but a central area also tends to be referred to (usually alongside the south). Figure 9 shows the beat boundaries and the two districts (colour-coded) and figure 10 shows the more informal north-south-central split, with the south and central areas seeming to be considered by officers as more socio-economically and culturally similar to one another, being more deprived than the 'richer' northern area. It can also be seen on these maps that beat boundaries, thus districts, are contiguous with major roads and the east-west railway line (which is paralleled by the main east-west thoroughfare), thus these boundaries are also physical.

The distinction between the two parts of the city was stark during my observations, particularly the further north area, which was predominantly residential, houses were set on good-sized plots and streets were tree-lined. Here in the far north there were tennis courts and recreation areas with lakes and walking trails. The north/central area was dominated by commercial, mainly retail, premises situated in blocks of activity along the multi-lane road. There were few sidewalks and this part of town was designed around the car driver, with large parking areas serving each block of four, five, or six large retail premises (supplemented with coffee and fast food restaurants). To the western edges of the city, the area became more like the far north, but moving towards the south, there was a dramatic change in layout and property style. First the commercial area was passed through, with large showrooms and stores. As an annual market-town, however, most of the time this area was empty, with shutters down and cars only passing through. There was no downtown, no bustling financial and professional district. After departing this area, the 'true south' of the city was encountered. Although there were a number of shopping areas, some of which were large, generally stores tended to be smaller and the number of strip malls increased. Fast food restaurants became modest 'fried chicken' or 'pizza' shops and small chain and independent convenience stores appeared (which were not apparent in the northern areas). Big chain gas stations were still present, but there were also smaller and independent ones, that were difficult to distinguish from the local convenience stores, other than for a couple of gas pumps outside. Housing in this area was smaller, more densely situated and was clearly low-rent. Some, but by no means all, of the housing and the areas surrounding it would be best described as shabby and there was a general feeling that this was a less safe area (though this was obviously a subjective perception). Other than the main roads running through this area, streets were narrow, sidewalks and scrub were ever-present and some areas were frequented by sex workers. Though people still drove, it was far more common to see people on foot than it was further north (where it was extremely rare).

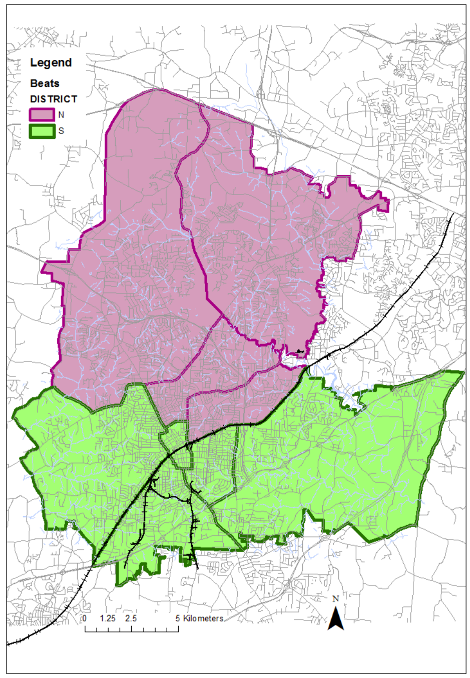


Figure 9: North/South districts (and beat boundaries)

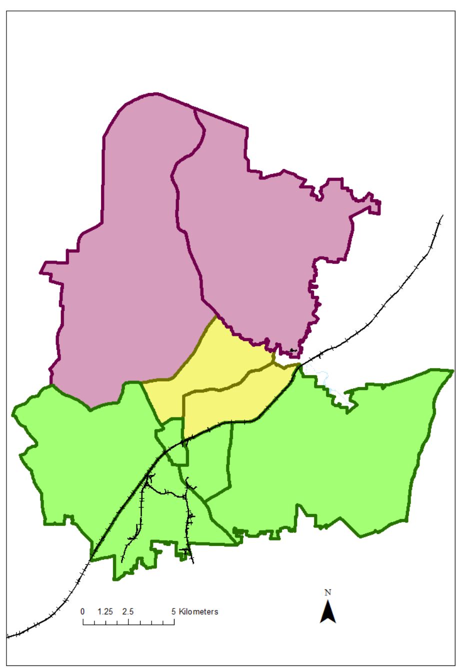


Figure 10: North/South/Central (and beat boundaries)

These same perceptions were echoed in the interviews (which were completed after most of the observations had taken place, therefore, this did not influence my impressions of the areas observed, though it may have shone a light on the relevance of certain elements). The northern part of the city was perceived as being ‘rich’ or generally as more well-off and populated by working families and professionals. One respondent commented that northern residents’ crime concerns were far more trivial, but seen by the residents themselves as serious problems, which they would call the police about:

> You know, they’ll call you about stupid stuff like this vandalism of a car and they want you to take a report on it just because…somebody threw something at the car like an egg or something…I mean that’s where you get all your petty calls like vandalism and dogs barking and it’s just ridiculous calls…[they think that] because somebody egged the vehicle, it is like equivalent to a robbery or something. (R6)

The southern part of the city was described by respondents as having a much more deprived population, many of whom did not, or could not, work, and who lived in initiative or project areas, or those characterised by social housing or very low rent accommodation. Convenience stores in this area were described by officers as ‘walk-ups’, that existed to serve local residents. Also in the south, I was informed there were crack houses and open air drug markets. The interview respondents also identified the south as home to a number of local gangs, and some groups associated with (or ‘chapters’ of) regional or pan-country gangs (such as the ‘Bloods’).

As might be expected, the nature of crime was perceived to differ across the two areas. However, the south was perceived to be the main ‘supplier’ of offenders (committing crimes in both areas), and some out-of-town offenders were seen to be attracted into the two different areas, but for different reasons. These patterns will be further discussed where relevant below.

9.3.2.2 The locations of (risky) facilities: an overview

Risky and non-risky facilities are spread throughout the city, though there are some noticeable patterns to highlight. As can be clearly seen from the map in figure 11, many facilities are situated along the major routes through the city, particularly the main north-south street. Visual inspection suggests that risky facilities also tend to be on main streets, suggesting that risky facility distribution is a reflection of all facility distribution. That said, it also appears (again purely from visual inspection) that non-risky facilities also occur more widely across the city, including off main routes, whereas risky premises seem to be mainly found (particularly for *retail other* and *petrol*) along major thoroughfares. This is consistent with opportunity-based explanations regarding accessibility. It may also be the case, as my observations seemed to suggest, that the larger retail and financial establishments tend to be based on main roads, with easy vehicular access, whilst smaller stores are located on side streets and in (semi-)residential locations. As it has been suggested that one explanation for the existence of risky facilities is that they are larger/higher turn-over premises, we would again expect to observe this pattern.

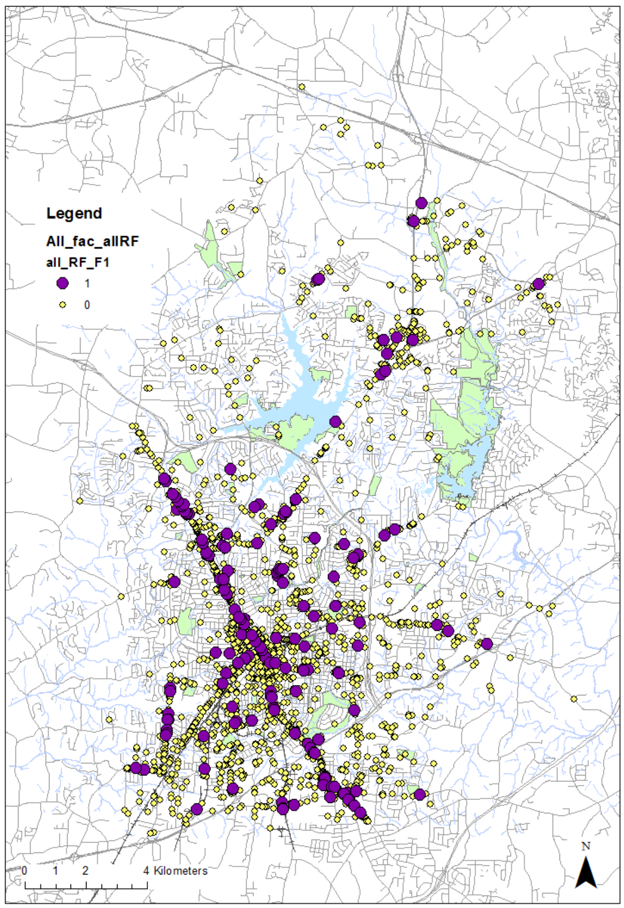


Figure 11: All facility crime (risky and non-risky facilities)

Having noted this general distribution, all facility crime was tested for the presence of hotspots of facility clusters (according to the method set out above). There were concentrations of high (none zero-crime) facility counts along the main north-south route as expected, with a particular hotspot around the railway line, stretching somewhat towards the north and a further hotspot to the south, that coincided with a small shopping area (see figure 12).[19] Risky facilities contributed to these facility hotspots, but they also appeared to occur relatively evenly spread between them as well (at least along the street network). When tested, there were no statistically significant risky facility hotspots (based on facility count), which indicates that risky facilities were not spatially clustered. Whilst this may be affected by sample size, taken all together this does suggest that risky facilities will occur more in the main facility locations (such as major thoroughfares and shopping plazas), but within this general distribution they do *not* seem more likely to occur in areas where there are particularly high proportions of facilities *per se*. In all, this basic analysis suggests that the locations of risky facilities in the city are more affected by the street network, than by the presence of (many) other facilities.

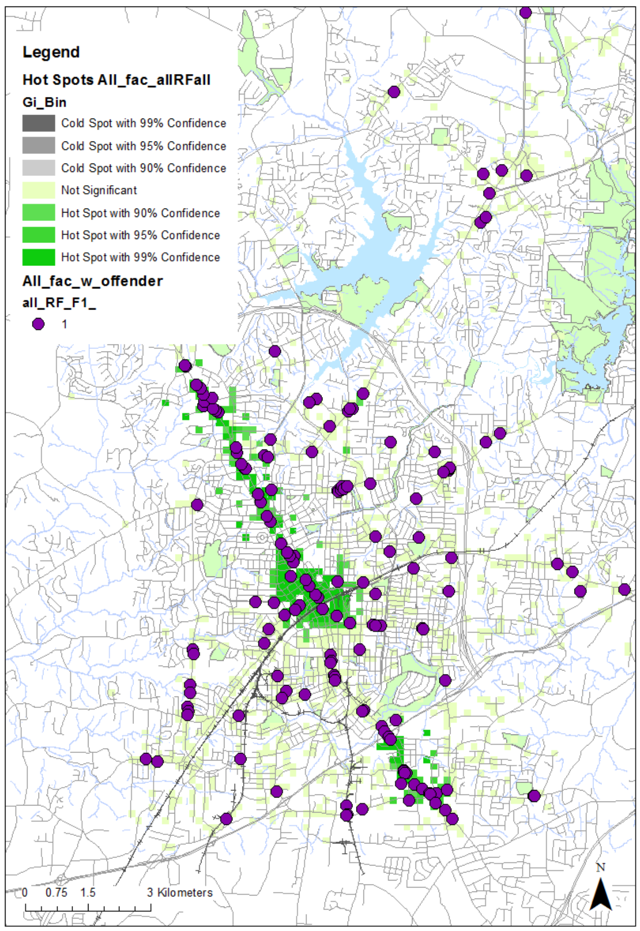


Figure 12: Hotspot of facilities, overlaid with risky facility point data

[19] It must be remembered that here I am using the term hopspot, but this is not based on crime count. Rather this is a hotspot of facilities, I.e. a location where there is a higher than expected count (of premises). The method does not account for the underlying topography, however, therefore the results are likely to be somewhat skewed by the locations *available* for premises to be situated, thus causing at least some degree of clustering.

This somewhat contrasts with the interview findings. There was some discussion that hinted at offenders targeting locations that were easy to access, but the road network itself was not mentioned. Respondents did highlight different patterns with regards the types of 'journey' that might be made to crime by local offenders (with particular reference to types of facility). In particular, they distinguished offences committed at large stores, particularly those in the shopping areas to the north of the city, as requiring transport. For example, when asked to clarify if those offending at places like North Walmart lived nearby, respondent 4 explained: "…we have one or two crap behind it, but the people who are having to get there go by bus or by car." The same respondent also commented that "[i]f they get cars they progress, or the bus.." This could mean that if offenders are travelling other than on foot, they will be more likely to target locations that are readily accessible from main thoroughfares, which is consistent with the general pattern of facility crime, and with the locations of risky facilities. However, it does not explain why risky facilities do not seem particularly more likely to occur where there are clusters of facilities than where there are not. Indeed, the respondents often talked about problematic locations as those places where there were multiple premises: "…this whole shopping area here…"(R1); "…the Palladium, the movie theatre, and all those stores and restaurants, same thing." (R3); "…it's just a big retail type area." (R4), implying that this would be the case.

9.3.2.3 Comparing risky and non-risky premises

Maps showing the locations of risky facilities and non-risky facilities for each category were considered and compared. There are some general patterns that can be observed across most of these maps. Firstly, risky facilities and non-risky facilities (in the same category) tend to occur near one another. In other words, both risky and non-risky premises are found in the same areas, suggesting that small-area differences are unlikely to be the explanation for their existence. This is consistent with Block & Block (1995) and Madensen & Eck (2008), though both of these studies looked at premises holding liquor licenses only. For some categories it can tend to look as if there are clusters of risky facilities, but even then there are non-risky facilities situated close by. Hotspot analysis was not carried out on the different facility categories as, in most cases, the number of risky facilities was too few (the required absolute minimum count to use the ArcGIS Online tool is 30).

As stated above, the addresses of known facility-crime offenders were also mapped, and tested for the presence of hotspots. There was one significant cluster, situated on the north-south district divide (falling either side of the railway line), just to the east of the main north-south thoroughfare. This is shown in figure 13. As already stated, it was not possible to map all offender address data, so this represents only a clustering of the addresses of those who were identified as committing facility crime (during the period studied). It is not possible to state if this is representative of the spread of all offender residences. However, this hotspot is consistent with the areas identified in interview as populated by local offenders, as well as overlapping with some of the areas of offender residence concentration that the respondents marked on the annotated maps.

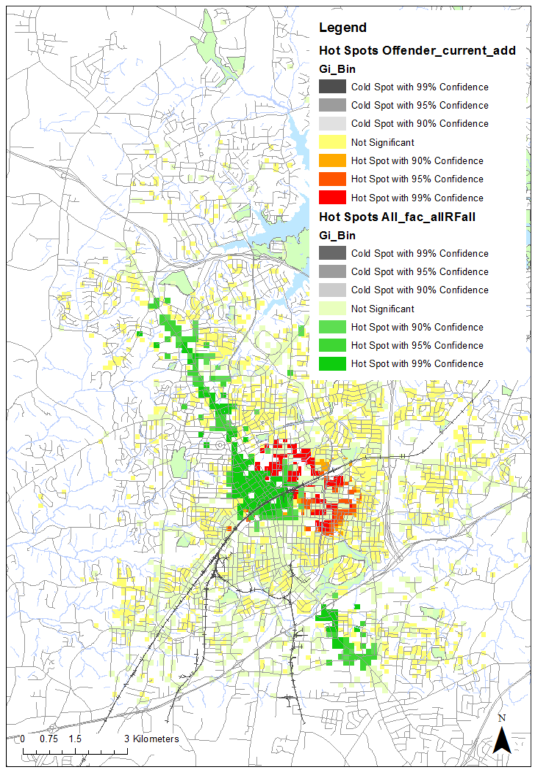


Figure 13: Offender residence hotspot and facility (count, not crime) hotspot (green)

When considering all risky facilities associated with an offender on the same map as the offender residence hotspot, this provides further information to help illustrate the patterns found through the journey-to-crime analysis, that offenders generally travel further to commit crime at risky facilities. either because the most suitable premises tend to be situated further away (by virtue of patterning of facility provision in relation to where offenders are more likely to live), or because they do not wish to offend at premises close by. This suggests two possible mechanisms for explaining why certain addresses are offended against (substantially) more than others. Firstly, a backcloth explanation, whereby the pattern results from the distribution of the 'best' targets (relative to areas of high offender concentration). Secondly, a risk-calculation explanation, whereby offenders deliberately travel further to avoid facilities nearer to where they live (where they perceive they are more likely to be recognised/caught). These are both consistent with opportunity theories. The fact that overall, crime facilities cluster next to the offender hotspot (also shown on figure 13, above) supports the second explanation, as in terms of quantity, there are targets available closer to home. However, observations and interview data highlight that the largest retail store targets (the two Walmarts and the main shopping mall) are not situated in this facility cluster, which supports the first possible explanation. The journeys taken to the two Walmarts are shown in figure 14, as illustration of this. It is also interesting that offenders tend to travel to offend at the two Walmarts from similar (possibly the same) locations, which means journeys north will be longer; but something clearly attracts offenders here. Further exploration is needed to test these hypothesis, but given the data, it appears that both are feasible and in fact may interact to produce the patterns seen.

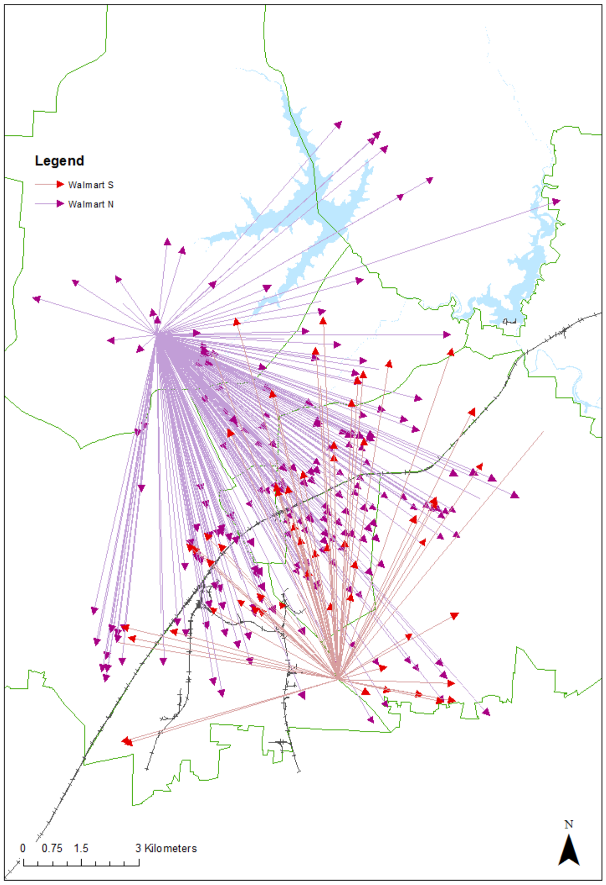


Figure 14: Journeys to offend at Walmart (North is purple, South is red) - arrows DO NOT show direction of travel, but rather highlight addresses

This pattern of travelling longer distances to risky facilities did not apply to all facility categories. The backcloth explanation certainly seems applicable to healthcare and pharmacies because, although pharmacies are located throughout the city, these premises (and certainly the largest) tend to be clustered in an area on the main thoroughfare, not far from the hospital (which is the major healthcare facility in the city), which is also relative close to the offender hotspot. The other facility type with shorter crime journeys, and shorter trips to risky facilities, was convenience stores (and possibly petrol stations, though we have seen that these two overlap. This is again consistent with the backcloth explanation, as neighbourhood convenience stores proliferate in the south and south/central areas. The interview data sheds some light on the routine activities that involve such premises. The stores themselves were described as 'walk ups', serving the needs of local residents, and sometimes being the only facility close-by. Respondents regularly identified this type of store as problematic.

> Another…local convenience store…They had a homicide there, the guy who used to own it, he was killed there. Numerous robberies, shopliftings and that one was a real hotspot for us because it was right beside that housing project… (R1)

It is also necessary to consider that even within one particular facility category, with one particular journey pattern, there can be several explanations for what might be taking place. This also highlights an advantage of mixed methods research, as such patterns would be much harder to detect (if possible) from the recorded crime data. When crime at local convenience stores was explained in the interviews, it was clear that different mechanisms were perceived to be at play. These can be summarised as (1) high volumes of local offenders, committing crime as they frequent local services (as part of their routine activities, a crime generator explanation); (2) local systemic crime (related to gang activity or drug-dealing) that took place in the vicinity of such premises, such as the parking areas; (3) complicit place managers who were in some way 'involved' in offending (from turning a blind-eye, to actively committing crime such as selling drugs); and (4) vulnerable place managers (which may co-occur with, or worsen, explanations (1) and (2)) who were themselves at risk of victimisation, or simply did not have the means or ability to be able to protect themselves or their premises. Although local convenience stores seemed to be relatively heavily victimised, there were still premises in this area that only experienced small amounts of crime. Further exploration of the mechanisms identified from this thematic analysis might help shed light on these findings.

9.3.2.4 Policy decisions

Another issue that was apparent in the interviews was the impact of other social policy decisions on crime patterns. In particular, school allocation policy. Respondents saw this as exacerbating other gang and territoriality issues in the area, by children being sent to schools not local to them, or by those in similar areas being sent to different schools. This was thought to result in tensions (and may also have affected patterns of offending outside schools as this policy also shaped the routine journeys taken to and from schools). Exploring this further is unfortunately beyond the scope of this study, but is illustrated nicely by a map showing the overlapping journeys-to-crime associated with the three main schools in the area, by offenders living in the vicinity of one another (figure 15). However, this does further demonstrate that it is not just the backcloth, but also the (non-crime) routine activities of individuals that shape patterns of risky facility formation and distribution. It also demonstrates how these patterns of activities can be produced as a result of far removed policy decisions.

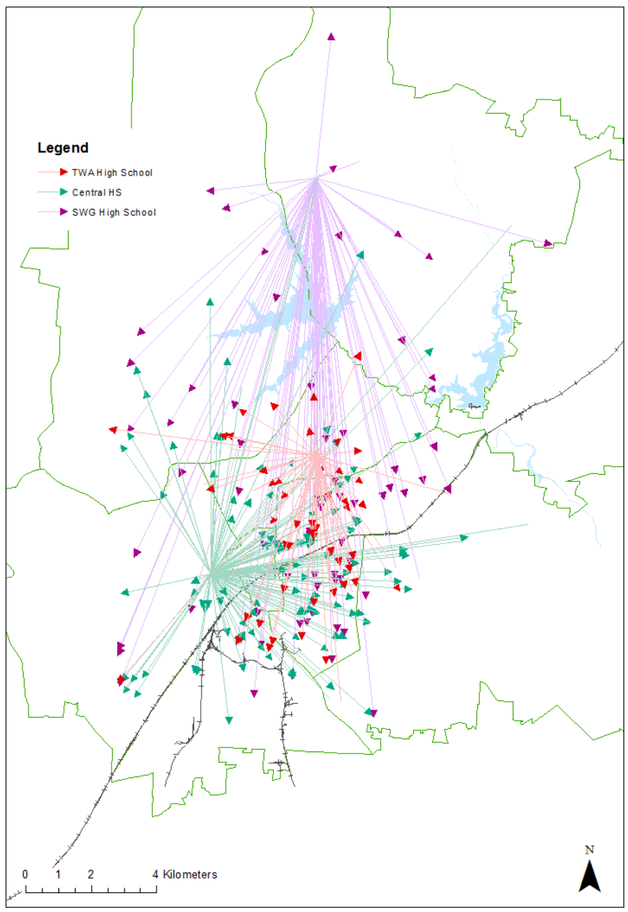


Figure 15: Journeys to three risky schools (the most risky is to the north)

Sources of offenders

Not apparent from the mapping, but an important theme in the interviews, were different sources of offenders. This is a limitation of the constraints of offender geocoding from the base data available that resulted in the exclusion of offenders who lived outside the city limits.

Respondents explained three major sources and patterns of retail offending. Local offenders would either steal for themselves (with some indication this might be closer to home, because fewer items were taken and it was easier to offend nearby, e.g. the southern Walmart) or they would steal to order for others, sometimes joining together to travel to suitable premises (e.g. the northern Walmart) to obtain the orders:

The other type of retail offender highlighted came from out of town. Known as 'boosters' (R6) these were individuals who would travel around the region, shoplifting all day: "..that's like a job to them and that's what they do. That's how they make money." (R6). Although local offenders might also be described in this way, generally it was people travelling into the area because they would not be recognised, offending until they became known to security personnel and then moving on to target other towns. This can help explain risky facilities, because the premises targeted tended to be the same chains, just in different places. This might be thought of as similar to a 'chainstore' boost repeat victimisation explanation (Bowers & Johnson, 2004; Tseloni & Pease, 2003).

9.4 Discussion

Quantitative analysis of the distance travelled allowed some quantification and comparison of the journeys made to commit crime at facilities, but did not tell us anything about what locations these distances were travelled from or to, or why. In order to address these issues, and attempt to answer the remaining research questions, content and thematic analyses of the qualitative data gathered through interviews, and the associated annotated maps (as well as informal conversations and observations) were also carried out. This allowed a consideration of features of offender crime journeys other than distance, and meant the findings were not strictly limited to data that relied on individuals being identified and charged with an offence. In addition, the recorded crime and offender data were mapped and displayed in a variety of ways using GIS software. The resulting maps were visually inspected for patterns, which were interpreted alongside the other findings. Taken in its entirety, this chapter is a further example of how the mixed methods approach that was employed allowed for triangulation, a more complete consideration of the issues and the opportunity to better explain findings than relying on quantitative data and analyses only.

Overall, the data suggest that patterns in facility crime, and risky facilities in particular, might be best explained by the constraints and influences of the environmental backcloth (Brantingham & Brantingham, 1993a), particularly clusters of facilities in shopping complexes and, crucially, major transport networks - notably the positioning of facilities along major roads. However, this is not the only explanation, as for some types of facility, proximity to local offenders seems to play a significant role. The findings on locations and journeys-to-crime are also consistent with routine activity and rational choice explanations of crime. This can be exemplifed by the patterns found when comparing daylight and darkness offending.

It was found that crime journeys during both darkness and daylight hours tend to be longer if they were to risky facilities than to non-risky facilities, and that when committing crime at risky facilities, offenders tend to travel longer distances during daylight hours than in the dark. A possible explanation for this is that offenders might feel they need to travel further not to be recognised or it could be because they are more likely to be offending while doing other things as well during the day (shopping, travelling to/from school, etc.). Alternatively, it might be because the places they offend at are more/less attractive at different times of the day. For example, big stores, even if open late, will be busy during the day so offenders may feel they are less likely to be noticed, whereas this would be less of a concern in small local stores as they are likely to be known anyway. This is in some ways consistent with the findings of Coupe & Blake (2006) in their study on the patterns of daylight versus darkness burglaries, as offenders travelled different distances, and to different locations, based on routine activity explanations. Burglars tended to travel to more 'up market' areas during daylight, when occupants were out at work (in the study area, this would equate to more offending in the north, which results in longer journeys given where offenders tend to reside). Targeting decisions in Coupe & Blake's (2006) study were also shown to be affected by 'cover', that is to say the chance of being seen, such that places where this was more likely tended to be targeted when the condition reduced these risks. In this case it was townhouses with less cover being targeted in the dark. In my study this could explain both targeting large stores and shopping complexes during the (busier) daytime periods and targeting local establishments when it was dark as the chances of being seen by people they knew (at premises or whilst travelling) would be reduced. On the other hand, offenders might travel shorter distances to offend at night just because this was more convenient, or because they stayed more locally as part of their normal (non-crime) evening routines, such as hanging out with friends in semi-public spaces (like the parking areas around convenience stores).

It has also been found through this analysis that non-crime related policies might inadvertently produce criminogenic situations, by influencing the routines of potential offenders - here school children (thus this would be consistent with Ratcliffe's (2006) temporal constraint theory), and by creating conflicts and tensions. These issues require further exploration, but again are evidence of the role in routine activities, and people's movements through the city, in explaining the existence of risky facilities.

One particular limitation of the journey-to-crime findings, is that they do not account for confounding variables that might better explain the distances travelled. For example, research has shown that distance travelled is associated with age, though this relationship may be quite complex, depending on available opportunities, access to transport, need to escape from guardians, routine activities, and so forth (Andresen et al., 2014; Baldwin & Bottoms, 1976; Clarke & Eck, 2003 (discussing Brumwell, unpublished); Repetto, 1974; Wiles & Costello, 2000), therefore it is possible that different types of facility might be predated by different age groups of offender, so that the difference in distance travelled is actually a result of different age profiles. When comparing risky and non-risky facilities, again the apparent difference in mean distance might actually be the result of age heterogeneity across premises. For example, younger offenders might be more likely to offend at a particular subset of retail premises, and their offending might be more prolific (or there may be more of them), resulting in a shorter mean distance recorded at the premises with the most offences, leading to the erroneous conclusion that offenders travel shorter distances to risky facilities (when in fact it is that risky facilities tend to host crimes committed by younger offenders – a very different, and extremely interesting finding). This suggests there is still much work to be done on exploring the demographic, offending history and possible other 'criminal career' features of the cohort of offenders responsible for 'making' a premises risky (compared with those addresses that are not).

This also raises the further concern that offender residence (in fact the best possible estimate of this based on current address) was used as the origin of the journeys analysed, when it is well established that what could be considered the 'active' journey to offend may have commenced at some other node. As has already been recognised, this is a significant problem for quantitative research of journey-to-crime, but the approach taken herein was not unusual, thus meaning the findings are consistent and comparable with much of the existing literature. However, as noted, research focused on origin and direction, both generally neglected is needed. This should include such research in the context of risky facility crime journeys as well.

A final limitation of the approach taken in the quantitative element of this research is the issue of aggregation and nesting. More sophisticated methods are needed to be able to discount bias in the outputs from offender characteristics, relative contributions and (hidden) heterogeneity, within this study, however the rudimentary checks carried out suggest that prolificity and variation in the number of crime trips contributed by individuals has not substantially affected the findings produced from aggregated data.

Beyond those issues discussed elsewhere in this chapter, there are no specific limitations relating to the qualitative and mixed elements of this research, and general limitations are rehearsed elsewhere. That said, a larger interview sample would have resulted in annotated maps of police-perceptions that could have been digitised and subjected to geographical analysis, alone and in comparison with the recorded crime data. In addition, there were a number of themes and issues raised in the interviews that could have been explored further here, as relevant to exploring the journeys taken to offend at facilities, the purposes for these and the wider (possible) relationship between (risky) facilities and offender residences and other anchor points. Unfortunately, this was beyond the constraints of the study, but it is hoped that those findings that have been presented will be the inspiration for further consideration of such issues (from a quantitative and a qualitative perspective).

9.5 Conclusion

The distances of journeys-to-crime have been considered and it was found that overall, journeys-to-crime were longer for risky facilities than for non-risky facilities, but that this was not the case for all facility categories, with convenience stores and pharmacies (plus, tentatively, healthcare facilities) having statistically significantly shorter journeys-to-crime at risky facilities than at non-risky facilities, whilst journeys-to-crime at educational establishments, supermarkets and other retail premises were the opposite (with statistically significantly longer journeys to commit crime at risky facilities compared to non-risky ones). Additionally, it was shown that journeys to commit crime at risky facilities varied in length by facility category, with convenience store and pharmacy journeys significantly shorter than many other facility types, and journeys to educational establishments, hotels, supermarkets and petrol stations longer. The latter finding was expected, although not inevitable, given the former. Overall it can be said that journeys to commit crime at convenience stores and pharmacies are relatively short and that this is even more the case when offending occurs at risky facilities; whilst journeys-to-crime at educational establishments and other types of retail premises (including pharmacies) are relatively long, and even more so when these are risky facilities. If this is generally representative of journeys made to commit crime at these types of facilities, it suggests that risky convenience stores and pharmacies are likely to be those that are situated near to a ready supply of offenders, whilst other retail premises and schools are more likely to be targeted by offenders journeying to them as part of their non-crime routine activities (these might be termed generator risky facilities) or purposely travelling to them with the deliberate aim of offending because they are seen as good (better?) targets than other, perhaps nearer, alternatives (these might be thought of as attractor risky facilities). That said, these patterns might not apply to all types of retail premises, but this cannot be ascertained from the data.

It was also found that crime trips to risky facilities tended to be longer (than to non-risky facilities) for both acquisitive crime and violent crime and that offences of violence *may* involve longer journeys to risky facilities than was the case for acquisitive crimes. In the same vein, risky facility crime trips are longer than those to non-risky facilities for daylight and darkness offending and crime journeys taken during daylight to risky facilities are longer than those taken during hours of darkness. As different patterns were found for different facility categories, then this might also be the case had these been analysed by crime type and time of offence, but the sample sizes were too small.

A number of possible explanations for these findings have been proposed, notably the constraints of the environmental backcloth and road networks, offending as part of non-crime routine activities and targeted offending at the facilities that provide the best opportunities and the least risks for *that time of crime* at *that time*.

This chapter, then, highlights that there remains considerable scope for exploring locational and environmental features in relation to risky facilities. Both the mapping of recorded crime data and the interviews highlighted the potential impact of street networks and facility clustering on the emergence of risky facilities and these need to be explored much more thoroughly than they have been to-date. Further, analysis of the role of co-location of facilities remains, effectively, in its infancy (Bowers, 2014) and this provides an opportunity to research and compare the relationship between, and impact of, a risky facility and other premises, both risky and non-risky alike.

Further, this study has demonstrated that the supply of offenders may be important in determining which premises become risky, but that the impact of this may be different for different types of facility. This introduces the possibility that there may be some facilities that attract offenders to them, some that become risky because they sit within offenders' non-crime routine activities and some that are risky because of sheer proximity to clusters of offender residences. These may be termed attractor risky facilities, generator risky facilities and proximate risky facilities and further exploration should form the basis of future research endeavours.

As already noted, I am unaware of any published research looking at journeys-to-crime associated only with facilities (where the crime trip to the facility is the focus of the research), beyond work by Mago et al. (2014) that looked at journeys-to-crime associated with shopping malls as crime attractors, but not looking at or comparing risky *facilities*. Given the interest in journey-to-crime more generally, this is an omission that needs to be rectified. Knowledge about the distances and directions travelled by offenders (and the ‘true’ origins of these trips – a more complex problem) to commit crime at facilities could help tease out the motivations and decisions of offenders, and how these and their routine activities and awareness spaces, as well as their ‘anchor’ points, could help explain the patterning of risky facilities within a given area. Based on the current research, it is also likely that different journey patterns will be found for different types of facilities. It is perfectly feasible that these are subject to similar explanations, but the patterns produced when they play out are different. For example, it may be a rational decision, based on least-effort, to offend at a convenience store close to home, and also to offend at a school on the far side of the city – because this is the school that the offender attends.

Future journey-to-crime research could also consider the impact on risky facilities of the distribution of offender residences within a city, and consider commuter/external offenders, which were excluded from the current research. A longer-term endeavour should also be to seek to separate and articulate the effects of offender journeys on risky facility emergence from the effects of facility presence (particularly those that are good crime opportunities) on the journeys offenders take (in other words to consider to what extent journeys and supply of offenders leads to risky facilities and to what extent the locations of facilities determine the length of journey offenders (have to) travel).

Chapter 10: Discussion of findings, original contribution and implications

10.1 Introduction

This study sought to investigate the phenomenon of risky facilities, a particular form of spatial (and temporal) crime concentration, where a small proportion of premises (within a given category of facility) contribute a substantially disproportionate proportion of crime (Clarke & Eck, 2003; 2007; Eck et al. 2007). Though the terminology has come into more regular usage, and the number of studies specifically employing this term is increasing, the concept itself remains under-explored. As a result, when considering nonresidential crime, particularly that associated with 'facilities', much of the literature involves focus on land use as part of the environmental backcloth, or looks at *types* of premises that are crime attractors or generators. In addition, much of the recent literature that refers to 'risky facilities', is either referencing Eck et al. (2007) as further evidence of spatial crime concentration or is a conflation of the attractor/generator and risky facility concepts. I contend that the risky facility is an important concept in its own right, that deserves both greater research attention, and to be disentangled from attractors and generators.

Alongside this, although there are numerous studies (as cited in the review of the literature and tabulated in Appendix 2) that find evidence of the risky facility phenomenon, in the main they focus on one, or a small number of, type(s) of facility (the most numerous being licensed premises, retail and schools), rather than considering the whole range of premises that might fall under such a definition. Further, many of the cited studies pre-date the emergence of the risky facility concept, therefore were not always specifically looking for this pattern. These studies also remain incomparable, as there is no empirical, agreed upon definition of what the risky facility pattern is, a relative measure is not always employed, and there is no established 'cut-off' for which premises should be classed as risky and which should not. This study sought to address all of these issues.

A further concern regarding the body of research building up around risky facilities, is that it has been too quick in endeavouring to explain this phenomenon, at the expense of firstly exploring how ubiquitous (one might say 'universal') it is, whether there is variability within the risky facility pattern across a range of features (such as type of premises, type of crime, the type of area in which they are situated, and the time of day) and (with limited exceptions) whether the pattern changes or persists over time. This study was developed as a result of these concerns.

A review of the literature also reveals that a decade after the first academic paper on risky facilities (Eck et al., 2007), some of the key concepts that have emerged out of the study of crime and place have not been explicitly considered in relation to patterning of risky facilities. In particular, I draw out the journey-to-crime as a potentially important element in determining the emergence and location of risky facilities, and note that there is no known research directly exploring the relationship between these two phenomena. Again, this study began to address this.

Finally, it is apparent that the vast majority of contemporary research on crime and place is quantitative in nature (though see some of Madensen's (2007) work with place managers; more broadly see Dymne, 2017; Greene, 2014; Kooi, 2015; St Jean, 2007; Telep, 2018). This is not unexpected given the desire to identify, test and compare spatial (and temporal) patterns and concentrations, and the relationships between these and other environmental or demographic characteristics (often across a city). However, the current study also sought to demonstrate that qualitative data could be useful in furthering our understanding, and more generally that mixed methods approaches could add substantial value and insight to the study of spatial crime patterns, particularly that of risky facilities.

In addition to all of this, the current study also aimed to test whether it is correct to refer to the risky facility distribution as following a 'power-law' (Eck et al., 2007), in order to consider whether this might have a part to play in identifying risky facilities (or to 'debunk' this idea). The quantitative analyses were also based on a dataset in which crime was attributed to particular premises without the problematic use of buffers, again arguably rare in this field of study.

The research was carried out, and presented, as three main studies. In each of the study chapters, specific data preparation and analysis were considered and the findings were presented and discussed. Therefore, this chapter aims to draw the different elements of the research together to discuss the main, and overarching findings, in light of the existing literature and the critique previously presented. As a result of doing this, I also consider the implications of these findings for those undertaking research related to risky facilities (and related concepts) and for practitioners (notably the police). I also highlight how the study advances our knowledge of risky facilities and the original contributions that have been made. Some limitations are mentioned, but as these are well-rehearsed in the previous chapters, they are excluded from specific discussion here, for brevity.

10.2 The research questions and a discussion of the findings

In carrying out this research study, I sought to address five research questions. In this section, I restate these, and summarise the relevant findings that answer them.

RQ1: How 'ubiquitous' is the concept of risky facilities?

RQ1.1: Is crime in facilities concentrated in some premises more than others? More specifically, do different classes of facilities in a given location show evidence of unequal distribution of crime, which may be considered consistent with the concept of 'risky facilities'?

RQ1.2: Is there variability in the concentration of crimes within facilities by: (a) facility type; (b) crime type; (c) time of offence?

The findings presented in chapter seven related to the ubiquity of the risky facility pattern. It was demonstrated that in the study area, crime was concentrated in some facility premises more than others, and that this distribution was consistent with the risky facilities pattern when considered against the (loose) 80-20 rule and when plotted as a 'J-curve' (Clarke & Eck, 2003; 2007; Eck et al., 2007). The findings of a large contribution of crime by a small proportion of premises was also consistent with Weisburd (2015) and subsequent studies seeking to test the proposed universal law of crime concentration at places. Specifically, the percentage of facilities contributing 25% and 50% of crime was in keeping with that found in the research to date (Gill et al., 2017; Habermann, et al., 2017; Lee et al., 2017; Levin et al., 2017; Wesiburd, 2015).

Having categorised facilities into one of 13 types (12 excluding retail other), the same general observation could be made regarding each of these categories. This part of the analysis was carried out on ten years of data, producing reasonable sample sizes even for the less numerous facilities (such as leisure facilities), and in all cases there was an unequal distribution of crime, again presenting as a J-curve and being not inconsistent with the 80-20 'rule'. However, it was apparent that some types of facility experienced a greater degree of inequality, with sharper, clearer J-curves, and a much smaller proportion of addresses accounting for 80% of crime (Clarke & Eck, 2003; 2007; Eck et al, 2007) and for 25% and 50% of crime (Weisburd, 2015). Indeed, for some types of facility (religious premises, restaurants and leisure establishments, and to a slightly lesser extent licensed premises and pharmacies), the proportion of premises contributing particularly 50% of crime, fell quite far outside of the expected bandwidth.

In response to my critique that research on risky facilities (and many other forms of spatial concentration) is currently incomparable, I also calculated the Gini-coefficient (based on the crime recorded for each address) for all facility crime, and for each category of facility. As a relative measure, this allowed me to directly compare the extent of inequality across types of facility, revealing that crime was most unequally distributed across educational establishments, supermarkets, petrol stations and convenience stores. If future studies follow the same approach, then it will be possible to ascertain the variability in crime inequality not only across facility types within a given area, but also across other areas (including other countries) and other types of area. For example, the majority of research on crime concentrations, and certainly that on risky facilities tends to focus on cities. However, the only minimum parameter for studying facility crime in this way is that there are sufficient numbers of each type of premises, and sufficient crime. What 'sufficient' might mean is open to testing, but crime counts, at least, can be inflated by using longer time periods (as long as the function of the premises themselves remains relatively stable, as discussed later).

Arguments for using the Gini coefficient, over other approaches, have already been rehearsed. In brief, it does not rely on an arbitrary proportion of crime (say 25%, but why not 30%), but its most important feature is the ability to compare across types of facility, time, and any other method of categorisation that is chosen (in this study, for example, across crime type and time of day). However, it does not indicate the nature of the distribution. Therefore, I recommend that future risky facility research utilises both the Gini coefficient to establish inequality in a way that is comparable across studies, and produces ranked crime counts (and cumulative percentages) and J-curve plots to more specifically consider the pattern that presents for that particular study.

In addition to comparing the extent of crime inequality across all facilities, and by facility type, this was also compared (for the same categories) by crime type and time of day. More specifically, acquisitive and violent crime were compared and daylight and darkness offending were compared. On the whole, it could be concluded that both types of crime were unequally distributed across premises (overall, and for each facility category) and both daylight and darkness offending were as well. Again, however, there was clear variability in the extent of this.

Acquisitive crime overall was more concentrated in certain premises than was violent crime, and this was also the case for each category of facility other than healthcare (where the two were very similar). This pattern was also consistent when the 25% (and other) contributions were considered. This raises an issue that requires further consideration (and testing): sample size. The number of premises experiencing violent crime, and the amount of such crime committed, was usually (often substantially) less than for acquisitive crime, but this was not always the case. Therefore, it is not possible to ascertain from the current analysis if the smaller Gini coefficients found for violent crime are an artefact of the smaller sample sizes, or if violent crime tends to be less concentrated within particular premises than does acquisitive.

However, this pattern did not hold on every occasion and, theoretically, it is conceivable that theft-related offences *would* be more highly concentrated as they may be more dependent on access to suitable targets (Clarke, 1999; Gill & Clarke, 2012) and a lack of capable guardianship through security and surveillance, (for example) (Cohen & Felson, 1979; Felson, 1995; 2008). In other words, acquisitive crime may be thought of as more driven by choice and 'rationality' than violence. A number of reasons for the risky facility pattern have been proposed (Eck et al., 2007; Madensen & Eck, 2008), but one of the favoured explanations is the role of place management (Eck, 1994; Eck et al., 2007; Madensen, 2007; Madensen & Eck, 2008). The way a facility is managed may, in part, determine how suitable it is for offending (whatever type of offending that might be). All other things being equal, we would expect good opportunities to be exploited more than poor ones. Opportunity and environment have been shown to play an important role in the presence (and location) of violence in licensed premises (as summarised by Graham & Homel, 2008), and this can be reduced through changes to management practices and layout, but it may be that opportunities and situations (or the lack thereof) that are associated with violent crime are more similar across premises (within the same category) than are good opportunities to commit acquisitive offences. If this is the case, then risky facilities would be synonymous with concentrations of opportunities.

It is suggested that further research is needed to test this pattern. The city in which this study was carried out had relatively low levels of facility-based violent crime, quite few licensed premises, and low rates of offending in licensed premises. Therefore, although violence was shown to be unequally distributed, this may not have been to the extent of other studies (such as Madensen, 2007). Analysis of other datasets, with higher proportions of recorded violence, and in cities with more typical night-time economies may add weight to, or challenge, the current findings. Such endeavours should then consider ways to capture and analyse the nature and spread of opportunities and crime-ripe situations for different types of crime across different categories of facility, and test whether these are consistent with the patterning of risky facilities in that area.

Similar patterns were also found when comparing daylight and darkness offences. In both cases crime was unequally distributed across premises (for all facilities, and for each class of facility), consistent with the risky facility J-curve. Generally, crime was less concentrated during darkness hours, which also tended to have lower crime counts but this was not always the case. Some facilities had more crime, and this was more concentrated, during darkness, such as licensed premises and restaurants, when usage is expected to be higher. However, for some facilities (convenience store, healthcare and hotel) the Gini coefficient was smaller for the period with the higher crime count. This again adds weight to the suggestion that although sample size might have an effect, the patterns found represent actual differences in crime concentrations and behaviour, for different facility types, different crimes and at different times of the day.

Overall, there is a general finding that crime tends to be more concentrated in categories with more premises and/or larger crime counts, and it makes sense that the more crime there is, the more concentrated it is possible to be, but this may not be the only explanation. It is recommended that further analysis be carried out to explore the mathematical relationship between the Gini coefficient, count, and sample size.

Taken together these findings are important as they highlight that even though the risky facility concept seems to be ubiquitous, there is variation in the extent of this concentration that needs to be taken into account both for identifying risky facilities (as discussed below) and for seeking to explain them. It is possible that more than one explanation might be needed for the emergence of risky facilities across different categories of premises type, and across other features (such as crime type and time of day, in the current study). The variation in concentration might also be related to different patterns of routine activities (again associated with particular types of facility, crimes or times) and (temporal) constraints (Ratcliffe, 2006). It is also plausible that what is perceived as a good opportunity (thus the (proportion of) premises that provide these) changes for different types of crime and at different times of the day. It is further conceivable this will vary for different types of offender (young, prolific, and so forth), which fell outside the scope of the current study, but should also be considered in future risky facility research.

All of this suggests that a more nuanced consideration of risky facilities is needed. Approaches such as that taken by the current study, should allow the development of a set of more 'realistic' explanatory hypotheses, that account for a variety of contexts and mechanisms (Pawson & Tilley, 1997; Sidebottom & Tilley, 2012), which should then be tested. Given the variation in concentration found herein, I also suggest that there may be different *types* of risky facility, as discussed further below.

RQ2: Are risky facilities stable?

RQ2.1: Is the distribution pattern persistent?

RQ2.2: Within each class of facilities, is the ordering of premises consistent?

As well as considering whether the risky facility pattern is contemporaneously ubiquitous (it is, but the extent of inequality varies), this study also sought to consider the *stability* of risky facilities, which could be seen as a further element of ubiquity. Two features were considered: persistency and consistency. The former referred to the persistency of the risky facility distribution over time, regardless of the specific premises. It was found that year-on-year all facilities, and categories of facility for which there were sufficient data, experienced an unequal distribution of crime, consistent with the risky facility pattern. For all facilities and many of the individual categories, the Gini coefficient showed there was very little difference in the extent of this inequality from year to year. The risky facility pattern is not, then, short-lived. This suggests that it is unlikely to be simply an artefact of the data (Eck et al., 2007). It also suggests that risky facilities are an important and abiding contributor to crime 'problems' in a given area and they therefore warrant further research attention and, crucially, need to be considered by practitioners as part of crime reductive efforts.

The other feature considered as part of research question two was the consistency of positioning of individual premises. Distinct from the persistency of the crime distribution *pattern*, these results showed that for all facilities, and for many of the categories analysed (excepting healthcare, leisure and religious premises) the rank position of crime contribution across addresses tended to be consistent. As this was concluded from the results of the Spearman's rank order correlation for all pairs of years, it suggests a tendency or general pattern, rather than 'proving' consistency for every address. Therefore, to explore this idea more specifically, the positions of the top-ranked educational establishments and supermarkets (two of the categories with the most unequally distributed crime counts) were inspected annually, across the ten years of data. This showed there *was* consistency of rank position for each address, particularly for the top-ranked premises, though as would be expected there was some movement. These findings, then, demonstrate that for this study area, the relative crime contribution was generally similar over a long period of time.

The importance of this is twofold. For practitioners, this suggests that if risky premises exist in the local area (as other findings indicate they will) they are unlikely to stop contributing such a large proportion of crime unless action is taken to intervene. In light of the other findings, this again reinforces the argument that identifying and responding to risky facilities is an important practical endeavour. It also suggests that such efforts are worthwhile, as these problematic premises will abide. This is also relevant for researchers as it suggests this pattern is not one that is short-lived, or inconsistent, again supporting calls for further research specifically focused on this phenomenon, and how it changes (or doesn't) over time.

The general pattern aside, when the top-ranked supermarkets were considered, it highlighted that some premises were *not* problematic for the whole time period – because they did not exist, either closing down during the period studied, or not opening until part-way through. As only the top ranked premises were inspected in this way, that they still appeared despite having less than ten years of data suggests that during their period of operation they might have been even more risky than was apparent. This raises further important issues for future research regarding the choice of time periods, but more crucially highlights that simply using police recorded crime data (however attributed to an address) may not be sufficient to fully explore the risky facility phenomenon, particularly when longer time periods, or longitudinal research, is being carried out.

Relatedly, it is also worth noting that the overall findings (across all research questions) relate to the full ten-year period covered by the data. Therefore, these are based on patterns that emerge from data aggregated over this length of time. It is necessary to recognise that different results may have been obtained if a different period or length of time had been used. In particular, a shorter time period, with lower crime counts, might have revealed less inequality of distribution and might also have been more affected by the exclusion of zero-crime facilities. Existing research on hotspots also suggests the possibility that this may have exhibited *greater but more short-lived* concentration (Gorr & Lee, 2018). Further comparison of patterns using different periods and aggregations of data should also be carried out to add to our understanding of how these decision impact on identification and selection of risky premises.

This longitudinal element of the research not only reveals interesting patterns about risky facilities, but is also an original contribution to existing knowledge. Criminal careers of 'places' (Sherman, 1995), and longitudinal trajectory analysis of street segments (Groff, Weisburd & Morris, 2009; Groff, Weisburd & Yang, 2010), have been considered by others, but neither of these concepts and approaches have been applied specifically to individual facilities. Given the findings of this study, I propose that there is scope to further explore risky facilities from these perspectives. This will help us to understand more about the nature of risky facilities, how they emerge (onset), why they persist (or do not) as long as they do (length), why they stop being risky (desistance), and how the severity and extent (frequency) of concentration changes during the lifecourse. Such research will also need to take into account changes at specific addresses, not only in premises type, but also in ownership/chain.

RQ3: How is the concentration of crime within facilities distributed?

RQ3.1: For each class of facility, does the concentration of crime best fit a proposed power-law distribution?

Research question three can be dealt with quite briefly, but this should not undermine the importance of its contribution. Somewhat casual mention of the risky facility pattern as following a power-law or being a form of Pareto distribution (Eck et al., 2007; Herrmann, 2015) or even the looser iteration of the 80-20 rule (Clarke & Eck, 2003, amongst others) may be a useful way of describing the risky facility form, but such terms should be used accurately, particularly as they may end up becoming lore. Endeavours to develop empirical methods of identification and definition might also be hampered, if it is assumed that crime is distributed across premises according to a power-law. Therefore, in this study I tested whether this was in fact the best fit.

Using methods developed by Clauset et al. (2009), I was able to reject the power-law as the best fit for crime distribution across all facilities (regardless of type) and a number of other types of premises (educational establishments convenience stores, financial institutions, licensed premises and petrol stations). This pattern varied somewhat when considered by crime type and by time of day, but in numerous cases the same could be said. The power-law distribution could not always be rejected, however. It is proposed that in some cases this was because the sample sizes became too small to achieve statistical significance, but it is possible that different types of distribution are apparent for different facility types, or for different offending patterns (type of crime and time). If the latter is the case, then the form of distribution may be a way of identifying the risky facility pattern, but given that the types of facility for which it could be rejected were also often (but not always) the ones with the greatest degree of inequality, the power-law function does not appear to be the most appropriate. I would suggest that pursuing further research regarding this distribution should be limited only to confirmatory analysis, perhaps using larger sample sizes (such as data from a large city). That said, it might also be interesting to consider if the heterogeneity of groupings also has an impact on the form of distribution (as the power-law function could not be rejected for the *retail other* category, which seems out of step with the other high crime/numerous premises categories).

RQ4: Can risky facilities be defined empirically?

RQ4.1 What quantitative methods could be used to empirically define risky facilities and what are the implications of using these?

RQ4.2: Which quantitative method is most appropriate?

RQ4.3: Where do operational police officers perceive risky facilities to occur and how does this compare to quantitative identification using recorded crime data?

Research question four was developed in response to my critique that identification and selection of risky facilities is at best arbitrary, and generally inconsistent. I proposed that risky facilities *ought* to be defined, thus identified, empirically. As far as I have been able to ascertain, no other studies using or testing the risky facility concept, have suggested employing a consistent, empirical definition or, therefore, what this might be. In response, I considered a number of ways risky facilities could be identified empirically, from both a quantitative and a more qualitative perspective. I argued that identifying risky facilities in a more uniform and consistent way is useful for practitioners, and very important for researchers, especially when seeking to compare across risky and non-risky facilities, and across different studies. The latter is crucial if we wish to establish and test rules and explanations regarding (micro-geographic) spatial and temporal crime concentrations. Replication and reinforcement are supported by methods that can be applied to multiple types of dataset, across different time periods and locations, resulting in a comparison of like with like. Put simply, if one researcher defines risky facilities as the top five premises by crime count in an area, whilst another defines them as the ranked addresses that cumulatively contribute 80% of crime, then my research suggests that very different quantities of risky facilities will be selected, with very different proportions of crime contribution. Carrying out further analysis seeking to establish explanatory variables, or features that predict the presence of such premises on these two different datasets is likely to produce different results. On the other hand, if the same method is employed across two different study areas, similar findings can be confirmatory of one another, whilst different results will raise further questions regarding the appropriateness of the hypotheses being tested.

In order to maximise the objectivity and replicability of any method used for selecting what premises should be considered, I firstly looked at quantitative methods that could be applied to recorded crime data. A small number of criteria were applied to develop the selection of methods to be assessed and compared, most notably that the approach be relatively simple. This excluded more sophisticated techniques, such as changepoint regression (Ratcliffe, 2012) that I have acknowledged may be more capable of ‘accurately’ identifying the point at which crime distribution begins to ‘level off’ from the heavy tailed ‘risky’ portion, but that I believe are too time/resource intensive and, frankly, too difficult to be recommended as the most appropriate method for identifying risky facilities. That said, I propose that research should be carried out, across a range of facility types and study locations using such methods, so that the results can be compared with the more prosaic methods considered in the current study, thus determining which, if any of these, produces results that are the most similar, most often.

The proposed methods were then considered conceptually, looking at such issues as whether they accounted for the underlying crime contribution, and if they reduced the likelihood of selecting risky facilities when the requisite pattern of unequal crime distribution across addresses was not present. After this, the methods were applied to the recorded crime dataset and the outputs were compared, to assess how appropriate each method was. Taking all of this into consideration I settled upon a preferred method (RF_F1) that selected risky facilities on the basis of those addresses with a recorded crime count that was greater than or equal to three times the mean. However, as there was little to choose between three times and four times the mean further testing and comparison on other datasets (and in other types of location) should be carried out to determine which is the most suitable.

As there is not yet a quantitative cut-off for defining premises as risky or not, it is not possible to say that any of the methods tested is ‘correct’: there is nothing to measure it against. An original contribution of this research study, therefore, is that not only have I proposed such a definition is needed, I have also assessed some of the possible approaches, discounted a number of these, and selected a method that I believe to be viable. Although more research is required to compare the outputs of using this approach across other datasets, and against more sophisticated, yet arguably less accessible, methods, my contribution has been a necessary first step towards establishing a universal empirical definition.

Qualitative analysis was also employed to consider both the appropriateness of the chosen quantitative method and whether the police recorded crime data revealed the ‘true’ pattern of risky facilities. Analysis of interviews with serving police officers and of their annotated maps provided some interesting points of comparison and clarification. In the main, it can be said that the recorded crime data revealed a greater range of *types* of facilities (respondents were generally focused on discussing convenience stores, supermarkets and other retail premises, and to a lesser extent schools, licensed premises and petrol stations) as well as a greater number, though given the interviews involved only six respondents and relied on recall, the latter point is expected. However, for the types of premises discussed in interviews, many of the top ranked addresses (or the commercial areas in which they were located) were identified. This acted as a form of confirmation that these were the ‘right’ risky facilities. For schools, however, this was less the case. They were highlighted by fewer respondents, and the top ranked address was not mentioned. Given the way the qualitative findings have been presented, extensive discussion regarding the discrepancies, patterns and implications of these results has already taken place (in chapter 8). Therefore, only the major key points are rehearsed here.

Most notable was the fact the city was perceived as being split in two halves (north and south), which were very different in terms of resident demographics, commercial provisions, and typical methods of moving around. Not only were these differences explicitly highlighted by the respondents (and noted during my own observations), but they were also discussed in relation to the patterning (and later journeys to) problematic premises (risky facilities), as discussed further below. The interviews and observations also revealed the different types of retail premises, convenience stores and petrol stations present in the data, and resulted in some limited, but necessary, re-categorisation for later analysis to reflect this. Both of these were crucial factors that would not have been apparent through just a quantitative analysis of the police recorded crime data.

There were also a small number of premises identified by the respondents as problematic, that did not appear as such in the recorded crime data. Various reasons for, and implications of, the 'discrepancies' (both ways) were presented in chapter 8. Overall, these differences highlight the importance, particularly for practitioners, of using multiple sources of data/information when determining which are the problematic premises in a city, and how best to respond to them. It also raised the issue of *perceptions* of risky facilities (and crime generators/attractors, i.e. types of facility) as well as of possibly associated offenders (and their motivations). This was flagged as potentially (negatively) affecting decisions on resourcing and prioritisation. A particular example of this was the possible role of familiarity and awareness spaces (Brantingham & Brantingham, 1981; 1993a; 1993b) in police perceptions of high crime locations/facilities (Ratcliffe & McCullagh, 2001), with some premises to the far north of the city missing from their considerations, despite appearing as risky in the recorded crime data. It was proposed that this might also explain the exclusion of the university from the interview list of risky facilities.

Taken together, the findings demonstrate that it is possible to identify risky facilities empirically, and that this can – and should – be done, in the main, through quantitative methods. I set out arguments as to why a universal definition was required and then, considering a range of different approaches, I proposed a working definition of premises with a (recorded) crime contribution of three times the mean. I also recognised that further comparative testing of this was required. Finally, I considered qualitative findings that both provided confirmation of the general appropriateness of the method chosen, whilst also highlighting some of the potential benefits, and problems, of not just relying on quantitative data when determining what facilities within a city are problematic.

RQ5: Where are risky facilities located and what types of crime journeys are made to them?

RQ5.1: How are risky facilities distributed across the study area?

RQ5.2: What distances of journeys-to-crime are associated with risky facilities (compared to non-risky facilities)?

RQ5.3: Do distances of crime journeys differ by (a) facility type; (b) crime type; (c) time of offence?

RQ5.4: What are operational police officers' perceptions of journeys-to-crime at risky facilities?

Research question five sought to address concerns over the lack of attention paid to locational characteristics of risky facilities across the city, the possible relationship to areas of offender supply, and the journeys (specifically distances) between them. As with the previous set of research questions, the nature of the findings and use of qualitative data have meant that much of the discussion has already been presented (in chapter 9), therefore here I draw out the key patterns only.

The consideration of journeys-to-crime at risky facilities has been noted as an original contribution of this research, with no known studies considering this previously. The distances of journeys in the current research are, however, generally consistent with the wider journey-to-crime literature, the mean distances falling within the range summarised by Vandeviver (2013). Overall, distances to facility-crime were relatively short. Importantly, journeys were compared across facility types and between risky and non-risky facilities (by facility type, crime type and time of day). It was found that the distance travelled to commit crime at different types of premises varied, and that there were statistically significant differences for some pairings of facility (for example journeys to convenience stores were shorter than those to educational establishments and journeys to pharmacies were shorter than those to supermarkets). There were a number of possible explanations for this that require substantial further research and testing, but one of the most likely is the distribution of such facilities throughout the city.

When considering risky and non-risky facilities, it was found that overall the distance travelled to risky facilities was greater than that travelled to non-risky facilities. However, when this was considered by facility category it was apparent that this masked different patterns for different types of premises. For some categories this was hampered by small samples, but it was possible to conclude (albeit limited by the representativeness of the data used) that journeys were longer to risky facilities (than non-risky facilities) for educational establishments, supermarkets and other types of retail premises, but that they were shorter to risky facilities (than non-risky facilities) for convenience stores, healthcare premises, pharmacies and restaurants. In chapter nine, possible explanations for these results were considered alongside the qualitative findings, but the environmental backcloth (Brantingham & Brantingham, 1981; 1993a; 1995; 2008) differing routine activities (Cohen & Felson, 1979; Felson, 2002; 2008), and risk-calculations (Cornish & Clarke, 1986; 2008) all seem feasible explanations

Similarly, differences in distances travelled to offend at risky versus non-risky facilities were found by crime type, with trips tending to be longer to risky facilities for both acquisitive and violent crime. It was also tentatively suggested, though could not be concluded from this research, that violent offences at risky facilities might be associated with longer journeys than acquisitive offences. Given the locations of licensed premises (for example) within the city, it was proposed that this difference might be the result of reporting differences, with local violence being less likely to come to the attention of the police. This requires further testing and consideration.

For time of day, again, both daylight and darkness offending seemed to involve longer trips to offend at risky facilities compared to non-risky facilities, and there was a small but statistically significant difference in the distances travelled to offend at risky facilities only, with daylight crimes involving slightly longer journeys. Again, it was proposed that this could be related to different routine activities and risk-calculations (Coupe & Blake, 2006).

The interviews and mapping (both from the qualitative and quantitative data), further highlighted some interesting patterns regarding the location of risky facilities throughout the study area. In particular, it was noted that there is a general north-south socio-demographic split in the city that also impacts on the type of commercial premises in the two areas. It was also found that there was a tendency for risky facilities to (generally, but not exclusively) be situated along main thoroughfares, consistent with opportunity theories regarding accessibility (Brantingham & Brantingham, 1993a; 1995) and (given the patterning of particularly retail property in this city) store size/footfall (Eck et al., 2007). The findings also suggested that risky facilities were not clustered together and did not exclusively occur in areas where (non-zero crime) facilities themselves were clustered. In other words, risky facilities were not *only* situated in concentrated commercial or retail areas.

When considering offender residences using both the quantitative and qualitative data, it was apparent that particular areas were more likely to contain concentrations of offender residences (one area was identified from the recorded (facility) offender data and others from the police-annotated maps). Interestingly the main area was situated next to the main facility cluster. However, there was no apparent relationship between this potential supply of offenders and the location of risky facilities. It was noted that the journey-to-crime findings suggested (overall) that offenders were likely to travel further to offend at risky-facilities, than at non-risky facilities, and this was supported by the mapping exercise. Some facility crime was committed near to offender residences in the main facility cluster, but many risky facilities involved travel outwards from this area (to the north or to the south). This was seen as potentially related to both ‘backcloth’ and opportunity supply explanations as offenders travel to where the best crime (and possibly non-crime) opportunities are, but also risk-calculation explanations with offenders avoiding those facilities closer to home, where they might be recognised (Rossmo, 2000; Rossmo & Rombouts, 2008).

However, this pattern did not hold for all types of facilities, particularly convenience stores. The journey-to-crime analysis showed that this type of risky facility was likely to be *closer* (than non-risky convenience stores) to offenders' home addresses, and this was again confirmed through the mapping exercise. Again it was noted that this could be explained as a result of opportunity theories: the locations of this type of store; the function(s) they served and the way they were used by local residents; and the way they were differently accessed (walk-ups). Importantly, however, this was further evidence that there may be different types of risky facilities and different mechanisms at play in explaining why they are formed (where they are). It further highlighted the usefulness of mixed methods and qualitative data (here interviews and observations of this type of location and store) in drawing out these patterns, nuances and mechanisms, to help make sense of quantitative findings.

Overall, the research on locations, offender residences and journeys (distance) to crime particularly highlighted the importance of the environmental and socio-demographic backcloths in understanding the patterning of risky facilities within a city, issues that have been mentioned in the literature, but not particularly well-explored. For example, Madensen (2007) and Madensen and Eck (2008) note that as risky and non-risky bars appear close to one another, this negates backcloth explanations. Though they were hampered in carrying out further analysis due to the spatial patterning of bars *generally* within their study area, this conclusion seems too simplistic. A similar pattern was found in my own research, with risky premises and non-risky ones often being located in close proximity, or even in the same strip, and with no evidence of geographic clustering of risky facilities. However, much more research is required, looking at a range of risky facility types and geographic distributions, before such conclusions can be drawn. Even when risky and non-risky premises (of the same or different types) are co-located, this itself could be the reason why one has emerged as problematic, or even acts as a protective feature for the other (Bowers, 2014). Further, it may be correct that the environment, topology, transportation network, socio-demographic characteristics, patterning of offender residences and even neighbourhood collective efficacy are not the *cause* of risky facilities, but it seems unlikely, given the findings of the current research, the emerging findings relating to street segments (Weisurd, Bushway et al., 2004; Weisburd, Groff et al., 2012; Weisburd, Shay et al., 2018) and some of the work on crime attractors and generators (as more loosely defined) (such as McCord, 2007) that these features do not play a role. Place management is clearly an important factor in risky facility emergence (Eck, et al., 2007; Madensen, 2007; Madensen & Eck, 2008), and is infinitely more malleable, but it is unlikely to create risky facilities *in isolation from the environmental and social backcloth*. As such, I recommend that the risky facility research agenda recognises the importance of a *range* of potential explanations, likely impacting upon one another and, as already noted, that there will also be variation in relative importance of these different explanations, and the mechanisms by which they lead to the emergence of risky facilities, across different premises types, crime types, and other features of the offence and location.

Another interesting, and unexpected, finding of this set of research questions was the impact of policy-decisions on the emergence and locations of (some) risky facilities, and likely the commission of offences in other types of location along travel routes. Very briefly, this related to school allocation policy in the city studied, the rivalries (across schools and neighbourhoods or residential areas, and the 'gangs' associated with them), and the particular distribution of nodes and paths (Brantingham & Brantingham, 1993b) this created. This suggests another important explanatory factor, as well as scope for further research.

As I have already noted, journey-to-crime is under-researched in relation to risky facilities, and it is hoped the analysis presented here will provide the impetus to develop a research agenda in this area. There are many issues that could be explored, which in doing so would help advance not only our understanding of how and why risky facilities emerge, but also how best to tackle them. In particular, more work needs to be carried out on distances travelled, including the consider of commuter/external offenders and whether these are drawn into an area by the presence of risky facilities, or they 'make' facilities risky by targeting specific areas where they are currently unknown, with well-known chains and premises that are most easily accessible from out-of-town being at greatest risk (as was indicated by some of my respondents).

In addition, any research on journey-to-crime should seek to account for the nesting and aggregation issues identified in the literature (Andresen et al., 2014; Townsley & Sidebottom, 2010; Smith et al., 2009; van Koppen & de Keijser, 1997), and to consider the other elements of the journey: direction and origin (Rengert, 2004), as well as other possible anchor points (Wiles & Costello, 2000), as it is anticipated from the current research that at least some of the crimes occurring at (or in facilities near to) schools are more likely to originate from the school itself, rather than the home address.

10.3 Methodological issues

It is also important to highlight (briefly, as these have already been well rehearsed) the implications and contributions associated with the methodological approach and decisions taken.

Firstly, this research was based on analysis of only crime recorded by police as occurring *at* a particular address. Whilst this will have resulted in a potentially large amount of 'missing' crime, it was in response to my critique of the use of buffers, especially for this kind of research looking at the concentrations of crime associated with specific addresses, many of which would be located close to other premises. Thus I argued it was important to be able to disentangle the crimes of individual premises, and that it was inappropriate to assume, without further research, that facilities would spill out crime into their environs (Bowers, 2014), or how far this would reach (Groff, 2011; Ratcliffe, 2012). Given that the findings of the research, particularly in relation to the existence of the risky facility distribution, are consistent with the wealth of existing knowledge on spatial crime concentrations, this approach has proven its feasibility. Whilst others have identified risky facilities in this way (e.g. Block & Block, 1995; Madensen, 2007), I am unaware of any published research that uses this method of attribution, across a whole city, for all types of facility, focusing on individual addresses. This is, therefore, a further original contribution and I recommend that this method is replicated to test the findings presented, in other cities, countries and types of location.

This research was based on crime counts, and it has been acknowledged that using a suitable denominator to consider rates is also worthy of further attention (Sidebottom & Bowers, 2010) (though the requirement of a high *rate* of crime is not a prerequisite of a risky facility). To prompt further consideration of this, I propose that certain categories of facility should be considered first, as the choice of suitable denominators is relatively uncontentious, and the data are likely quite readily accessible. Based on the categorisation used herein, I would suggest these include hotels (using number of rooms, or available beds, though this does not account for actual occupancy levels), educational establishments (using number of pupils on roll), and licensed premises (using licensed maximum capacity, though this does not account for actual customer numbers). Other ideas for denominators include number of covers for restaurants (though this would not apply to take-aways) and for healthcare the number of registered practitioners or registered patients (though the latter would not be applicable for hospitals, where a combination of beds and outpatient appointments might need to be used). Other more widely applicable, but much harder to obtain, denominators could include turnover or footfall.

Qualitatively gathered data were also analysed to further assess methods of identifying risky facilities as well as to explore the locations of such facilities and their relationship with the environment around them, including the supply of offenders. As noted earlier in this chapter, qualitative research is used within the approach we might refer to as place-based criminology, to which the current research belongs, and there is more recent evidence of an interest in utilising such approaches. However, these methods are still currently rare. The research discussed here utilised qualitative data collection and analysis to help further understanding of risky facilities, and to allow exploration of issues that might be difficult to operationalise or measure when relying on quantitative approaches. Qualitative data also facilitated a greater understanding of how those tasked with policing risky facilities perceived them, in terms of the types of premises they thought of when asked about what was problematic in their city and specific addresses that stood out as cause for concern (and how this compared to those selected from the recorded crime data). It also allowed a tentative insight into their thoughts on *why* particular places were risky, with a focus on locations, and types and methods of offending.

The use of qualitative approaches is highlighted, therefore, as a further particular contribution. I have incorporated and critiqued under-used (in this field) methods, and the findings these have produced. In particular, I have sought to do this through adopting a (partial) mixed-methods research strategy following what might be categorised as a convergent parallel design. The mixed-methods approach, has utilised quantitative and qualitative data in a number of ways, including to triangulate the data resulting in an *enhanced* interpretation of the findings (and a concomitant increase in their reliability and validity); as a way to increase the *completeness* of the research, by considering issues that would have been much more difficult to explore using mono-methods (in particular the more usual quantitative approaches); to *aid explanation* and *illustrate* some of the quantitative findings; and to improve the *utility* of the findings, such that more appropriate implications (and recommendations) for policy and practice have been identified.

10.4 Conceptual issues

Throughout this study I have argued that risky facilities require greater attention, and to be disentangled from other micro-level crime concentrations, particularly crime attractors and crime generators. Going back to the original definitions (Brantingham & Brantingham, 1995), the terms crime attractor and crime generator should be reserved for *types of premises* that tend to have higher levels of crime because either offenders are attracted by their reputation as good places to commit crime (attractors) or because large numbers of routine activities overlap at that node, bringing many suitable targets and motivated offenders together (perhaps in the absence of capable guardians) (Cohen & Felson, 1979). Thus the extant research suggests, for example, that licensed premises (e.g. Donnelly et al., 2006; Frisbie et al., 1977; Ronceck & Bell, 1981; Ronceck & Maier, 1991), schools (e.g. Roman, 2003; Roncek, 200; Roncek & Lobosco, 1983), shopping malls (e.g. LaGrange, 1999) and transport hubs (e.g. Kooi, 2007; Newton, 2004), amongst other places, tend to fit (one or both of) these definitions. Albeit supported by research, these are theoretically-driven concepts and relate to a whole class of facility. In the current study, for example, the facilities that were associated with the highest amounts of crime, thus could be deemed attractors/generators, were convenience stores, supermarkets, educational establishments, and petrol stations. If the number of such premises in the area (excluding zero-crime facilities) was taken into account, pharmacies would also be included here.

Risky facilities, on the other hand, again returning to the original definition, are those premises (specific addresses) within a class of facility, that contribute a disproportionate amount of crime (Clarke & Eck, 2003; Eck et al., 2007). Risky facilities, then, are individual premises, not types. They need to be empirically identified from the crime distribution within one type of facility and in a given area. In this study, for example, educational establishments demonstrated the risky facility distribution pattern and, specifically, using my proposed criteria for selection (three times the mean) ten premises were identified as risky facilities (the High School on Barrow, the High School on McGuinn, the High School on Ferndale, and so on). To drive the point home, schools are crime attractors/generators (regardless of location). The High School on Barrow is a risky facility in the local area.

Maintaining this distinction in the literature is important, because risky facilities, by their definition should be specific premises, are unique to a given area, and may arise for reasons not related to the attractor and generator concepts. For example, they may appear because of quirks of the data, because of the environment in which they are situated, because they are large or because of poor place management (Eck et al., 2007; Madensen & Eck, 2008). All of these are valid potential explanations and the existence of address that contribute a particularly large proportion of local crime is not negated because they fail to have a 'bad' reputation or they are small and not particularly busy. The other problem with using the attractor/generator terminology (and indeed the research literature) when discussing specific locations is that everyone 'knows' schools and licensed premises are crime attractors/generators, but of course as the risky facility research, including that presented in this study, shows, not all schools or licensed premises *do* experience high crime and certainly the majority of them will contribute a small proportion of the crime associated with that type of facility. Therefore, to focus resources and intervention activities on tackling premises 'just' because they fall under an attractor or generator facility category is likely to be wasted and ineffective, without first identifying which of those premises are the risky facilities for that area.

That said, considering my findings regarding variation in inequality of distribution by facility types, crime types and time of the day, those relating to journeys-to-crime, and the work of Bowers (2014), I propose that attractor and generator could be used as qualifiers, or adjectives, for the different types of risky facility that may arise as a result of different mechanisms and circumstances. Therefore I suggest the possibility of *attractor* risky facilities, premises to which people are drawn because of their reputation for providing good opportunities to offend; *generator* risky facilities, premises that are frequented by many people as part of their routine activities, some of whom will be potential offenders; and *proximate* risky facilities, premises that experience high proportions of crime simply (or mostly) because of being (appropriately) close to a supply of offenders (most likely clusters of residences, but possibly other anchor points). In reality, there are also likely to be *mixed* risky facilities, premises that fall into more than one of the other categories. There may also be a further category of *dependent* risky facilities, which experience disproportionate amounts of crime because of their co-location with other (risky) premises, of the same or different types, but this is not explicitly apparent from my research. Given the remaining possible explanations from Eck et al. (2007), there may also be *artefact* (or temporary) risky facilities, produced randomly within the data (but unlikely to persist) and *administrative* risky facilities (though these two could be combined), those premises that appear risky only because of crime reporting/recording patterns. These qualifiers need refining and testing, but using them in research would help to highlight the variation in types, patterns and, ultimately, explanations for risky facilities, and adds further direction to the research agenda.

10.5 Conclusion

The research reported in this study has sought to go back to the early work on risky facilities to explore some of the issues that I believe have been neglected by the rush to explain the phenomenon. Though it may be contended that this type of spatial concentration is now well established, the findings presented here suggest that there are a number of patterns that could yet be explored. These patterns can then help suggest some of the mechanisms that might be in play, thus highlight more possible explanations to test. In addition, I have proposed that in embarking on such a programme of research, a consistent method of identifying and selecting facilities as risky or not ought to be employed, and I have made suggestions as to what this should be.

References

Abbot, P. & Sapsford, R. (2006). Ethics, politics and research. In R. Sapsford & V. Jupp. Eds. *Data Collection and Analysis*. 2nd ed. London: Sage, pp291-312.

Alker, H.R. & Russett, B.M. (1964). On measuring inequality. *Systems Research and Behavioral Science*, 9/3: 201-18.

Altizio, A. & York, D. (2007). *Robbery of Convenience Stores*. COPS Problem-specific Guides Series, No. 49. Washington, DC: US Department of Justice.

Andresen, M.A., Frank, R. & Felson, M. (2014). Age and the distance to crime. *Criminology and Criminal Justice*, 14/3: 314-33.

Anselin, L., Griffiths, E. & Tita, G. (2008). Crime mapping and hotspot analysis. In R. Wortley & L. Mazerolle. Eds. *Environmental Criminology and Crime Analysis*. Cullompton: Willan Publishing, pp97-116.

Ashby, M. & Bowers, K.J. (2015). Concentrations of railway metal theft and the location of scrap-metal dealers. *Applied Geography*, 63: 283-91.

Ashby, M. P.J., Bowers, K.J., Borrion, H. & Fujiyama, T. (2014). The when and where of an emerging crime type: The example of metal theft from the railway network of Great Britain. Security Journal, 13 Oct 2014: 1-23; doi: 10.1057/sj.2014.43 retrieved (7 Jun 2016) Baldwin, J. & Bottoms, A.E. (1976). *The Urban Criminal.* London: Tavistock.

Armitage, R. & Pease, K. (2008). Predicting and Preventing the Theft of Electronic Products. *European Journal on Criminal Policy and Research*, 14(1): 11-37.

Arksey, H. & Knight, P. (1999). *Interviewing for Social Scientists*. London: Sage.

Baldwin, J. & Bottoms, A. (1976). *The Urban Criminal: A study in Sheffield.* London: Tavistock.

Bazeley, P. (1999). The *bricoleur* with a computer: Piecing together qualitative and quantitative data. *Qualitative Health Research*, 9/2: 279-87.

Becker, H. (1963). *Outsiders.* New York: Free Press.

Bellair, P. (27 July 2017). Social Disorganization Theory. *Oxford Research Encyclopedia of Criminology*. Retrieved 8 Jun 2018, from http://criminology.oxfordre.com/view/10.1093/acrefore/9780190264079.001.0001/acrefore-9780190264079-e-253.

Bentley, I. & Kerr, T. (2008). Lock It – Protect It. Presentation to the 2008 POP Conference, Birmingham. Retrieved from http://www.ukpopconference.co.uk/presentations/2008/POP%20Conference%20-%20Lock%20It%20Protect%20It%20-%20Workshop%20-%20ian%20bentley.ppt (3/3/09; link no longer available).

Bernasco, W. (2010). Modeling micro-level crime location choice: Application of the discrete choice framework to crime at places. *Journal of Quantitative Criminology*, 26/1: 113-38.

Bernasco, W. & Block, R. (2010). Robberies in Chicago: A block-level analysis of the influence of crime generators, crime attractors, and offender anchor points. *Journal of Research in Crime and Delinquency*, 48/1: 33-57.

Bichler, G., Christie-Merrall, J. & Sechrest, D. (2011). Examining juvenile delinquency within activity space: Building a context for offender travel patterns. *Journal of Research in Crime and Delinquency*, 48: 472-506.

Biderman, A.D. & Reiss Jr, A.J. (1967). On exploring the “dark figure” of crime. *The Annals of the American Academy of Political and Social Science*, 374/1: 1-15.

Block, C.R. (1995). STAC hot-spot areas: A statistical tool for law enforcement decisions. In C.R. Block, M. Dabdoub & S. Fregly. Eds. *Crime Analysis Through Computer Mapping*. Washington, DC: Police Executive Research Forum. Pp15-32.

Block, R.L. & Block, C.R. (1995). Space, place and crime: Hot spot areas and hot places of liquor-related crime. In J.E. Eck & D. Wesiburd. Eds. *Crime and Place*. Crime Prevention Studies. Vol 4. Monsey, NY: Criminal Justice Press.

Block, R., Galary, A. & Brice, D. (2007). The journey to crime: Victims and offenders converge in violent index offences in Chicago. *Security Journal*, 20/2: 123-27.

Bottoms, A.E. & Wiles, P. (1986). Housing tenure and residential community crime careers. In A.J. Reiss & M. Tonry. Eds. *Communities and Crime.* Crime and Justice a Review of the Research Vol. 8. Chicago: University of Chicago Press.

Bottoms, A.E. & Wiles, P. (2002). Environmental criminology. In M. Maguire, R. Morgan & R. Reiner. Eds. *The Oxford Handbook of Criminology*. 3rd ed. Oxford: Oxford University Pres, pp620-56.

Bowers, K.J. & Johnson, S.D. (2004). Who commits near repeats? A test of the boost explanation. *Western Criminology Review*, 5/3: 12-24.

Bowers, K.J. (2014). Risky facilities: Crime radiators or crime absorbers? A comparison of internal and external levels of theft. *Journal of Quantitative Criminology*, 30/3: 389-414.

Boyatzis, R.E. (1998). *Transforming Qualitative Information: Thematic analysis and coding development*. London: Sage.

Braga, A.A. (2008). *Police Enforcement Strategies to Prevent Crime in Hot Spot Areas*. COPS Crime Prevention Research Review, No. 2. Washington, DC: US Department of Justice.

Braga, A.A., Hureau, D.M. & Papachristos, A.V. (2010). The relevance of micro places to citywide robbery trends: A longitudinal analysis of robbery incidents at street corners and block faces in Boston. *Journal of Research in Crime and Delinquency*, 48/1: 7-32.

Brandon Tuma, N. (2009). Modeling change. In M. Hardy & A. Bryman. Eds. *The Handbook of Data Analysis*. Paperback ed. London: Sage, pp309-330.

Brantingham, P. & Brantingham, P. (1995). Criminality of place: Crime generators and attractors. *European Journal on Criminal Policy and Research*, 3/3: 5-26.

Brantingham, P.J. & Brantingham, P.L. Eds. (1981). *Environmental Criminology*, New York: Sage.

Brantingham, P.J. & Brantingham, P.L. (1984). *Patterns in Crime*. New York, NY: Macmillan.

Brantingham, P.J. & Brantingham, P.L. (2008). Crime Pattern Theory. In R. Wortley, L. Mazerolle. Eds. *Environmental Criminology and Crime Analysis*. Cullompton: Willan, pp78-93.

Brantingham, P.L. & Brantingham, P.J. (1993a). Environment, Routine and Situation: Toward a pattern theory of crime. In R.V. Clarke & M. Felson (Eds.), *Routine Activity and Rational Choice* (pp259-94). New Brunswick, NJ: Transaction Publishers.

Brantingham, P.L. & Brantingham, P.J. (1993b). Nodes, paths and edges: Considerations on the complexity of crime and the physical environment. *Journal of Environmental Psychology*, 13/1: 3-28.

Brantingham, P.L. & Brantingham, P.J. (1999). A Theoretical Model of Crime Hotspot Generation. *Studies on Crime and Crime Prevention*, 8(1): 7-26.

Brantingham, P.L., Brantingham, P.J., Vajihollahi, M. & Wuschke, K. (2009). Crime analysis at multiple scales of aggregation: A topological approach. In D. Wesiburd, W. Bernasco & G.J. Bruinsma. Eds. *Putting Crime in its Place*. New York, NY: Springer, pp87-107.

Brantingham, P.J., Dyreson, D.A. & Brantingham, P.L. (1976). Crime seen through a cone of resolution. *American Behavioral Scientist*, 20/2: 261-73.

British Society of Criminology (2015). *British Society of Criminology Statement of Ethics 2015*. Retrieved from http://www.britsoccrim.org/documents/BSCEthics2015.pdf (28/1/18).

Bruce, L. (2004). Why most bank robberies are on Fridays. North Palm Beach, FL: Bankrate.

Bruinsma, G.J.N., Pauwels, L.J.R., Weerman, F.M. & Bernasco, W. (2013). Social disorganization, social capital, collective efficacy and the spatial distribution of crime and offenders: An empirical test of six neighbourhood models for a Dutch city. *British Journal of Criminology*, 53/5: 942-63.

Burgess, E.W. (1925). The growth of the city: an introduction to a research project. In R.E. Park, E.W. Burgess & R.D. McKenzie. Eds. *The City*. Chicago, Il: University of Chicago Press, pp47-62.

Burrell, A. & Erol, R. (2006). *Violence in the night-time economy: Tracking hotspots over time.* Research report. London: UCL Jill Dando Institute of Crime Science.

Bursik, R.J., Jr. (1988). Social disorganization and theories of crime and delinquency: Problems and prospects. *Criminology*, 26: 519-51.

Bryman, A. (2015). *Social Research Methods*. 5th ed. Oxford: Oxford University Press.

Calder, J. & Sapsford, R. (2006). Statistical techniques. In R. Sapsford & V. Jupp. Eds. *Data Collection and Analysis*. 2nd ed. London: Sage, pp208-42.

Canter, D. & Gregory, A. (1994). Identifying the residential location of rapists. *Journal of Forensic Science*, 34: 169-75.

Canter, P.R. (1997). Geographic Information Systems and Crime Analysis in Baltimore County, Maryland. In D. Weisburd & T. McEwan. Eds. *Crime Mapping and Crime Prevention.* Monsey, NY: Criminal Justice Press, pp 159-90.

Capone, D. & Nichols, W. (1976). Urban structure and criminal mobility. *American Behavioral Scientist*, 20/2: 199-213.

Capone, D.L. & Nichols, W.W. (1975). Crime and distance: An analysis of offender behaviour in space. *Proceedings of the Association of American Geographers*, 7/1: 45-9.

Cayo, M. R. & Talbot, T.O. (2003). Positional error in automated geocoding of residential addresses. *International Journal of Health Geographics*, 2(10): 1-12.

Chainey, S. & Ratcliffe, J. (2005). *GIS and Crime Mapping*. Chichester: John Wiley & Sons.

Chainey, S., Reid, S. & Stuart, N. (2002). When is a hotspot a hotspot? A procedure for creating statistically robust hotspot maps of crime. In D. Kidner, G. Higgs & S. White. Eds. *Socio-economic Applications of Geographic Information Science*. London: Taylor & Francis, pp21-36.

Chapin, F.S. Jr (1974). *Human Activity Patterns in the City: Things people do in time and space.* New York, NY: John Wiley and Sons.

Cheatwood, D. (1995). The effects of weather on homicide. *Journal of Quantitative Criminology*, 11/1: 51-70.

City of Chicago (2018). Chicago History. Retrieved from https://www.cityofchicago.org/city/en/about/history.html (13/06/18).

Clarke, R.V. (1992). *Situational Crime Prevention: Successful Case Studies*. Albany, NY: Harrow & Heston.

Clarke, R.V. (1997). *Situational Crime Prevention: Successful Case Studies.* 2nd ed. Albany, NY: Harrow & Heston.

Clarke, R.V. (1999). *Hot Products: Understanding, anticipating and reducing demand for stolen goods*. Police Research Series paper 112. London: Home Office.

Clarke, R.V. & Eck, J.E. (2003). *Become a problem-solving crime analyst in 55 small steps*. London: UCL Jill Dando Institute of Crime Science.

Clarke, R.V. & Eck, J.E. (2007). *Understanding Risky Facilities*. COPS Problem_Oriented Guides for Police Problem-Solving Tools Series No. 6. Washington, DC: US Department of Justice.

Clarke, R.V. & Newman, G. Eds. (2005). *Designing out Crime from Products and Systems.* Crime Prevention Studies, Vol. 18. Monsey, NY: Criminal Justice Press.

Clarke, R.V., Perkins, E. & Smith Jr., D.J. (2001). Explaining Repeat Residential Burglaries: An analysis of property stolen. In G. Farrell & K. Pease. Eds. *Repeat Victimization*. Crime

Prevention Studies Vol. 12. Monsey, NY: Criminal Justice Press, pp119-132.
Clauset, A., Shalizi, C.R. & Newman, M.E.J. (2009). Power-Law Distributions in Empirical Data. SIAM Review, 51/4: 661-703. DOI: 10.1137/070710111
Clery Center (2018). The Clery Act. Retrieved from https://clerycenter.org/policy-resources/the-clery-act/ (14/01/18).
Cohen, L.E. & Felson, M. (1979). Social Change and Crime Rate Trends. *American Sociological Review*, 44(4): 588-608.
Cohn, E.G. (1990). Weather and crime. *British Journal of Criminology*, 30/1: 51-64.
Coleman, A. (1986). *Utopia on Trial.* London: Hilary Shipman.
Coleman, A. (1989). Disposition and Crime. In D.J. Evans & D.T. Herbert. Eds. *The Geography of Crime.* London: Routledge.
Corder, G.W. & Foreman, D.I. (2014). *Nonparametric Statistics: A step-by-step approach.* 2nd ed. Hoboken, NJ: Wiley.
Cornish, D. B. (1994). Crimes as scripts. In Proceedings of the international seminar on environmental criminology and crime analysis (Vol. 1, pp. 30-45). Tallahassee, FL: Florida Criminal Justice Executive Institute.
Cornish, D.B. & Clarke, R.V.G. (1986). *The Reasoning Criminal*. New York, NY: Springer-Verlag.
Cornish, D.B. & Clarke, R.V. (2003). Opportunities, Precipitators and Criminal Decisions: A reply to Wortley's critique of situational crime prevention. In M.J. Smith & D.B. Cornish. Eds. *Theory for Practice in Situational Crime Prevention*. Crime Prevention Studies, Vol. 16. Monsey, NY: Criminal Justice Press.
Cornish, D.B. & Clarke, R.V. (2008). The rational choice perspective. In R. Wortley, L. Mazerolle. Eds. *Environmental Criminology and Crime Analysis*. Cullompton: Willan, pp21-47.
Cotton, J.L. (1986). Ambient temperature and violent crime. *Journal of Applied Social Psychology*, 16/9: 786-801.
Coupe, T. & Blake, L. (2006). Daylight and darkness targeting strategies and the risks of being seen at residential burglaries. *Criminology*, 44/2: 431-64.
Crawford, A. (1998). *Crime Prevention and Community Safety.* Harlow: Longman.
Creswell, J.W. & Creswell, J.D. (2018). *Research Design.* 5th ed. London: Sage.
Crowe, T. (2000). *Crime Prevention Through Environmental Design*. 2nd ed. Boston: Butterworth-Heinman.
Cullen, I. & Godson, V. (1975). Urban networks: The structure of activity patterns. *Progress in Planning*, 4: 1-96.
Dalton, H. (1920). Measurement of the Inequality of Income. The Economic Journal, 30: 348-61.
Dawson, P. & Cuppleditch, L. (2007). An impact assessment of the Prolific and other Priority

Offender programme. Home Office Online Report 08/07. Retrieved from library.college.police.uk/docs/hordsolr/rdsolr0807.pdf (30/5/17).

Dedel, K. (2005). *School Vandalism and Break-ins*. COPS Problem-specific Guides Series, No. 35. Washington, DC: US Department of Justice.

Deltas, G. (2003). The Small-Sample Bias of the Gini Coefficient: Results and Implications for Empirical Research. The Review of Economics and Statistics, 85/1: 226-34.

Donkin, S. & Wellsmith, M. (2006). Cars Stolen in Burglaries: The Sandwell Experience. *Security Journal*, 19(1): 22-32.

Donnelly, N., Poynton, S., Weatherburn, D., Bamford, E. & Nottage, J. (2006). *Liquor outlet concentrations and alcohol-related neighbourhood problems.* Alcohol Studies Bulletins, 8. Sydney: BOCSAR NSW.

Durkheim, E. (1892) [1984]. *The Division of Labor in Society* (W.D. Halls, Trans.). New York, NY: The Free Press.

Durkheim, E. (1897) [2006]. *On Suicide.* (R. Buss, Trans.). London: Penguin Books.

Dymne, C. (2017). Hot spots of robberies in the city of Malmo: A qualitative study of five hot spots, using the routine activity theory, and crime pattern theory. Masters thesis submitted to the University of Malmo, Sweden. Retrieved from http://muep.mau.se/bitstream/handle/2043/23167/Hot%20spots%20of%20robberies%20final%20product.pdf?sequence=2&isAllowed=y (7/1/18).

Eck, J.E. (1994). Drug markets and drug places: A case-control study of the spatial structure of illicit drug dealing (Doctoral dissertation), University of Maryland.

Eck, J.E., Chainey, S., Cameron, J.G., Leinter, M. & Wilson, R.E. (2005). *Mapping Crime: Understanding Hot Spots*. Washington, DC: US Department of Justice..

Eck, J.E., Clarke, R.V. & Guerette, R.T. (2007). Risky Facilities: Crime concentrations in homogeneous sets of establishments and facilities. In G. Farrell, K.J. Bowers, S.D. Johnson & M. Townsley. Eds. *Imagination for Crime Prevention: Essays in Honour of Ken Pease*. Crime Prevention Studies, Vol. 21. Monsey, NY: Criminal Justice Press, pp225-64.

Eck, J.E. & Gallagher, K. (2016). Problem-oriented Policing: Evidence v. framing in implementation success. In T.G. Blomberg, J.M. Brancale, K.M. Beaver & W.D. Bales. Eds. *Criminology and Criminal Justice Policy*. London: Routledge. Pp129-38.

Eck, J.E. & Weisburd, D. (1995). Crime places in crime theory. In J.E. Eck & D. Weisburd. Eds. *Crime and Place*. Crime Prevention Studies, Vol. 4, pp1-34. Monsey, NY: Criminal Justice Press.

Ekblom, P. (1997). Gearing up against crime: A dynamic framework to help designers keep up with the adaptive criminal in a changing world. *International Journal of Risk, Security and Crime Prevention*, 2/4: 249-65.

Ekblom, P. (2000). The conjunction of criminal opportunity: A tool for clear ‘joined-up’ thinking about community safety and crime reduction. In S. Ballintyne, K. Pease & V. McLaren. Eds. *Secure Foundations: Key Issues in Crime Prevention, Crime Reduction and Community Safety.* London: Institute for Public Policy Research, pp30-66.

Ekblom, P. (2001). *The Conjunction of Criminal Opportunity: A framework for crime reduction toolkits*. Police and Reducing Crime Unit Report. London: Home Office.

Ekblom, P. & Sidebottom, A. (2007). What do you mean "Is it secure?" Redesigning language to be fit for the task of assessing the security of domestic. *European Journal on Criminal Policy and Research*, 14/1: 61-87.

Ellingworth, D., Farrell, G. & Pease, K. (1995). A victim is a victim is a victim? Chronic victimization in four sweeps of the British Crime Survey. *British Journal of Criminology*, 35/3: 360-5.

Emmison, M. & Smith, P. (2000). *Researching the Visual*. London: Sage.

Evans, D.J., Fyfe, N.R. & Herbert, D.T. (1992). *Crime, Policing and Place.* London: Routledge.

Everson, S. & Pease, K. (2001). Crime Against the Same Person and Place: Detection opportunity and offender targeting. In G. Farrell & K. Pease. Eds. *Repeat Victimization*. Crime Prevention Studies Vol. 12. Monsey, NY: Criminal Justice Press, pp199-220.

"facility" (2018). In *Oxford Living Dictionaries*. Retrieved from https://en.oxforddictionaries.com/definition/facility (28/1/18).

Farrell, G. (2005). Progress and prospects in the prevention of repeat victimization. In N. Tilley. Ed. *Handbook of Crime Prevention and Community Safety*. Cullompton: Willan Publishing, pp143-70.

Farrell, G. & Buckley, A. (1999). Evaluation of a UK police domestic violence unit using repeat victimisation as performance indicator. *The Howard Journal of Crime and Justice*, 38/1: 42-53.

Farrell, G. & Pease, K. (2001). Why repeat victimization matters. In G. Fareel & K. Pease. Eds. *Repeat Victimization*. Crime Prevention Studies, Vol. 12. Monsey, NY: Criminal Justice Press, pp1-4.

Farrell, G. & Pease, K. (2008). Repeat victimization. In R. Wortley, L. Mazerolle. Eds. *Environmental Criminology and Crime Analysis*. Cullompton: Willan, pp117-135.

Farrell, G., Sousa, W. & Weisel, D.L. (2002). The time-window effect in the measurement of repeat victimization: A methodology for its examination, and an empirical study. In N. Tilley. Ed. *Analysis for Crime Prevention*, Crime Prevention Studies 13. New York: Criminal Justice Press, pp15-27.

Farrington, D.P. (2003). Methodological Quality Standards for Evaluation Research. *The Annals of the American Academy of Political and Social Science*, 587/1: 49-68.

Fay, M.P. & Proschan, M.A. (2010). Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics Surveys*, 4: 1-39.

Federal Bureau of Investigation (2011). *Crime in the United States 2010: Clearances.* Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2010/crime-in-the-u.s.-2010/clearances (30/01/18).

Felson, M. (1986). Routine activities, social controls, rational decisions, and criminal outcomes. In D.B. Cornish & R.V.G. Clarke. Eds. *The Reasoning Criminal: Rational Choice Perspectives on Offending*. New York, NY: Springer-Verlag, pp302-337.

Felson, M. (1987). Routine activities and crime prevention in the developing metropolis. *Criminology*, 25/4: 911-32.

Felson, M. (1995). Those Who Discourage Crime. In J.E. Eck & D. Weisburd. Eds. *Crime and Place*. Crime Prevention Studies, Vol. 4. Monsey, NY: Criminal Justice Press, pp53-66.

Felson, M. (2002). *Crime and Everyday Life*. 3rd ed. Thousand Oaks, CA: Sage.

Felson, M. (2008). Routine activities approach. In R. Wortley, L. Mazerolle. Eds. *Environmental Criminology and Crime Analysis*. Cullompton: Willan, pp70-77.

Felson, M. & Cohen, L.E. (1980). Human ecology and crime: A routine activity approach. *Human Ecology*, 8/4: 389-406.

Felson, M. & Poulsen, E. (2003). Simple indicators of crime by time of day. *International Journal of Forecasting*, 19/4: 595-601.

Forrester, D., Chatterton, M. & Pease, K. (1988). *The Kirkholt Burglary Prevention Project, Rochdale.* Crime Prevention Unit Paper 13. London: Home Office.

Foster, P. (2006). Observational research. In R. Sapsford & V. Jupp. Eds. *Data Collection and Analysis*. 2nd ed. London: Sage, pp57-92.

Fowler, F.J. & Mangione, T.W. (1982). *Neighbourhood Crime, Fear and Social Control.* Washington, DC: US Government Printing Office.

Fox, J.A. & Tracy, P.E. (1988). A Measure of Skewness in Offense Distribution. Journal of Quantitative Criminology, 4/3: 259-74.

Franzosi, R.P. (2009). Content analysis. In M. Hardy & A. Bryman. Eds. *The Handbook of Data Analysis*. Paperback ed. London: Sage, pp547-66.

Frisbie, D.W., Fishbine, G., Hintz, R., Joelson, M. & Nutter, J.B. (1977). *Crime in Minneapolis: Proposals for prevention*. St Paul, MN: Governor's Commission on Crime Prevention and Control.

Getis, A. & Ord, J.K. (1992). The Analysis of Spatial Association by use of Distance Statistics. *Geographical Analysis*, 24: 189-206.

Getis, A. & Ord, J.K. (1996). Local spatial statistics: An overview. In P. Longley & M. Batty. Eds. *Spatial Analysis: Modelling in a GIS environment*. Chichester: John Wiley & Sons, pp261-77.

Gill, C., Wooditch, A. & Weisburd, D. (2017). Testing the "Law of Crime Concentration at Place" in a Suburban Setting: Implications for Research and Practice. *Journal of Quantitative Criminology*, 33/3: 549-45.

Gill, C.E. (2011). Missing links: how descriptive validity impacts the policy relevance of randomized controlled trials in criminology. *Journal of Experimental Criminology*, 7: 201-24.

Gill, M. & Clarke, R. (2012). Slowing thefts of fast-moving goods. In P. Ekblom. Ed. *Design Against Crime: Crime proofing everyday products*. Crime Prevention Studies, Vol. 27. Boulder, CO: Lynne Rienner, pp229-238.

Gill, M. & Matthews, R. (1994). Robbers on robbery: Offenders' perspectives. In M. Gill. Ed. *Crime at Work: Studies in security & crime prevention*. Basingstoke: Palgrave Macmillan, pp11-28.

Gill, C., Wooditch, A. & Weisburd, D. (2017). Testing the "law of crime concentration at place in a suburban setting: Implications for research and practice. *Journal of Quantitative Criminology*, 33/3: 519-45.

Gillespie, C.S. (2013). The poweRlaw package: Examples. Retrieved from https://cran.r-project.org/web/packages/poweRlaw/vignettes/b_powerlaw_examples.pdf (02/08/2013 - file since updated)

Gillespie, C.S. (2015). Fitting Heavy Tailed Distributions: The poweRlaw Package. Journal of Statistical Software, 64/2: 1-16. Retrieved from https://www.jstatsoft.org/index.php/jss/article/view/v064i02/v64i02.pdf (12/4/16)

Gillespie, C.S. (2016a). Package 'poweRlaw': Analysis of heavy tailed distributions. Retrieved from https://cran.r-project.org/web/packages/poweRlaw/poweRlaw.pdf (12/4/16)

Gillespie, C.S. (2016b). The poweRlaw package: a general overview. Retrieved from https://cran.r-project.org/web/packages/poweRlaw/vignettes/a_introduction.pdf (12/4/16)

Gillespie, C.S. (2016c). The poweRlaw package: Examples. Retrieved from https://cran.r-project.org/web/packages/poweRlaw/vignettes/b_powerlaw_examples.pdf (12/4/16)

Gillespie, C.S. (2016d). The poweRlaw package: Comparing distributions. Retrieved from https://cran.r-project.org/web/packages/poweRlaw/vignettes/c_comparing_distributions.pdf (12/4/16)

Gini, C. (1912). Variabilità e Mutabilità. Bologna: Tipografia di Paolo Cuppini. Cited by Xu, K. (2004) inf.

Gini, C. (1921). Measurement of Inequality of Incomes. The Economic Journal, 31: 124-6.

Goldsmith, V., McGuire, P.G., Mollenkopf, J.H. & Ross, T. Eds. (2000). *Analyzing Crime Patterns: Frontiers of practice.* Thousand Oaks, CA: Sage.

Goldstein, H. (1990). *Problem-oriented Policing*. New York, NY: McGraw-Hill.

Goldstein, M. L., Morris, S.A. & Yen, G.G. (2004). Problems with fitting to the power-law distribution. The European Physical Journal B, 41: 255-8. DOI: 10.1140/epjb/e2004-00316-5 Retrieved from http://download.springer.com/static/pdf/958/art%253A10.1140%252Fepjb%252Fe2004-

00316-
5.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1140%2Fepjb%2Fe2004-00316-5&token2=exp=1460459420~acl=%2Fstatic%2Fpdf%2F958%2Fart%25253A10.1140%25252Fepjb%25252Fe2004-00316-5.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1140%252Fepjb%252Fe2004-00316-5*~hmac=379e7f0f2db2babfa31728692923b97f7415b50e5002022ec78aab69ec3907f5 (12/4/16).

Gore, W. & Lee, Y. (2018). Chronic and temporary crime hot spots. In D. Weisburd & J.E. Eck. Eds. *Unraveling the Crime-Place Connection: New Directions in Theory and Policy.* New Brunswick, NJ: Transaction Publishers, pp41-63.

Gottfredson, M.R. & Hirschi, T. (1990). *A General Theory Crime*. Palo Alta: Stanford University Press.

Graham, K., Bernards, S., Osgood, D.W., Homel, R. & Purcell, J. (2005). Guardians and handlers: The role of bar staff in preventing and managing aggression. *Addiction*, 100/6: 755-66.

Graham, K. & Homel, R. (2008). *Raising the Bar: Preventing aggression in and around bars, pubs, and clubs*. Cullompton: Willan.

Greene, J.R. (2014). New directions in policing: Balancing prediction and meaning in police research. *Justice Quarterly*, 31/2: 193-228.

Groff, E. (2011). Exploring ‘near’: Characterizing the spatial extent of drinking place influence on crime. *Australian and New Zealand Journal of Criminology*, 44/2: 156-79.

Groff, E., Weisburd, D. & Morris, N.A. (2009). Where the action is at places: Examining spatio-temporal patterns of juvenile crimes at places using trajectory analysis and GIS. In D. Weisburd, W. Bernasco & G.J. Bruinsma. Eds. *Putting Crime in its Place; Units of Analysis in Spatial Crime Research*. New York, NY: Springer, pp61-86.

Groff, E.R., Weisburd, D. & Yang, S-M. (2010). Is it important to examine crime trends at a local “micro” level?: A longitudinal analysis of street to street variability in crime trajectories. *Journal of Quantitative Criminology*, 26/1: 7-32.

Grove, J.M. & Burch, W.R. Jr (1997). A social ecology approach and applications of urban ecosystem and landscape analyses: a case study of Baltimore, Maryland. *Urban Ecosystems*, 1: 259-75.

Grubesic, T.H. (2006). On the Application of Fuzzy Clustering for Crime Hot Spot Detection. *Journal of Quantitative Criminology*, 22(1): 77-105.

Guardian online (2017). Betsy DeVos to overhaul ‘failed’ guidelines on campus sexual assault. Retrieved from https://www.theguardian.com/education/2017/sep/07/betsy-devos-campus-sexual-assault-policy (14/1/17).

Guerry, A-M (1833). *Essai sur la statistique morale de la France.* Paris: Crochard.

Guerrry, A-M (1864). *Statistique morale de l'Anglettere comparee avec la statistique morale de la France, d'apres les comptes de l'administration de la justice criminelle en Anglettere et en France.* Paris: J-B. Bailliere et fils.

Habermann, C.P., Sorg, E.T. & Ratcliffe, J.H. (2017). Assessing the validity of the law of crime concentration across different temporal scales. *Journal of Quantitative Criminology*, 33/3: 547-67.

Hardy, M. (2009). Summarizing distributions. In M. Hardy & A. Bryman. Eds. *The Handbook of Data Analysis*. Paperback ed. London: Sage, pp35-64.

Hart, A. (2001). Mann-Whitney test is not just a test of medians: differences in spread can be important. *British Medical Journal*, 323/7309: 391-3.

Hart, T. & Zandbergen, P. (2014). Kernel density estimation and hotspot mapping: Examining the influence of interpolation method, grid cell size, and bandwidth on crime forecasting. *Policing: An International Journal of Police Strategies & Management*, 37/2: 305-23.

Hart, T.C. & Zandbergen, P.A. (2012). *Effects of Data Quality on Predictive Hotspot Mapping*, Washington, DC: US National Institute of Justice. Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/239861.pdf (08/01/18).

Henry, S. (2007). Deviance: Constructionist perspectives. In G. Ritzer. Ed. *The Blackwell Encyclopedia of Sociology*. Oxford: Blackwell, pp145-63.

Herrmann, C.R. (2015). The dynamics of robbery and violence hot spots. *Crime Science*, 4:33. Retrieved from https://link.springer.com/article/10.1186/s40163-015-0042-5 (26/6/18).

Hillier, B. & Hanson, J. (1984). *The Social Logic of Space.* Cambridge: Cambridge University Press.

Hinton, P. R. (2014). *Statistics Explained*. 3rd ed. Hove: Routledge.

Hirschfield, A. & Bowers, K.J. (1997). The effect of social cohesion on levels of recorded crime in disadvantaged areas. *Urban Studies*, 34/8: 1275-95.

Hirschfield, A. & Bowers, K. Eds. (2001). *Mapping and Analysing Crime Data: Lessons from research and practice.* London: Taylor & Francis.

Hirschfield, A., Yarwood, D. & Bowers, K. (2001). Crime Pattern Analysis, Spatial Targeting and GIS: The development of new approaches for use in evaluating community safety initiatives. In G. Clarke & M. Madden. Eds. *Regional Science in Business*. Berlin: Springer-Verlag.

Hirschi, T. (1969). *Causes of Delinquency*. New Brunswick, NJ: Transaction Publishers.

Homel, R., Carvolth, R., Hauritz, M., McIlwain, G. & Teague, R. (2004). Making licensed venues safer for patrons: What environmental factors should be the focus of interventions? *Drug and Alcohol Review*, 23/1: 19-29.

Home Office (2008). The Prolific and Other Priority Offenders Strategy. Available at webarchive.nationalarchives.gov.uk/20081023143640/http://crimereduction.homeoffice.gov.uk/ppominisite01.htm (last accessed 30/5/17).

Home Office (2009). Prolific and Other Priority Offender Programme: Five Years on: Maximising the impact. London: Ministry of Justice. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/118035/PPO-5-years-on.pdf (accesed 30/5/17).

Hope, T. (1986). Liquor licensing and crime prevention. Home Office Research Bulletin No. 20. London: Home Office.

Hunter, R.D. (1999). Convenience store robbery revisited: A review of prevention results. *Journal of Security Administration*, 22/1: 1-13.

Ignatans, D. & Pease, K. (2015). Distributive justice and the crime drop. In M. Andersen & G. Farrell. Eds. *The Criminal Act: Festschrift for Marcus Felson*. London: Palgrave Macmillan, pp77-87.

Ignatans, D. & Pease, K. (2016). On whom does the burden of crime fall now? Changes over time in counts and concentration. *International Review of Victimology*, 22(1): 55-63.

Ignatans, D., & Roebuck, T. (2018). Do More Immigrants Equal More Crime? Drawing a Bridge Between First Generation Immigrant Concentration and Recorded Crime Rates. *Crime, Security and Society*, 1(1). DOI: 10.5920/css.2018.03.

Jacobs, J. (1961). *The Death and Life of Great American Cities.* Harmonsworth: Penguin.

Jeffery, C.R. (1971). *Crime Prevention Through Environmental Design.* California: Sage.

Johnson, S.D. & Bowers, K.J. (2004). The burglary as clue to the future: The beginnings of prospective hot-spotting. *European Journal of Criminology*, 1/2: 237-55.

Johnson, R. & Onwuegbuzie, A.J. (2004). Mixed Methods Research: A research paradigm whose time has come. *Educational Researcher*, 33/7: 14-26.

Johnson, R.B., Onwuegbuzie, A.J. & Turner, L.A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1/2: 112-33.

Kiess, H.O. (1996). Statistical Concepts for the Behavioural Sciences. 2nd ed. Needham Heights, MA: Allyn & Bacon.

Kirk, M.G. (2008). The spatio-temporal effects of spectator events on crime (Doctoral dissertation), Simon Fraser University, Vancouver.

Kleemans, E.R. (2001). Repeat Burglary Victimisation: Results of empirical research in the Netherlands. In G. Farrell & K. Pease. Eds. *Repeat Victimization*. Crime Prevention Studies Vol. 12. Monsey, NY: Criminal Justice Press, pp53-68.

Kooi, B.R. (2007). *Policing Public Transportation: An environment and procedural evaluation of bus stops*. New York, NY: LFB Scholarly Publishing.

Kooi, B. (2015). Security concerns at hot-spot bus stop locations. *Journal of Applied Security*

Research, 10/3: 277-307.

Kposowa, A.J. & Breault, K.D. (1998). Temporal factors and homicide in the United States. *Journal of Interpersonal Violence*, 13/5: 590-604.

Kruskal, W.H. & Wallis, W.A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47/260: 583-621.

Kubrin, C.E. & Weitzer, R. (2003). New directions in social disorganization theory. *Journal of Research in Crime and Delinquency*, 40: 347-402.

Kumar, N. & Waylor, C.R.M. (2002). Proximity to alcohol-serving establishments and crime probabilities in Savannah, Georgia: A statistical and GIS analysis. *Southeastern Geographer*, 42/2: 125-42.

Lagrange, T.C. (1999). The impact of neighborhoods, schools, and malls on the spatial distribution of property damage. *Journal of Research in Crime and Delinquency*, 26/4: p393-422.

Landau, S. & Fridman, D. (1993). The seasonality of violent crime: The case of robbery and homicide in Israel. *Journal of Research in Crime and Delinquency*, 30/2: 161-91.

Lander, B. (1954). *Towards an understanding of juvenile delinquency.* New York, NY: Columbia University Press.

La Vigne, N.G. (2007). *Traffic Congestion Around Schools*. COPS Problem-specific guides series, No. 50. Washington, DC: US Department of Justice.

Lee, Y.J., Eck, J.E., SooHyun, O. & Martinez, N.N. (2017). How concentrated is crime at places? A systematic review from 1970 to 2015. *Crime Science*, 6(6). DOI 10.1186/s40163-017-0069-x. Retrieved from https://crimesciencejournal.springeropen.com/articles/10.1186/s40163-017-0069-x. (accessed 30/5/17).

Levin, A., Rosenfeld, R. & Deckard, M. (2017). The law of crime concentration: An application and recommendations for future research. *Journal of Quantitative Criminology*, 33/3: 635-47.

Levine, N. and Associates (2002). *Crimestat: A Spatial Statistics Program for the Analysis of Crime Incident Locations*. Washington, DC: US National Institute of Justice.

Lindström, P. (1997). Patterns of School Crime: A Replication and Empirical Extension. British Journal of Criminology, 37(1): 121-30.

Loukaitou-Sidaris, A., Liggett, R. & Iseki, H. (2002). The geography of transit crime: Documentation and evaluation of crime incidence on and around the Green Line stations in Los Angeles. *Journal of Planning, Education and Research*, 22/2: 135-51.

Lu, Y. (2003). Getting away with the stolen vehicle: An investigation into journey-after-crime. *The Professional Geographer*, 55/4: 422-33.

Macintyre, S. & Homel, R. (1997). Danger on the dance floor: A study of interior design, crowding and aggression in nightclubs. In R. Homel. Ed. *Policing for Prevention: Reducing crime, public intoxication and injury*, Crime Prevention Studies Vol. 7. Money, NY: Criminal Justice Press, pp91-113.

Madensen, T.D. (2007). Bar management and crime: Toward a dynamic theory of place management and crime hotspots. Doctoral thesis submitted to University of Cincinnati, Ohio.

Madensen, T. D. & Eck, J.E. (2008). Violence in bars: Exploring the impact of place manager decision-making. *Crime Prevention and Community Safety*, 10(2): 111-25.

Mago, V.K., Frank, R., Reid, A. & Dabbaghian, V. (2014). The strongest does not attract all but it does attract the most – evaluating the criminal attractiveness of shopping malls using fuzzy logic. *Expert Systems*, 31: 121-35.

Maguire, M. & McVie, S. (2017). Crime data and criminal statistics: A critical reflection. In A. Liebling, S. Maruna & L. McAra. Eds. *The Oxford Handbook of Criminology*. 6th ed. Oxford: Oxford University Press, pp163-89.

Mailley, J., Whitehead, S. & Farrell, G. (2006). Progress and prospects in the prevention of mobile phone theft. *Justice of the Peace*, 170: 404-7.

Mann, H.B. & Whitney, D.R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18/1: 50-60.

Massey Jr, F.J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46/253: 68-78.

Matsueda, R., Kreager, D. & Huizinga, D. (2006). Deterring delinquents: A rational choice model of theft and violence. *American Sociological Review*, 71/1: 95-122.

Mawby, R.I. (1977). Defensible Space. *Urban Studies*, 14: 169-180.

May, T. (2011). *Social Research: Issues, Methods and Process.* 4th ed. Maidenhead: Open University Press.

Mayhew, H. (1862). *London Labour and the London Poor.* London: Cass.

Mayhew, P. (1991). Crime in public view: Surveillance and crime prevention. In P.J. Brantingham & P.L. Brantingham. Eds. *Environmental Criminology*. Beverley Hills, CA: Sage. (Reprinted by Waveland Press, Prospect Heights, Il).

Mayhew, S. (2015). power law. In A Dictionary of Geography. Oxford: Oxford University Press. Retrieved from http://www.oxfordreference.com.libaccess.hud.ac.uk/view/10.1093/acref/9780199680856.001.0001/acref-9780199680856-e-4064. (21/6/18)

Mazeika, D. & Summerton, D. (2017). The impact of geocoding method on the positional accuracy of residential burglaries reported to the police. *Policing: An International Journal of Police Strategies and Management*, 40(2): 459-70.

McCleary, R. & Chew, K.S.Y. (2002). Winter is the infanticide season: Seasonal risk for child homicide. *Homicide Studies*, 6/3: 228-39.

McCord, E.S. & Ratcliffe, J.H. (2007). A micro-spatial analysis of the demographic and criminogenic environment of drug markers in Philadelphia. *Australian and New Zealand Journal of Criminology*, 40/1: 43-63.

McCord, E.S., Ratcliffe, J.H., Garcia, M. & Taylor, R.B. (2007). Nonresidential crime attractors and generators elevate perceived neighbourhood crime and incivilities. *Journal of Research in Crime and Delinquency*, 44/3: 295-320.

McKenzie, R.D. (1925). The ecological approach to the study of the human community. In R.E. Park, E.W. Burgess & R.D. McKenzie. Eds. *The City*. Chicago, Il: University of Chicago Press, pp63-79.

Miller, J. Caplan, J.M. & Ostermann, M. (2016). Assessing the Effects of Local Crime Hotspots on Parole Recidivism. *The Prison Journal*, 96(3): 437-61.

Mohd Razali, N. & Yap, B. (2011). Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *Journal of Statistical Modeling and Analytics*, 2/1: 21-33.

Morgan, F. (2001). Repeat Burglary in a Perth Suburb: Indicator of short-term or long-term risk? In In G. Farrell & K. Pease. Eds. *Repeat Victimization*. Crime Prevention Studies Vol. 12. Monsey, NY: Criminal Justice Press, pp83-118.

Morris, T. (1957). *The Criminal Area*. London: Routledge and Kegan Paul.

Nelson, A.L., Bromley, R.D.F. & Thomas, C.J. (2001). Identifying micro-spatial and temporal patterns of violent crime and disorder in the British city centre. *Applied Geography*, 21/3: 249-74.

Newman, M. (2017). Power-law distribution. *Significance*, 14: 10-11. Retrieved from https://rss-onlinelibrary-wiley-com.libaccess.hud.ac.uk/doi/full/10.1111/j.1740-9713.2017.01050.x (21/6/18).

Newman, O. (1972). *Defensible Space*. New York, NY: Macmillan.

Newton, A. (2004). Crime and disorder on buses: towards an evidence base for effective crime prevention (Doctoral dissertation), University of Liverpool.

Newton, A. (2008). A study of bus route crime risk in urban areas: The changing environs of a bus journey. *Built Environment*, 34/1: 88-103.

Newton, A.D., Partridge, H. & Gill, A. (2014). In and around: Identifying predictors of theft within and near to major mass underground transit. *Security Journal*, 27/2: 132-46.

Oberwittler, D. & Wikström, P-O.H. (2009). Why small is better: Advancing the study of the tole of behavioral contexts in crime causation. In D. Wesiburd, W. Bernasco & G.J. Bruinsma. Eds. *Putting Crime in its Place*. New York, NY: Springer, pp35-59.

Office for National Statistics (2017). Crime in England and Wales: year ending Dec 2016. London: ONS. Available at https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingdec2016/pdf (accessed 30/5/17).

Openshaw, S. (1984). The modifiable areal unit problem. *Concepts and Techniques in Modern Geography*, 38: 1-41.

Ord, J.K. & Getis, A. (1995). Local Spatial Autocorrelation Statistics: Distributional issues and an application. *Geographical Analysis,* 27: 286-306.

Park, R.E., Burgess, E.W. & McKenzie, R.D. (Eds.) (1925). *The City.* Chicago, Il: University of Chicago Press.

Pawson, R. & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.

Pease, K. (1998). *Repeat Victimisation: Taking stock*. London: Home Office.

Pires, S. & Clarke, R. (2012). Are parrots CRAVED? An analysis of parrot poaching in Mexico. *Journal of Research in Crime and Delinquency*, 49(1): 122-46.

Piscitelli, F. & Albanese, J.S. (2000). Do casinos attract criminals? A study at the Candian-US border. *Journal of Contemporary Criminal Justice*, 16/4: 445-56.

Piquero, A.R. (2000). Assessing the relationships between gender, chronicity, seriousness, and offense skewness in criminal offending. Journal of Criminal Justice, 28: 103-15.

Piquero, A.R. & Buka, S.L. (2002). Linking juvenile and adult patterns of criminal activity in the Providence cohort of the National Collaborative Perinatal Project. Journal of Criminal Justice, 30: 259-72.

Polvi, N. Looman, T., Humphries, C. & Pease, K. (1991). The time course of repeat burglary victimization. *British Journal of Criminology*, 31/4: 411-4.

Poyner, B. (2005). *Crime-free Housing in the 21st Century.* Cullompton: Willan Publishing.

Power, A. (1989). Housing, Community and Crime. In D. Downes. Ed. *Crime and the City.* Basingstoke: Macmillan.

“premises” (2018). In *Oxford Living Dictionaries*. Retrieved from https://en.oxforddictionaries.com/definition/premises (28/1/18).

Pratt, T.C. & Cullen, F.T. (2005). Assessing macro-level predictors and theories of crime: A meta-analysis. *Crime and Justice*, 32: 373-450.

Pyle, G.F. (1976). Spatial and Temporal Aspects of Crime in Cleveland, Ohio. *American Behavioural Scientist*, 20: 180-183.

Quetelet, A. (1831). Research on the propensity for crime at different ages. Reproduced in P. Bean (2003). Ed. *Crime: Critical Concepts in Sociology.* London: Routledge.

Quigley, B.M., Leonard, K.E. & Collins, R.L. (2003). Characteristics of violent bars and bar patrons. *Journal of Studies on Alcohol*, 64/6: 765-72.

Ratcliffe, J.H. (2000). Aoristic analysis: the spatial interpretation of unspecific temporal events. *International Journal of Geographical Information Science*, 14/7: 669-79.

Ratcliffe, J.H. (2002). Aoristic signatures and the spatio-temporal analysis of high volume crime patterns. *Journal of Quantitative Criminology*, 18/1: 23-43.

Ratcliffe, J.H. (2004). Geocoding crime and a first estimate of a minimum acceptable hit rate. *International Journal of Geographical Information Science*, 18(1): 61-72.

Ratcliffe, J.H. (2006). A temporal constraint theory to explain opportunity-based spatial offending patterns. *Journal of Research in Crime and Delinquency*, 43/3: 261-91.

Ratcliffe, J.H. (2012). The spatial extent of criminogenic places: A changepoint regression of violence around bars. *Geographical Analysis,* 44: 302-20.

Ratcliffe, J.H. & McCullagh, M.J. (1999). Hotbeds of Crime and the Search for Spatial Accuracy. *Journal of Geographical Systems,* 1: 385-98.

Ratcliffe, J.H. & McCullagh, M.J. (2001). Chasing Ghosts? Police perceptions of high crime areas. *British Journal of Criminology*, 41: 330-41.

Ratcliffe, J.H. & Rengert, G.F. (2008). Near-repeat patterns in Philadelphia shootings. *Security Journal*, 21/1-2: 58-76.

Raudenbush, S.W. & Sampson, R.J. (1999). Ecometrics: Toward a science of assessing ecological settings, with application to the systematic social observation of neighbourhoods. *Sociological Methodology*, 29/1: 1-41.

Rengert, G.F. (1992). The journey to crime: Conceptual foundations and policy implications. In D.J. Evans, N.R. Fyfe & D.T. Herbert. Eds. *Crime, Policing, and Place: Essays in environmental criminology*. Abingdon: Routledge, pp109-17.

Rengert, G. (2004). The journey to crime. In G. Bruinsma, H. Elffers & J. Keijser. Eds. *Punishment, Places and Perpetrators: Developments in Criminology and Criminal Justice Research*. Cullompton: Willan Publishing, pp169-81.

Rengert, G.F. & Lockwood, B. (2009). Geographical units of analysis and the analysis of crime. In D. Wesiburd, W. Bernasco & G.J. Bruinsma. Eds. *Putting Crime in its Place*. New York, NY: Springer, p109-22.

Rengert, G.F., Piquero, A.R. & Jones, P.R. (1999). Distance decay re-examined. *Criminology*, 37: 427-45.

Repetto, T. (1974). *Residential Crime*. Cambridge, MA: Ballinger.

Reynald, D.M. (2010). Guardians on guardianship: Factors affecting the willingness to supervise, the ability to detect potential offenders, and the willingness to intervene. *Journal of Research in Crime and Delinquency*, 47/3: 358-90.

Reynald, D.M. (2011). Factors associated with the guardianship of places: Assessing the relative importance of the spatio-physical and sociodemographic contexts in generating opportunities for capable guardians. *Journal of Research in Crime and Delinquency*, 48/1: 110-42.

Rhodes, W.M. & Conly, C. (1981). Crime and mobility: An empirical study. In P.J. Brantingham & P.L. Brantingham. Eds. *Environmental Criminology*. Beverly Hills, CA: Sage, pp167-88.

Roach, J. & Pease, K. (2013). *Evolution and Crime*. London: Routledge.

Robson, C. & McCartan, K. (2016). *Real World Research*. 4th ed. Chichester: Wiley.

Rock, D., Greenberg, D.M. & Hallmayer, J. (2003). Cyclical changes in homicide rates: A reanalysis of Brearley's 1932 data. *Journal of Interpersonal Violence*, 18/8: 942-55.

Roman, C.G. (2003). Schools as Generators of Crime: Routine activities and the sociology of place (Doctoral dissertation), American University, Washington, DC.

Roncek, D.W. (2000). Schools and crime. In V. Goldsmith, P.G. McGuire, J.H. Mollenkopf & T.A. Ross. Eds. *Analyzing Crime Patterns: Frontiers of practice*. Thousand Oaks, CA: Sage.

Roncek, D.W. & Bell, R. (1981). Bars, blocks and crime. *Journal of Environmental Systems*, 11: 35-47.

Roncek, D.W. & Faggiani, D. (1985). High schools and crime: A replication. *The Sociological Quarterly*, 26/4: 491-505.

Ronceck, D.W. & Lobosco, A. (1983). The effect of high schools on crime in their neighbourhoods. *Social Science Quarterly*, 64/3: 598-613.

Roncek, D.W. & Maier, P.A. (1991). Bars, blocks and crime revisited: Linking the theory of routine activities to the empiricism of "hot spots". *Criminology*, 29/4: 725-53.

Roncek, D.W. & Pravatiner, M.A. (1989). Additional evidence that taverns enhance nearby crime. *Socioogy and Social Research*, 73/4: 185-8.

Rossmo, D.K. (2000). *Geographic Profiling*. Boca Raton, FL: CRC Press.

Rossmo, D.K. & Rombouts, S. (2008). Geographic Profiling. In R. Wortley, L. Mazerolle. Eds. *Environmental Criminology and Crime Analysis*. Cullompton: Willan, pp136-149.

Rotton, J. & Cohn, E.G. (2000). Weather, disorderly conduct, and assaults: From social contract to social avoidance. *Environment and Behavior*, 32: 642-71.

Rotton, J. & Cohn, E. (2001). Weather, disorderly conduct, and assaults: From social contact to social avoidance. *Environment and Behaviour*, 32/5: 651-73.

Sallybanks, J. & Brown, R. (1999). *Vehicle Crime Reduction: Turning the corner.* Police Research Series, Paper 119. London: Home Office.

Sampson, R.J. (2012). *Great American city: Chicago and the enduring neighbourhood effect.* Chicago, Il: University of Chicago Press.

Sampson, R.J. & Groves, W.B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology*, 94: 774-802.

Sampson, R.J. & Raudenbush, S.W. (2004). Seeing disorder: Neighbourhood stigma and the social construction of "broken windows". *Social Psychology Quarterly*, 67/4: 319-42.

Sampson, R.J., Raudenbush, S.W. & Earls, F. (1997). Neighbourhoods and violent crime: A multi-level study of collective efficacy. *Science*, 277: 918-24.

SBD (n.d.). About. Retrieved from http://www.securedbydesign.com/about-secured-by-design/

Seaman, M.A., Levin, J.R. & Serlin, R.C. (1991). New developments in pairwise multiple comparisons: Some powerful and practicable procedures. *Psychological Bulletin*, 110/3: 577-86.

Shapiro, S.S. & Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52/3: 591-611.

Shaw, C.R. & McKay, H.D. (1931). *Social Factors in Juvenile Delinquency: A Study of the Community, the Family, and the Gang in Relation to Delinquent Behavior, for the National Commission on Law Observance and Enforcement.* US Government Printing Office.

Shaw, C.R. & McKay, H.D. (1942). *Juvenile delinquency and urban areas.* Chicago, Il: University of Chicago Press.

Sherman, L.W. (1995). Hot Spots of Crime and the Criminal Careers of Places. In J.E. Eck & D. Weisburd. Eds. *Crime and Place*. Crime Prevention Studies, Vol. 4. Monsey, NY: Criminal Justice Press, pp35-52.

Sherman, L.W., Gartin, P.R. & Buerger, M.E. (1989). Hot Spots of Predatory Crime: Routine activities and the criminology of place. *Criminology*, 27(1): 27-55.

Sherman, L.W. & Weisburd, D. (1995). General deterrent effects of police patrol in crime "hot spots": A randomized, controlled trial. *Justice Quarterly*, 12(4): 625-48.

Sidebottom, A. (2013). On the application of CRAVED to livestock theft in Malawi. *International Journal of Comparative and Applied Criminal Justice*, 37(3): 195-212.

Sidebottom, A. & Bowers, K. (2010). Bag Theft in Bars: An analysis of relative risk, perceived risk and *modus operandi*. *Security Journal*, 23(3): 206-24.

Sidebottom, A. & Tilley, N. (2012). Further improving reporting in crime and justice: an addendum to Perry, Weisburd and Hewitt (2010). *Journal of Experimental Criminology*, 8/1: 49-69.

Skogan, W.G. (1977). Dimensions of the dark figure of unreported crime. *Crime & Delinquency*, 23/1/: 41-50.

Smith, C., Bowers, K.J. & Johnson, S.D. (2006). Understanding bag theft within licensed premises in Westminster: Identifying initial steps towards prevention. *Security Journal*, 19: 3-21.

Smith, S.J. (1986). *Crime, Space and Society.* Cambridge: Cambridge University Press.

Smith, W., Bond, J. & Townsley, M. (2009). Determining how journeys-to-crime vary: Measuring inter- and intra-offender crime trip distributions. In D. Weisburd, W. Bernasco & G. Bruinsma. Eds. *Putting Crime in its Place: Units of Analysis in Spatial Crime Research*. New York, NY: Springer, pp217-36.

Spelman, W. (1993). Abandoned buildings: Magnets for crime. *Journal of Criminal Justice*, 21/5: 481-95.

Spelman, W. & Eck, J.E. (1987). *Problem-Oriented Policing*. National Institute of Justice Research in Brief. Washington, DC: US Department of Justice.

Spelman, W. & Eck, J.E. (1989). *Sitting ducks, ravenous wolves and helping hands: new approaches to urban policing*. Austin, TX: Lyndon B. Johnson.

St. Jean, P.K.B. (2007). *Pockets of Crime: Broken windows, collective efficacy, and the criminal point of view*. Chicago, IL: University of Chicago Press.

Steenbeek, W. & Hipp, J.R. (2011). A longitudinal test of social disorganization theory: Feedback effects among cohesion, social control, and disorder. *Criminology*, 49/3: 833-71.

Steenbeek, W. & Weisburd, D. (2016). Where the action is in crime? An examination of variability of crime across different spatial units in the Hague, 2001-2009. *Journal of Quantitative Criminology*, 32/3: 449-69.

Stitt, B.G., Nichols, M. & Giacopassi, D. (2003). Does the presence of casinos increase crime? An examination of casino and control communities. *Crime & Delinquency*, 49/2: 253-84.

Stumpf, M.P.H. & Porter, M.A. (2012). Critical Truths About Power Laws. Science, 335: 665-6. DOI: 10.1126/science.1216142 Retrieved from https://people.maths.ox.ac.uk/porterm/papers/critical.pdf (12/4/16).

Sykes, G.M. & Matza, S. (1957). Techniques of Neutralization: A Theory of Delinquency. *American Sociological Review*, 22/6: 664-70.

Taylor, N. (2002). *Robbery against service stations and pharmacies: recent trends*. Canberra: Australian Institute for Criminology.

Taylor, R.B. (1998). Crime and small scale places: What we know, what we can prevent, and what else we need to know. In R.B. Taylor, G. Bazemore, B. Boland, T.R. Clear, R.P.J. Corbett, J. Feinblatt, G. Berman, M. Sviridoff & C. Stone. Eds. *Crime and Place: Plenary papers of the 1997 conference on criminal justice research and evaluation*. Washington, DC: National Institute of Justice, pp1-22.

Teddie, C. & Tashakkori, A. (2009). *Foundations of Mixed Methods Research*. London: Sage.

Telep, C. (2018). Not just what works, but *how* it works: Mechanisms and context in the effectiveness of place-based policing. In D. Weisburd & J.E. Eck. Eds. *Unraveling the Crime-Place Connection*. London: Routledge, pp237-59.

Thomas, G. (2017). *How to Do Your Research Project*. 3rd ed. London: Sage.

The R Foundation (n.d.). What is R? Retrieved from https://www.r-project.org/about.html (12/4/2016).

Tilley, N., Pease, K., Hough, M. & Brown R. (1999). *Burglary Prevention: Early lessons from the Crime Reduction Programme*. Crime Reduction Research Series Paper 1. London: Home Office.

Tita, G.E. & Greenbaum, R.T. (2009). Crime, neighbourhoods, and units of analysis: Putting space in its place. In D. Wesiburd, W. Bernasco & G.J. Bruinsma. Eds. *Putting Crime in its Place*. New York, NY: Springer, pp145-70.

Toby, J. (1957). Social disorganization and stake in conformity: Complementary factors in the

predatory behaviour of hoodlums. *Journal of Criminal Law and Criminology*, 48/1: 12-7.
Tompson, L. & Bowers, K.J. (2015). Testing time-sensitive influences of weather on street robbery. *Crime Science*, 4/8. Retrieved from https://crimesciencejournal.springeropen.com/track/pdf/10.1186/s40163-015-0022-9
Townsley, M. & Farrell, G. (2007). Repeat victimization of prison inmates. In G. Farrell, K.J. Bowers, S.D. Johnson. & M. Townsley Eds. *Imagination for Crime Prevention: Essays in Honour of Ken Pease*. Crime Prevention Studies Vol. 21. New York, NY: Criminal Justice Press, pp265-276.
Townsley, M., Homel, R. & Chaseling, J. (2003). Infectious burglaries: A test of the near repeat hypothesis. *British Journal of Criminology*, 43/3: 615-33.
Townsley, M. & Sidebottom, A. (2010). All offenders are equal, but some are more equal than others: variation in journeys to crime between offenders. *Criminology*, 48/3: 897-917.
Tseloni, A. & Pease, K. (1997). Repeat Victimisation and Policing: Whose home gets victimised, where and how often? In P. Wikstrom, W. Skogan & L. Sherman. Eds. *Problem Oriented Policing and Crime Prevention*. New York, NY: Wadsworth.
Tseloni, A. & Pease, K. (2003). Repeat personal victimization: 'Boosts' or 'flags'? *The British Journal of Criminology*, 43/1: 196-212.
Van Daele, S. & Vander Beken, T. (2011). Outbound offending: The journey to crime and crime sprees. *Journal of Environmental Pscyhology*, 31/1: 70-8.
Vandeviver, C. (2013). Distance matters: A look at crime trip distances. In A. Crawford, J de Maillard, J. Shapland, A. Verhage & P. Ponsaers. Eds. *Crime, violence, justice and social order: Monitoring contemporary security issues*. Antwerp: Maklu, pp229-53.
Vandeviver, C., Van Daele, S. & Vander Beken, T. (2015). What makes long crime trips worth undertaking? Balancing costs and benefits in burglars' journey to crime. *British Journal of Criminology*, 55/2: 399-420.
van Koppen, P.J. & Jansen, R.W.J. (1999). The time to rob: variations in time of number of commercial robberies. *Journal of Research in Crime and Delinquency*, 36/1: 7-29.
van Koppen, P. & De Keijser, J. (1997). Desisting distance decay: On the aggregation of individual crime trips. *Criminology*, 35: 505-15.
Van Ourti, T. & Clarke, P. (2011). A Simple Correction to Remove the Bias of the Gini Coefficient Due to Grouping. The Review of Economics and Statistics, 93/3: 982-94.
Webb, B. (1997). Steering Column Lock and motor Vehicle Theft: Evaluations from three countries. In R.V. Clarke. Ed. *Situational Crime Prevention: Successful Case Studies.* 2nd ed. New York, NY: Harrow and Heston,
Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology*, 53(2): 133-57.
Weisburd, D., Bruinsma, G.J.N. & Bernasco, W. (2009). Units of analysis in geographic

criminology: Historical development, critical issues, and open questions. In D. Wesiburd, W. Bernasco & G.J. Bruinsma. Eds. *Putting Crime in its Place*. New York, NY: Springer, pp3-31.

Weisburd, D., Bushway, S., Lum, C. & Yang, S.-M. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of Seattle. *Criminology*, 42: 283-322.

Weisburd, D. & Eck, J.E. (2018). Theoretical foundations and frontiers of understanding high crime places: An introduction. In D. Weisburd & J.E. Eck. Eds. *Unraveling the crime-place connection.* London: Routledge, pp1-15.

Weisburd, D., Groff, E., Yang, S-M. (2012). *The Criminology of Place: Street segments and our understanding of the crime problem*. Oxford: Oxford University Press.

Weisburd, D. & McEwen, T. Eds. (1997). *Crime Mapping and Crime Prevention*. Monsey, NY: Criminal Justice Press.

Weisburd, D., Shay, M., Amram, S. & Zamir, R. (2018). The relationship between social disorganization and crime at the micro geographic level: Findings from Tel Aviv-Yafo using Israeli census data. In D. Weisburd & J.E. Eck. Eds. *Unraveling the Crime-Place Connection*. London: Routledge, pp97-121.

Weisburd, D., Wyckoff, L.A., Ready, J., Eck, J.E., Hinkle, J.C. and Gajewski, F. (2006). Does crime just move around the corner? A controlled study of spatial displacement and diffusion of crime control benefits. *Criminology*, 44/3: 549-92.

Weisel, D.L. (2005). *Analyzing Repeat Victimization*. COPS Problem-solving Tools Series, No.4. Washington, DC: US Department of Justice.

Weisel, D. L. (2007). *Bank Robbery*. COPS Problem-specific Guides Series, No.48. Washington, DC: US Department of Justice.

Wellsmith, M. & Burrell, A. (2005). The Influence of Purchase Price and Ownership Levels on Theft Targets: The example of domestic burglary. *British Journal of Criminology,* 45(5): 741-64.

Wellsmith, M., Birks, D. & Donkin, S. (2007). Applying the concept of risky facilities to problems of violence in the night-time economy. Project Report. UCL Jill Dando Institute of Crime Science, London.

White, R.C. (1932). The relation of felonies to environmental factors in Indianapolis. *Social Forces*, 10: 498-509.

White, R. (2008). *Crimes Against Nature: Environmental Criminology and Ecological Justice*. Cullompton: Willan.

Whitehead, S., Mailley, J., Storer, I.J., McCardle, J., Torrens, G.E. & Farrell, G. (2007). In safe hands: A review of mobile phone anti-theft designs. *European Journal on Criminal Policy and Research*, 14/1: 39-60.

The White House (2011). Vice President Biden Announces New Administration Efforts to Help

Nation's Schools Address Sexual Violence. Retrieved from https://obamawhitehouse.archives.gov/the-press-office/2011/04/04/vice-president-biden-announces-new-administration-effort-help-nation-s-s (14/01/18).

Wikström, P-O. (1991). *Urban Crime, Criminals and Victims.* New York: Springer-Verlag.

Wilcox, P. & Eck, J.E. (2011). Criminology of the unpopular. *Criminology and Public Policy*, 10(2): 473-82.

Wilcox, P., Quisenberry, N., Cabrera, D.T. & Jones, S. (2004). Busy places and broken windows? Toward defining the role of physical structure and process in community crime models. *Sociological Quarterly*, 45: 185-207.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1/6: 80-3.

Wiles, P. & Costello, A. (2000). *The Road to Nowhere: Evidence for travelling criminals*. Home Office Research Study No. 207. London: Home Office Research Development and Statistics Directorate.

Wilson, H. (1980). Parental Supervision re-examined. *British Journal of Criminology*, 27: 275-293.

Wilson, J.Q. & Kelling, G.L. (1982). Broken Windows: The police and neighbourhood safety. *Atlantic Monthly*. Mar, 249/3: 29-38.

Wilson, M. & Sapsford, R. (2006). *Asking questions*. In R. Sapsford & V. Jupp. Eds. *Data Collection and Analysis*. 2nd ed. London: Sage, pp93-123.

Wilson, W.J. (1987). *The truly disadvantaged*. Chicago, Il: University of Chicago Press.

Wilson, W.J. (1996). *When work disappears*. Chicago, Il: University of Chicago Press.

Wolfgang, M., Figlio, R.M. & Sellin, T. (1972). *Delinquency in a Birth Cohort*. Chicago, Il: University of Chicago Press.

Wortley, R. (2001). A classification of techniques for controlling situational precipitators of crime. *Security Journal*, 14/4: 63-82.

Xu, K. (2004). How has the Literature on Gini's Index Evolved in the Past 80 Years? Department of Economics, Dalhousie University retrieved (26 Jul 2016) from http://www.dal.ca/content/dam/dalhousie/pdf/faculty/science/economics/RePEc/dal/wparch/howgini.pdf (Chinese version of paper is (2003) China Economic Quarterly, 2: 757-78.)

Zahm, D. (2007). Using Crime Prevention Through Environmental Design in Problem-Solving. Problem-Oriented Guides for Police Problem-Solving Tools Series. Guide No. 8. Washington: US DoJ. Retrieved from http://www.popcenter.org/tools/pdfs/cpted.pdf

Zhu, L., Gorman, D.M. & Horel, S. (2004). Alcohol outlet density and violence: A geospatial analysis. *Alcohol and Alcoholism*, 39/4: 369-75.

Zipf, G.K. (1949). *Human Behaviour and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley.

Appendices

Appendix 1.1: Copy of information sheet, consent form, original semi-structured interview schedule

Information sheet

My name is Melanie Wellsmith and I am a PhD student at University College London and a senior lecturer in criminology at the University of Huddersfield, England. As part of my Doctoral research, I am carrying out research into high crime facilities within XX. I am doing secondary analysis of police recorded crime data, observations of high crime areas and interviewing police officers.

The purpose of the interview is to obtain data relating to the nature, incidence and placement of crime in XX, with respect to public facilities, such as stores and licensed premises. I am trying to establish the extent to which opportunity and location drives how much crime takes place and where and when it takes place.

I would like to invite you to participate in this research, which will involve me discussing with you for approximately 15-20 minutes, the areas of XX that you believe experience the highest concentrations of crime (you may think of these as ‘problem’ areas) and specific premises that you consider to be problematic.

The following contact details are provided should you have any further questions about this research:

Researcher: **Melanie Wellsmith LLB PGCert PGDip MSc FHEA,** Department of Behavioural and Social Sciences, University of Huddersfield, Queensgate, Huddersfield, UK, HD1 3DH. Tel: Email: Web: www2.hud.ac.uk/hhs/staff/shummw.php

Doctoral supervisor: **Dr Kate J. Bowers**, Reader, UCL Jill Dando Institute of Crime Science, University College London, Second Floor Brook House, 2-16 Torrington Place, London, UK, WC1E 7HN.
Tel: 0044 20 31083032 Email: k.bowers@ucl.ac.uk Web: www.jdi.ucl.ac.uk/

Consent to participate (respondent's copy)

Before participating in this study, I need to ensure you are aware of the following:

1) The purpose of the study (as above)
2) That the data gathered will be used as part of my doctoral studies and may appear in aggregated, analysed or individual form (including direct quotes) in my thesis
3) That your details will be kept confidential and known only to myself. If mentioned or quoted in the thesis you will referred to by a respondent number
4) That your individual responses will not be shared with **** PD
5) That aggregated data and the results of the analysis will be provided to **** PD
6) That you may choose to withdraw from the study at any time up until the point when the data has been included in the analysis. To withdraw you should contact me on the details above
7) That you may, without prejudice, refuse to answer any questions asked during the interview or to ask for the interview to be stopped at any time
8) That you can request a copy of your completed interview schedule/notes made (to be sent within 1 month of participating in the research)
9) That you can request an electronic copy of the thesis for which these data are being collected

Please sign, print and date to show you understand and agree to the above:

Signature...
Date..

Print your name...

I would also like to electronically voice record your responses to assist me in capturing the points you make. These recordings will be destroyed as soon as the thesis has been examined, will not be heard by anyone other than the researcher/transcriber and will not be included in the thesis. If you do not wish the interview to be recorded, this does **<u>not</u>** preclude you from participating in the research. Please sign again below to indicate whether or not you consent to the interview being recorded.

I do/do not agree [delete as applicable] to my responses being electronically recorded:

Signature...

Interview schedule

Respondent Information:

Rank:
Role (and beat if applicable):
Number of years service:
Previous roles/experience:

****

Maps and high crime areas

*1. The first thing I would like to ask you, is where *you* consider to be high-crime areas. These can be individual blocks, or larger areas, but could you please show me on the map, the areas you think are the most problematic in terms of crime and disorder:*

<Prompt for (i) generally (ii) acquisitive crime (iii) violent crime (iv) nuisance/disturbance (v) drugs (vi) gang-related activity>

<Can have multiple areas, but must number them as above>

*2. Now could you indicate the areas where *you* believe offenders tend to live:*

<Prompt for (i) general (ii) addicts (iii) gang members>

3. I am particularly interested in crime that occurs within or directly outside particular facilities, crime that can be associated with a particular premises or addresses. So, for example, I have looked at grocery, convenience stores and gas stations, while I've been here, but I am also interested in other types of premises; shops generally, licensed premises, restaurants and so on. What I am trying to find out is where these problem premises are and why you think that might be the case. We know that for each class of facility, there will be some addresses where there is a lot of crime and a lot where there is very little. Some of this will show in the crime figures, but of course not all crime is reported. The data cannot always explain why these places are problems either.

What I would like you to do is tell me where, from your experience, the problem premises are for the following types of facility and for different types of crime.

<Establish addresses if not already told. Use Google maps and street atlas if necessary>

*4. Looking at these premises you have highlighted as problems, could you tell me why *you* believe this is the case. What is it about these premises (as opposed to others you have not* *highlighted) that leads to them being problematic?*

<If necessary prompt with features of the location in which they are, the size of the premises, other facilities nearby, the way it is managed, the way it is laid out, the goods it has>

5. Is there anything else you would like to discuss with regards to crime problems in facilities?

Thank you. That completes the interview. Do you have any final points you wish to make, or any questions you would like to ask me?

Appendix 1.2: Evidence of ethical approval

Copy of confirmation that ethics approval not required from UCL Research Ethics Committee (from Head of Department)

Appendix 2: Studies identifying possible 'risky facilities'

Facility	Reference	Type	**Notes** (compiled from consideration of original source and further sources as cited)	Further sources
Apartments	Clarke & Bichler-Robertson (1998)	All crimes/ incidents	Evidence of risky facilities within property owned by one landlord (US)	Eck et al. (2007)
	Eck et al. (2007)	Violent and property crime	J-curves were present to varying degrees for different crime types across apartment complexes (US)	
Building sites	Clarke & Goldstein (2002)	Theft	Less than 4% of houses under construction accounted for all reported thefts (US)	Eck et al. (2007)
Business premises (general)	Burrows et al. (1999)	Violent and property crime	Survey in Scotland found J-curve concentrations of crime (UK)	Eck et al. (2007)
	Fisher & Looye (2000)	All crime	Random sample survey found just 12.5% of respondent businesses experienced any crime (US)	Eck et al. (2007)
	Gill (1998)	All crime	Various data sources relating to crime against businesses consistently evidenced J-curve concentrations (across different sources and crime types) (UK)	Eck et al. (2007)
	Hopkins & Ingram (2001)	All crime	Scottish Business Crime Survey found evidence of concentrations across different offences (UK)	Eck et al. (2007)
	Johnston et al. (1994)	Violent and property crime	Risk varied considerably across different industrial estates (UK)	Eck et al. (2007)
	Mirlees-Black & Ross (1995)	Violent and property crime	National survey found J-curves across different commercial premises, e.g. 8% of manufacturers experienced nearly three quarters of all crime (UK)	Eck et al. (2007)
	Perrone (2000)	All crime	National survey found 1% of small businesses experienced 66% of crime (AUS)	Eck et al. (2007)
	Taylor & Mayhew (2002)	All crime	National survey found very strong J-curve concentrations for small businesses across a range of different crimes (AUS)	Eck et al. (2007)

	Townsley et al. (2000)	Burglary	3% of business experienced 20% of recorded burglaries (AUS)	Eck et al. (2007)
	Walker (1996)	All crime	National survey found a quarter of businesses had been burgled (AUS)	Eck et al. (2007)
Car parks	Laycock & Austin (1992)	Vehicle crime	Five car parks accounted for half of recorded crime (UK)	Eck et al. (2007)
	Smith et al. (2003)	Vehicle crime	Evidence of J-curve concentrations for recorded offences, e.g. 10% of car parks experienced 35% of crime (UK)	Eck et al. (2007)
	Webb et al. (1992 as calculated by Eck et al. 2007)	Vehicle crime	Small number of car parks experienced higher rates of recorded crime (UK)	Eck et al. (2007)
Computers/ networks	Moitra & Konda (2004)	Various types of computer network attack	Small proportion of networks experienced much higher volume of crime	Townsley & Farrell (2007)
Country borders	Rossmo et al. (2008)	Illegal border crossings	Found J-curve in number of illegal entries into US per 'mile marker' for border area studied (US)	
Educational establishments	Bowers et al. (1998)	Burglary	Small proportion of schools had higher than average recorded revictimisation rates (UK)	Eck et al. (2007)
	Burquest et al. (1992 as calculated by Eck et al. 2007)	Burglary and criminal damage	18% of schools accounted for nearly half of recorded crime (UK)	Eck et al. (2007)
	Hope (1986)	Burglary	Within a sample of 59 schools chosen as likely to be most problematic, one quarter of establishments experienced one half of burglaries (UK)	
	Lindstrom (1997)	All crime	Evidence of J-curves for schools, concentrations varying across a range of crime types (SWE)	Eck et al. (2007)
	Snyder & Sickmund (1999)	Violent crime	Survey of school administrators showed relatively small proportion reported any incidents (US)	Eck et al. (2007)
Financial institutions (banks/building societies)	Austin (1988)	Burglary and robbery (inc. attempts)	Large sample study found only 5% of building societies experienced crime (UK)	Eck et al. (2007)

	Weisel (2007)	Robbery	Banks in urban areas experience a disproportionate amount of robberies (US, CAN, UK)	
	Matthews et al. (2001)	Robbery (inc. attempts)	Small proportion of banks experienced much higher rate of offending (UK)	Eck et al. (2007)
	Saylor & Janus (1981)	Robbery	12% of bank branches accounted for more than one-third of robberies (US)	Weisel (2007)
Healthcare	Bowers et al. (1998)	Burglary	Small proportion of facilities had higher than average recorded revictimisation rates (UK)	Eck et al. (2007)
Holiday accommodation	Eck et al. (2007)	All incidents?	J-curve concentrations of calls for service were found to varying degrees across different types of motel (US)	
	Oakland Police Department (2003)	All incidents	One hotel facility had much higher incident and arrest rate (US)	Eck et al. (2007)
	Mawby & Jones (2004)	Burglary other than dwelling	35% of offences occurred across only 3% of hotels studied (UK)	
On-licensed premises	Eck et al. (2007)	All incidents	Purposive data show a J-curve for bars, with 20% of premises accounting for 62% of calls for service (US)	
	Homel & Clark (1994)	Aggressive behaviour, violence (as observed by researchers)	Under one-third of sites accounted for 83% of observed incidents (AUS)	Eck et al. (2007)
	Newton & Hirschfield (2008)	Violent crime	Found evidence of J-curves across a number of towns, e.g. in two areas, three premises accounted for over 40% of crime (UK)	
	Ramsey (1986)	All crime	J-curve found for recorded crime in pubs and clubs (UK)	Eck et al. (2007)
	Sherman et al. (1992)	Violent crime	Around 15% of premises accounted for half of recorded incidents (US)	Eck et al. (2007)
	Sidebottom & Bowers (2010)	Bag theft	Five bars (of 26 studied) accounted for 59% of recorded offences. A J-curve was also present for rates of theft (per 100 seats), although premises were ordered slightly differently (UK)	

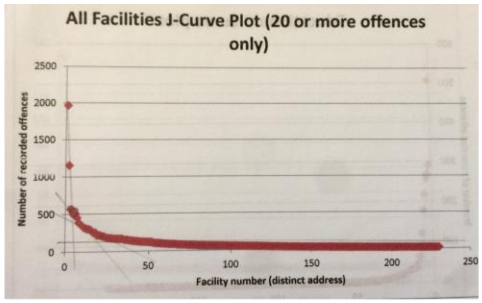
	Smith et al. (2006)	Bag theft	Found J-curve concentrations across small number of premises within a chain, as well as some establishments experiencing extremely high incidence rates compared to the average (UK)	
	Wellsmith et al. (2007)	Violent crime	Concentrations of recorded crime found across on-licensed premises, varying by town (UK)	
	Madensen & Eck (2008)	Violent crime (inc. threats)	20% of bars accounted for 75% of physical violence (US)	
Penal institutions	Clarke & Martin (1975)	Absconding	Small number of training schools had much higher than average absconding rates (UK)	Eck et al. (2007)
	Laycock (1977)	Absconding	Some borstals had much higher than average absconding rates (UK)	Eck et al. (2007)
Petrol (service) stations	Chakraborti et al. (2002 as calculated by Eck et al. 2007)	All crime	Evidence of J-curve concentrations across different reporting companies (UK)	Eck et al. (2007)
	La Vigne (1994)	All crime	J-curve concentrations in calls for service for range of incidents, e.g. 10% of gas stations accounted for 50% of calls for drive offs and drug crimes (US)	Eck et al. (2007)
	Taylor (2002)	Robbery	National survey and crime statistics show 37% of establishments accounted for 62% of robberies (AUS)	
Public transport	Newton (2004)	Vandalism	A quarter of bus shelters experienced 70% of incidents (UK)	Eck et al. (2007)
	Burrell (2007)	Violent crime	Some evidence of clustering of offences in a disproportionate number of transport hubs or routes.	
Retail (general)	Eck et al. (2007)	Shoplifting	Purposive data show a J-curve for stores, with just over 20% experiencing 85% of shoplifting reports (US)	
	Mirlees-Black & Ross (1995)	Violent and property crime	National survey found 3% of retailers experienced 59% of crime (UK)	Eck et al. (2007)
	Natarajan (2007)	Shoplifting	20% of stores contributed 85% of crime (US)	

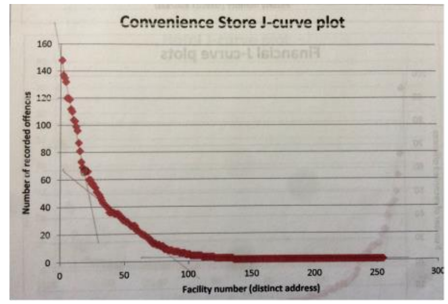
Retail (convenience stores)	National Association of Convenience Stores (1991)	Violent crime	Survey found less than 7% of stores experienced 65% of robberies (US)	Eck et al. (2007)
Retail (fast food/take-away)	Spelman (1995)	All crime	J-curve concentrations in police calls for service (US)	Eck et al. (2007)
Retail (pharmacies)	Taylor (2002)	Robbery	48% of pharmacies reported 64% of robberies (AUS)	
Sports facilities	Bowers et al. (1998)	Burglary	Almost one quarter of facilities had higher than average recorded revictimisation rates (UK)	Eck et al. (2007)
Telephone boxes	Hirschfield & Bowers (1998)	Hoax calls	J-curve concentrations were found for the origin of hoax calls to the police, e.g. 20% of calls came from 3% of telephones (UK)	Eck et al. (2007)

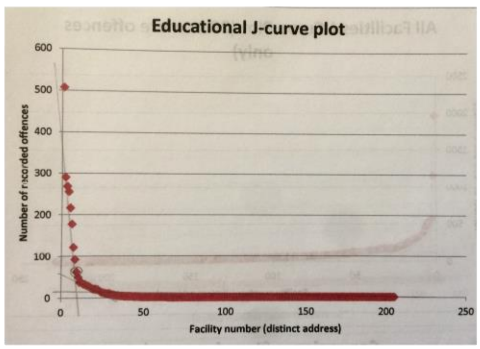
Appendix 3: Example R code

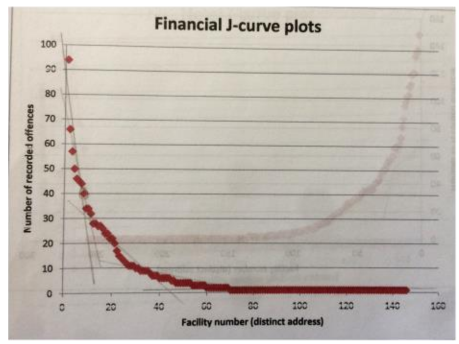
```
setwd ("G:/PhD from 160913/Analysis/RF distributions/PL test/DarkCounts")

read.csv("fin.csv", header=FALSE)

fin <- read.csv("fin.csv", header=FALSE)

names(fin) <- c("nobars", "incidents")

combo <- fin$incidents

plot (combo)

barplot (combo)

nocomb<- fin$nobars

## chooseCRANmirror()

##install.packages ("poweRlaw", dependencies = TRUE)

library("poweRlaw", lib.loc="C:/Users/Melanie/Documents/R/win-library/3.0")

m_pl = displ$new(combo)

est = estimate_xmin(m_pl)

m_pl$setXmin(est)

est

m_ln = dislnorm$new(combo)

est = estimate_xmin(m_ln)

m_ln$setXmin(est)

m_pois = dispois$new(combo)

m_pois$setXmin(est)

est

plot(m_pl)

lines(m_pl, col = 2)

lines(m_ln, col = 3)

lines (m_pois, col = 4)

bs_p = bootstrap_p(m_pl, no_of_sims = 5000, threads=1)

bs_p$p

plot(bs_p)

m_ln$setXmin(m_pl$getXmin())

est = estimate_pars(m_ln)

m_ln$setPars(est)

comp= compare_distributions(m_pl, m_ln)

comp
```

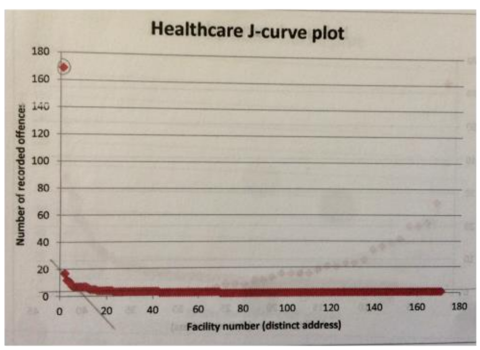
Appendix 4: Annotated J-curves (Method G)

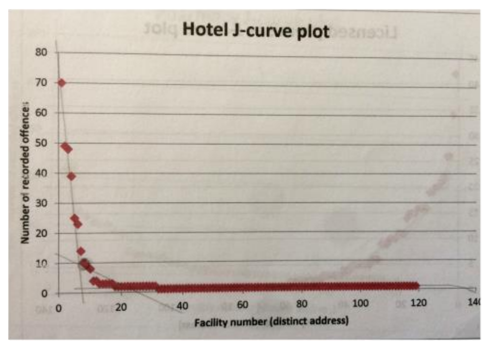


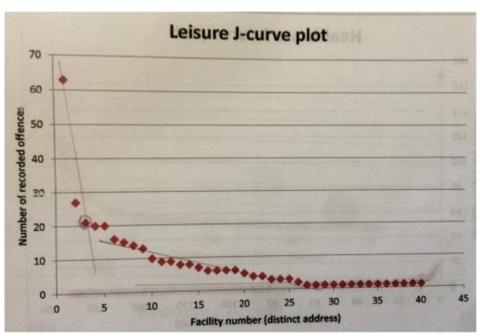


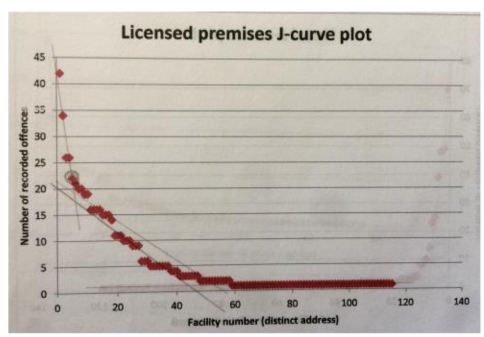


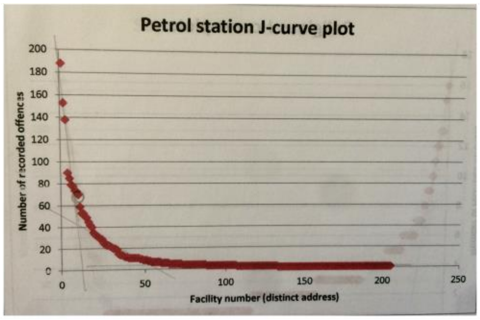


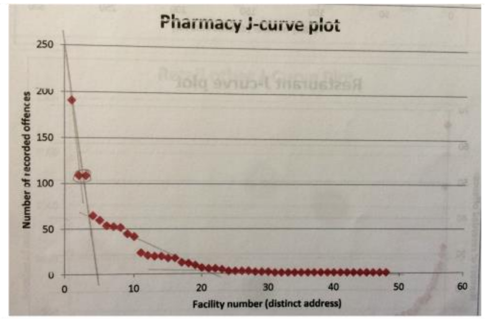


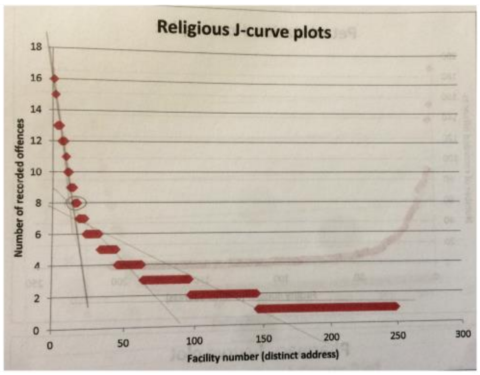


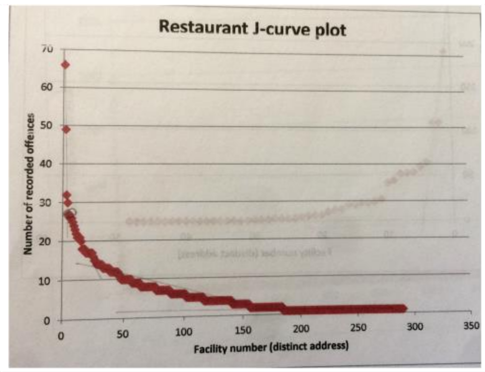


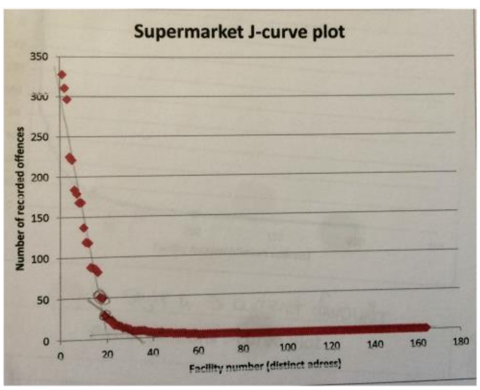


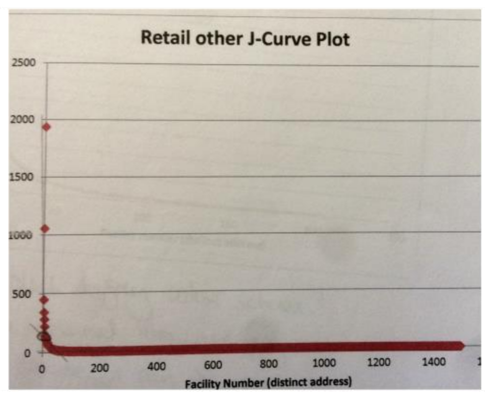

Petrol station J-curve plot
Number of recorded offences
Facility number (distinct address)


Pharmacy J-curve plot
Number of recorded offences
Facility number (distinct address)


Religious J-curve plots
Number of recorded offences
Facility number (distinct address)


Restaurant J-curve plot
Number of recorded offences
Facility number (distinct address)

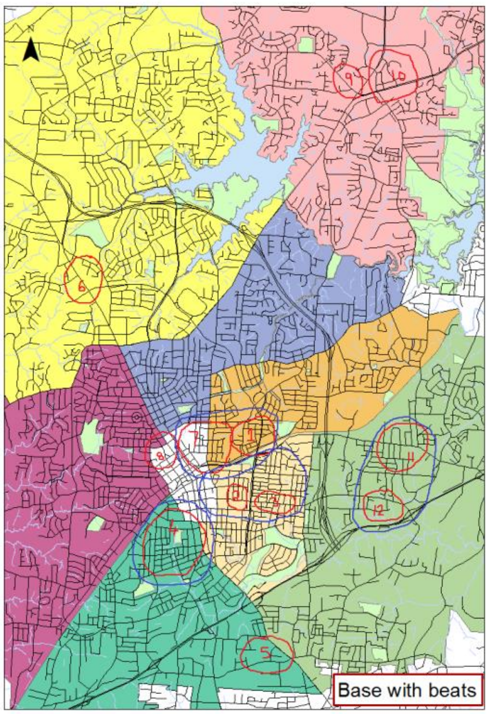

Supermarket J-curve plot
Number of recorded offences
Facility number (distinct adress)


Retail other J-Curve Plot
Facility Number (distinct address)

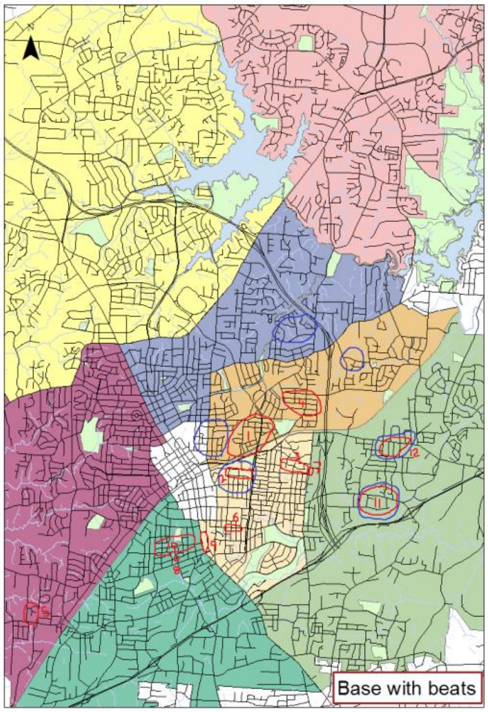
Appendix 5: Replicated police-annotated maps

These are digitally produced versions, drawn by the author as exact copies of the maps that were hand-annotated by respondents during interview. Red = high crime areas/premises, blue = areas of offender residences. Full-size reproductions are available on request.

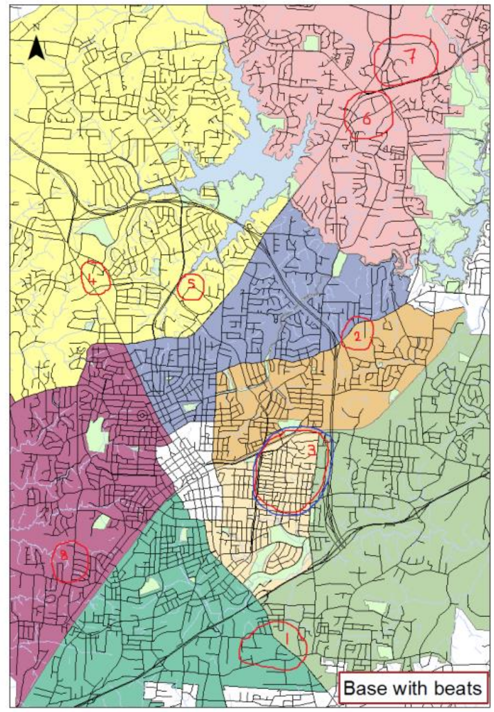
Respondent 1:



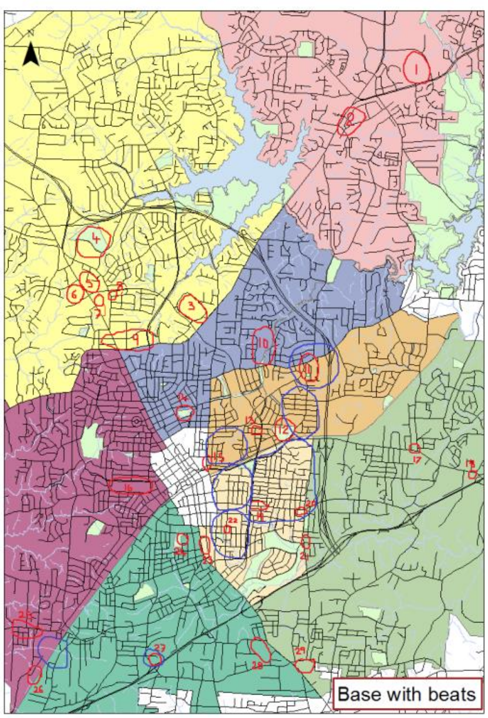
R2:


Base with beats

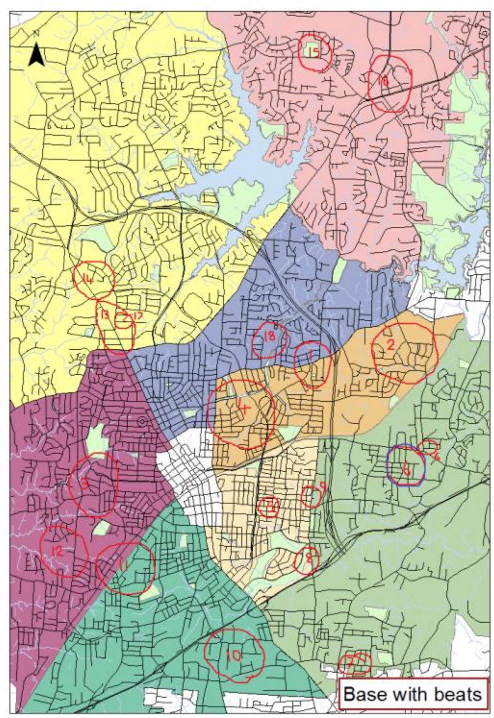
R3:


Base with beats

R4:


Base with beats

R5:


Base with beats

R6: